Universitat de Barcelona
Facultat de Matemàtiques i Informàtica

GRAU DE MATEMÀTIQUES

Treball final de grau

HISTÒRIA DE LA PROBABILITAT ALS SEGLES XIX i XX

Autor: Eduard Antentas Oliveras

Director: Dr. Josep Vives Santa Eulàlia
Realitzat a: Departament de Matemàtiques i Informàtica

Barcelona, 13 de juny de 2022

Abstract

The probability theory is the mathematical foundation to explain many aspects of our society and world: from demographic studies to ensembles of particles. And it was during the XIX and XX centuries that this mathematics branch developed most and became really important, thanks to renowned figures such as Laplace, Chebyshev, Kolmogórov and more; they unified the criteria and axioms that would be the basis of the modern statistics.

Resum

La teoria de la probabilitat conforma la base matemàtica per explicar molts fenòmens que regeixen la nostra societat actual i la naturalesa: des d'estudis demogràfics fins als conjunts de partícules. I és durant els segles XIX i XX que aquesta branca de les matemàtiques va esdevenir més important, gràcies a la contribució de figures com Laplace, Chebyshev, Kolmogórov, i molts d'altres; ells van unificar criteris i van crear l'axiomàtica que donaria lloc al desenvolupament de la teoria de la probabilitat actual, així com als inicis de l'estadística moderna.

2020 Mathematics Subject Classification. 60-00, 60-01, 60-02, 60-03, 60A05, 60C05, 60E05, 60E10, 60F05, 60F15, 60G15, 60G25, 60G50, 60G55, 60G57, 60J05, 60J10, 60J20, 60J35, 60J46, 60J50, 60J65

Agraïments

Vull agrair a la meva família pel seu suport incondicional durant aquests anys, així com als meus amics i companys de classe. És gràcies a tots ells que avui escric aquest projecte de final de grau. Vull mencionar especialment al meu tutor del projecte, el Dr. Josep Vives Santa Eulàlia, doncs m'ha donat molta confiança i suport durant el desenvolupament d'aquest treball. M'ha proporcionat les eines bibliogràfiques i tècniques necessàries per a tenir una base tangible sobre la qual construir aquesta memòria. Als meus pares i la meva germana: m'heu donat suport i animat en els meus daltabaixos, ajudant-me en tot el necessari perquè només m'hagués de preocupar per estudiar i aprendre. Al Jan i al Reig: el suport incondicional i ajuda desinteressada que m'heu donat no té preu, i és per això que sou el millor que m'emporto d'aquesta etapa universitària. A l'Àlex: per resoldre tots els dubtes que em sorgien, a qualsevol hora i qualsevol moment, amb una humilitat i paciència que fan pal·lès la teva bondat i excepcional intel·lecte. I a molts més companys, que no tinc espai per mencionar aquí, però són plenament conscients del meu afecte cap a ells i la importància que tenen en la meva vida. A la Universitat de Barcelona com a institució: per donar la oportunitat als joves com jo de formar-nos i créixer acadèmicament i personalment enmig d'una ciutat tan viva i pròspera com es Barcelona, disposant de tranquil·les i boniques facultats en el cor de la mateixa.

Índex

1 Introducció

1.1 El projecte

El prominent matemàtic i estadístic Karl Pearson (1857-1936), el 1878, exposava el següent: *"I do feel how wrongful it was to work for so many years at statistics and neglect its history"*. És evident la importància que té entendre el passat d'una ciència per entendre-la en la seva completesa, i les probabilitats i l'estadística no en són una excepció. Cada autor dels estudiats per fer aquest treball (Stigler [33], Gorroochurn [15], Hald [16], Sheynin [32], Maistrov [23]) tenen el seu punt de vista particular pel que fa a les etapes de la història de la probabilitat a partir del segle XIX, però tots ells tenen en comú un esquema bastant generalitzat:

1. El desenvolupament de la probabilitat com una disciplina matemàtica aplicada (de Laplace i Poisson a Chebyshev). Primera meitat del segle XIX principalment.
2. La demostració rigorosa dels teoremes de pas al límit (Teorema del límit central, Llei dels grans nombres, etc.) i la gradual transició de la probabilitat cap al regne de les matemàtiques pures, gràcies sobretot a l'escola russa de Sant Petersburg (Chebyshev, Màrkov, Liapunov). Segona meitat del segle XIX principalment.
3. Establiment d'una axiomatització i principals protagonistes d'aquesta, amb un paper destacat per part de Khintxin i Kolmogórov, que van donar nom a l'escola de probabilitats de Moscou. Principis del segle XX.

Per tant, aquest treball té l'objectiu de traçar una línia temporal al llarg del segle XIX i principis del XX, repassant els autors més importants en el desenvolupament de la teoria de la probabilitat, així com analitzant i comprenent l'evolució d'aquesta branca de les matemàtiques en aquest període històric. També estudia el naixement i desenvolupament de la física estadística durant la segona meitat del segle XIX (a l'annex), degut principalment a Boltzmann i Gibbs. Es presenten breument també les diverses aplicacions socials de la probabilitat: demografia, principis morals, legislació, assegurances, loteries, etc. En línies generals, el punt de partida és el revolucionari llibre de Laplace de 1812 titulat *Théorie Analytique des Probabilités* [19], i el punt final es troba en l'axiomatització de la teoria de la probabilitat per part de Kolmogórov, amb el també revolucionari llibre *Foundations of the Theory of Probability* [17] de 1933. El fil conductor d'aquest treball de fi de grau (tfg) són diversos llibres sobre la història de la probabilitat i la estadística (detallats a la bibliografia), escrits per autors contemporanis de renom com Stigler [33], Gorroochurn [15], Hald [16], Sheynin [32], Maistrov [23]; cal remarcar que cadascun té un enfocament i perspectiva diferent a nivell històric i també matemàtic. Per aquesta raó, s'ha intentat recopilar informació a parts iguals i aportar una perspectiva objectiva i neutral, considerant que l'autor en aquest cas és un estudiant de l'últim curs en matemàtiques i en física. Aquests llibres d'història de les matemàtiques s'han complementat amb alguns articles originals dels propis científics estudiats (Laplace, Màrkov, Chebyshev, Kolmogórov, etc.).

1.2 Estructura de la Memòria

La memòria s'ha estructurat en tres parts cronològicament ordenades: primera meitat del segle XIX, segona meitat del segle XIX, i principis del segle XX. I dins de cadascuna d'elles, s'han desenvolupat els diversos matemàtics més rellevants de cada època, incloent les seves contribucions en la teoria de la probabilitat i exposant alguns dels seus teoremes, axiomàtica i demostracions en general. Pel que fa a la part més formal i matemàtica, s'ha intentat mantenir l'essència conceptual dels articles originals usant terminologia i notació més moderna. D'aquesta manera, esdevé més fàcil per al lector entendre les idees i demostracions, i a la vegada comparar autors per a fer-se una idea general del context i desenvolupament històric d'aquesta branca de les matemàtiques als segles XIX i XX. Així, la memòria s'estructura al voltant dels protagonistes del desenvolupament de la teoria de la probabilitat, establint d'aquesta manera un fil temporal ordenat i coherent, mentre que tots els conceptes teòrics i cos matemàtic són introduïts adequadament. Finalment, s'ha afegit un annex sobre el naixement i desenvolupament de la física estadística durant la segona meitat del segle XIX.

2 Primera meitat del segle XIX

2.1 Pierre-Simon Laplace

La teoria de la probabilitat i els mètodes estadístics van néixer com resposta a les necessitats pràctiques socials i científiques de l'època. Laplace va predir l'aplicabilitat general d'aquests mètodes estadístics, mantenint així la teoria de la probabilitat com una eina important per augmentar tot tipus de coneixements humans. Abans d'entrar en matèria, però, cal exposar breument la seva trajectòria vital i professional.

Pierre-Simon Laplace (Normandia, 1749- París, 1827) va ser un astrònom, físic i matemàtic francès. Seguint la línia de la mecànica Newtoniana, va desenvolupar la famosa transformada que porta el seu nom, així com una equació de càlcul vectorial. Com a astrònom, va exposar la teoria nebular del sistema solar; i com a estadístic va establir les bases de la teoria analítica de la probabilitat. A nivell filosòfic i moral, va ser un ferm defensor del determinisme científic. Va néixer en el si d'una família rural de Normandia, i va estudiar a la Universitat de Caen, on ja va destacar notablement en matemàtiques. D'aquesta manera, el 1767 va ser anomenat professor a l'Escola Militar de Paris, on entre els seus alumnes hi consta Napoleó Bonaparte. El 1785, Laplace va esdevenir membre de l'Acadèmia de Ciències, així com al 1795, membre de la càtedra de matemàtiques del Nou Institut de les Ciències i les Arts, que de fet presidiria el 1812. El 1795, Laplace va començar a publicar els llibres que constituirien la seva obra mestra: *Traité de mécanique céleste*. També va tenir posicions polítiques influents, com per exemple la de Ministre d'Interior francès el 1799. A nivell ideològic, sempre va ser monàrquic, però es va saber adaptar molt als corrents polítics de cada moment. Per exemple, durant el mandat de Napoleó Bonaparte, Laplace es va posicionar des d'un primer moment al seu costat, rebent així la Legió d'Honor (1805) i el títol de comte de l'Imperi (1806).

El 1812, Laplace va publicar la *Théorie analytique des probabilités* [19], en la qual estableix molts resultats fonamentals de l'estadística. També és destacable *Essai philosophique sur les probabilités* [20] que va escriure el 1814. El 1816 va ser escollit membre de l'Acadèmia Francesa, i el 1817 va rebre el títol de marquès, doncs amb la restauració dels Borbons va ser prou hàbil com per posicionar-se al seu costat.

2.1.1 *Théorie analytique des probabilités*

En aquest volum publicat el 1812, Laplace presenta i recopila tots els seus resultats fonamentals de teoria de la probabilitat. Suposa un punt d'inflexió en l'estudi d'aquesta teoria, i per això s'ha escollit com a punt de partida d'aquest treball de fi de grau. Va normalitzar resultats previs d'altres autors, millorant les demostracions, donant peu a l'estudi de regularitats estadístiques, i també aplicant la probabilitat a estimacions d'errors i observacions, entre d'altres aspectes. La primera meitat d'aquest llibre tracta mètodes i problemes probabilístics, mentre que la segona es basa en mètodes estadístics i aplicacions d'aquests. Cal mencionar que algunes de les demostracions de Laplace no són considerades gaire rigoroses d'acord amb els estàndards actuals, i que són de lectura difícil per la complexitat tècnica usada, però això no treu importància a les conclusions a les que va arribar aquest il·lustre matemàtic francès. A continuació es fa un breu repàs del contingut d'alguns capítols de la obra, així com s'exposen alguns dels exemples que hi apareixen.

Al primer capítol, Laplace introdueix la clàssica definició de probabilitat (introduïda prèviament per De Moivre): la probabilitat $P(A)$ d'un esdeveniment A és igual al quocient del nombre de possibles resultats d'un experiment favorables a un esdeveniment A entre el nombre total de resultats possibles de l'experiment en qüestió. Es suposa que tots els possibles resultats són equiprobables. Laplace introdueix un significat subjectiu a aquesta definició a través del principi d'insuficiència. Segons aquest, s'assigna un cert valor (raonable) a la probabilitat desconeguda d'un esdeveniment. Conseqüentment, si un sistema està compost per diversos esdeveniments, amb les probabilitats de cadascun d'ells desconegudes, s'assumeix que tots ells són equiprobables. Per exemple, si una moneda és asimètrica, però es desconeix en quin dels dos costats està afavorida, es suposarà que en cada tirada la probabilitat d'obtenir cara és igualment una meitat. Per tant, degut a la ignorància respecte el costat favorable, la probabilitat pot augmentar i decréixer en mateixa mesura.

En aquest mateix capítol, Laplace exposa teoremes d'addició i multiplicació d'esdeveniments independents. En aquest sentit, afirma que si dos esdeveniments són independents, la probabilitat de la seva existència combinada és igual a producte de les respectives probabilitats. Per exemple, si la probabilitat d'obtenir un nombre en una tirada d'un dau és $\frac{1}{6}$, per tant la probabilitat d'obtenir els mateixos dos resultats tirant dos daus a la vegada és de $\frac{1}{36}$.

També s'exposen teoremes de probabilitat condicional. En efecte, si considerem coneguda *a priori* la probabilitat d'un esdeveniment i també la d'un succés compost d'aquest primer i un altre d'esperat, llavors la segona probabilitat dividida entre la primera és igual a la probabilitat del succés esperat. Per tant, aquí Laplace inclou el conegut teorema de Bayes:

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{P(B)}$$

on $P(A_i)$ són les probabilitats *a priori*, $P(B|A_i)$ la probabilitat de B condicionada a A_i, i $P(A_i|B)$ les probabilitats *a posteriori*.

Al segon capítol, Laplace resol diversos problemes, entre els quals trobem el següent:

Considerem l'extracció de tiquets d'una urna, on cadascun d'ells està numerat de 0 a n. Volem determinar la probabilitat que la suma dels k nombres extrets sigui igual a s. És a dir que

$$t_1 + t_2 + ... + t_k = s \tag{2.1}$$

amb la condició que $t_i \leq n \ \forall i$, i que

$$P(t_i) = \frac{1 - l^{n+1}}{n+1} \qquad l = \begin{cases} 0 \ per \ t_i \leq n \\ 1 \ per \ t_i > n \end{cases}$$

D'aquesta forma, Laplace calcula el nombre de combinacions possibles que compleixen la igualtat (2.1). També va considerar el cas límit de s, quan $n \to \infty$.

Ara, si considerem el terme $(x + x^2 + x^3 + x^4 + ...)^n$, llavors se sap que un cop fet el desenvolupament, el coeficient del terme x^s coincideix amb el nombre de vegades que, en tirar n daus, obtenim una puntuació igual a s. Laplace aconsegueix generalitzar aquest càlcul establint el mètode de les funcions generatrius, el qual avui en dia s'usa en gran mesura i en molts àmbits de les matemàtiques. Estableix que la funció generatriu de la seqüència $f_0, f_1, ..., f_n, ...$ ve donada per l'expressió

$$f(t) = \sum_{n=0}^{\infty} f_n t^n$$

Partint de la hipòtesi que aquesta sèrie de potències convergeix almenys per una $t \neq 0$.

Per altra banda, al llarg del seu tractat Laplace resol molts problemes probabilístics diferents, entre els quals trobem el problema de l'agulla de Buffon. Seguidament se'n presenten alguns de destacables.

- Volem trobar el nombre m de participacions en la loteria francesa de manera que la probabilitat que hi apareguin tots els 90 nombres sigui igual a una meitat. Laplace obté com a resultat que $m = 85,53$. I també obté una generalització del problema. Això és, si la loteria consta de s nombres, i en cada participació n'apareixen n, troba la probabilitat que apareixin tots s nombres en m participacions.

- Suposem que fem $m + n$ observacions, en les quals l'esdeveniment A es dóna m vegades, mentre que el B succeeix n cops, amb $m > n$. Laplace es pregunta quina és la probabilitat que $P(A) > P(B)$, i obté com a resposta que és

$$p = \frac{(\mu - m)(\mu - m + 1)\cdots\mu}{m!}\cdot$$
$$\cdot\left[\frac{(\frac{1}{2})^{\mu}}{\mu} + \frac{m\frac{(}{1}2)^{\mu-1}}{\mu(\mu-1)} + \frac{m(m-1)\frac{(}{1}2)^{\mu-2}}{\mu(\mu-1)(\mu-2)} + \cdots + \frac{m(m-1)\cdots 1(\frac{1}{2})^{\mu-m}}{\mu(\mu-1)\cdots(\mu-m+1)(\mu-m)}\right]$$

  on $\mu = m + n + 1$.

- Ara, Laplace es planteja trobar la probabilitat P que un esdeveniment observat m vegades seguides es repeteixi k vegades més, i obté que la resposta és

$$P = \frac{m+1}{m+k+1}$$

  .

- També s'analitza el problema de repartiment d'apostes (el clàssic problema dels punts): dos jugadors A i B juguen n partides, i el joc s'interromp quan A n'ha guanyat x i B n'ha guanyat y. Llavors, les apostes es reparteixen proporcionalment a la probabilitat que té cada jugador de guanyar les n partides. I és aquesta probabilitat la que Laplace obté, com a funció de $V = n - x$ i $W = n - y$, essent per al jugador A:

$$P = \left(\frac{1}{2}\right)^{V}\left[1 + V\frac{1}{2} + \frac{V(V+1)}{12}\left(\frac{1}{2}\right)^{2} + \cdots + \frac{V(V+1)\cdots(V+W-2)}{(W-1)!}\left(\frac{1}{2}\right)^{W-1}\right]$$

  I per al jugador B, la probabilitat de guanyar és $1 - P$.

- Laplace es proposa també trobar la millor combinació d'observacions per determinar una quantitat desconeguda, partint de la hipòtesi que el nombre d'observacions és indefinit. D'aquesta forma, conclou que la millor combinació s'obté amb el mètode dels mínims quadrats.

2.1.2 Teorema de Moivre-Laplace

Aquest és el teorema central i més important del llibre *Théorie analytique des probabilités* [19] de Laplace, entorn d'aquest resultat es desenvolupen i es deriven tots els seus treballs en probabilitat. S'obté a partir d'una nova demostració i reformulació del reconegut teorema de James Bernouilli. Laplace aconsegueix provar un dels teoremes del pas al límit més importants en la probabilitat, convertint-se així en una de les figures més influents en aquesta branca de les matemàtiques. Aquest teorema és un cas particular del teorema de límit central. En llenguatge actual:

Teorema 2.1. *La funció de probabilitat de successos amb probabilitat p, en un conjunt de n esdeveniments, convergeix a la funció de probabilitat d'una distribució normal amb mitjana np i desviació estàndard $\sqrt{np(1-p)}$.*

En la versió de Laplace al 1812:

Teorema 2.2. *Sigui p $(0 < p < 1)$ la probabilitat d'ocurrència d'un esdeveniment E en n experiments independents, i sigui m el nombre de vegades que ocorre E; llavors la probabilitat de la desigualtat*

$$z_1 < \frac{m-np}{(npq)^{1/2}} < z_2 \qquad (q = 1-p)$$

és arbitràriament molt propera a

$$\frac{1}{\sqrt{2\pi}} \int_{z_1}^{z_2} e^{\frac{-z^2}{2}} dz$$

La versió local d'aquest teorema s'enuncia com segueix:

Teorema 2.3. *La probabilitat de m esdeveniments E en un total de n experiments és igual a*

$$\frac{1}{\sqrt{2\pi npq}} e^{\frac{-(m-np)^2}{2npq}}$$

2.1.3 Teorema del Límit Central

Aquesta és la generalització del teorema de Moivre-Laplace, i de fet Laplace va ser el primer en obtenir-ne una demostració, l'any 1810. Més tard, el 1812, en la *Théorie analytique des probabilités* [19] el prova primer per una distribució uniforme, simètrica i discreta; després una de simètrica i contínua, i finalment una d'arbitrària. El teorema del límit central (abreviat TLC) es pot enunciar de la següent forma:

Teorema 2.4. *Siguin $x_1, \ldots, x_n$ variables independents i idènticament distribuïdes, amb funció de densitat $f(x)$, mitjana μ, i variància σ^2. Llavors, $s_n = x_1 + \ldots + x_n$ és una distribució normal amb mitjana $n\mu$ i variància $n\sigma^2$*

Per una distribució uniforme i discreta, Laplace considera la variable discreta x que pren valors enters entre $-a$ i a, amb funció característica

$$\psi(t) = (2a+1)^{-1} \frac{\sin\left[\frac{1}{2}(2a+1)t\right]}{\sin\left(\frac{t}{2}\right)}$$

De manera que

$$\psi(t) = E(e^{itx}) = \int_{-\infty}^{\infty} e^{itx} f(x) dx$$

on E denota esperança matemàtica, i f és la funció de densitat de la variable x.

En aquests termes, Laplace presenta la seva fórmula d'inversió per una distribució discreta simètrica:

$$c_s = \frac{1}{\pi}\int_0^{\pi} \psi^n(t) \cos stdt \tag{2.2}$$

Per una distribució uniforme i contínua, Laplace pren la funció característica

$$\psi(t) = (2a)^{-1}\left(\cos at + \frac{\cos\frac{t}{2}\sin at}{\sin\frac{t}{2}}\right)$$

Pel que fa a la demostració formal del TLC, la trobem en un article de Laplace del 1810, i també en els capítols 18 i 22 de la *Théorie analytique des probabilités* [19], segons el tipus de distribució usada. A continuació provem el TLC pel cas discret.

Demostració. Sigui x una variable aleatòria discreta, que pren valors enters $-a, -a+1, \ldots, a'-1, a'$, i té densitat de probabilitat $f\left(\frac{x}{a+a'}\right)$, de manera que

$$\sum_{x=-a}^{a'} f\left(\frac{x}{a+a'}\right) = 1$$

Anomenem als moments $\mu_d = E(x^d)$, $d = 0, 1, \ldots$ i a la variància $\sigma^2 = \mu_2 - \mu_1^2$. Ara, Laplace desenvolupa la funció característica com segueix

$$\psi(t) = E(e^{itx}) = 1 + i\mu_1 t - \frac{1}{2}\mu_2 t^2 + \cdots$$

i prenent logaritmes neperians

$$ln\ \psi(t) = i\mu_1 t - \frac{1}{2}\sigma^2 t^2 + \cdots$$

Ara, partint de la fórmula d'inversió (2.2) Laplace obté que la probabilitat de la suma de les n variables és

$$P(s_n = s+\mu) = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-it(s+\mu)}\psi^n(t)dt = \frac{1}{2\pi}\int_{-\pi}^{\pi} exp\left(-its + it(n\mu_1 - \mu) - \frac{1}{2}n\sigma^2 t^2 + \cdots\right)dt$$

Considerant que s és de l'ordre de $\sqrt{n}$ i definint $\mu = n\mu_1$, queda que

$$P(s = s_n - n\mu_1) = \frac{1}{2\pi}\int_{-\pi}^{\pi} exp\left(-its - \frac{1}{2}n\sigma^2 t^2 + \cdots\right)dt =$$

$$\frac{1}{2\pi\sigma\sqrt{n}} exp\left(\frac{-s^2}{2n\sigma 2}\right)\int_{-\pi\sigma\sqrt{n}}^{\pi\sigma\sqrt{n}} exp\left[-\frac{1}{2}\left(z + \frac{is}{\sigma\sqrt{n}}\right)^2\right]dz = \frac{1}{2\pi\sigma\sqrt{n}} exp\left(\frac{-s^2}{2n\sigma 2}\right)$$

Això implica que $s_n - n\mu_1$ segueix una distribució normal $(0, n\sigma^2)$=(mitjana, variància). A més, aquest resultat és vàlid per qualsevol distribució discreta, sempre i quan els moments existeixin. $\square$

Pel cas continu, Laplace considera la variable aleatòria $y = \frac{x}{a+a'}$, obtenint així la distribució de $\sum y_i$ a partir de la distribució de s_n. Amb un procediment similar a l'anterior demostració, Laplace prova que $\sum_{i=1}^{n}[y_i - E(y)]$ segueix una distribució normal $(0, \frac{n\sigma^2}{(a+a')^2})$.

Laplace aplica aquest teorema del límit central a distribucions de sumes de variables aleatòries, simplificant molt els càlculs. D'aquesta manera, s'estableix un nou mètode per aproximar grans mostrejos de dades a una distribució concreta (distribució normal), que avui dia encara s'usa molt freqüentment.

2.1.4 Doctrina filosòfica, principis morals i visió social

Les principals idees i metodologies de Laplace són presentades en el seu *Essai philosophique sur les probabilités* [20]. Segons el seu punt de vista, a la naturalesa domina la necessitat per sobre de l'atzar, i ho expressa en aquest llibre de la següent forma:

"All events, even those which on account of their insignificance do not seem to follow the great laws of nature, are a result of it just as necessarily as the revolutions of the sun" [[20], p. 37].

Laplace era un ferm defensor del determinisme filosòfic, molt comú en aquella època, i que afirma que tot esdeveniment físic i humà està determinat abans que succeeixi per la relació causa-conseqüència. Partint d'aquesta doctrina, Laplace descriu un estat "ideal" de la ciència, i ho expressa molt clarament en el següent paràgraf:

"Given for one instant an intelligence which could comprehend all the forces by which nature is animated and the respective situation of the being who compose it - an intelligence sufficiently vast to submit these data to analysis - it would embrace in the same formula the movements of the greatest bodies of the universe and those of the lightest atom; for it, nothing would be uncertain and the future, as the past, would be present to its eyes" [[20], p. 4].

En eliminar l'atzar, Laplace es veu forçat a definir la probabilitat d'un esdeveniment sense fer referència explícitament a aquest concepte, argumentant que la probabilitat és relativa, degut en part a la nostra ignorància i en part al nostre coneixement. És a dir, si no coneixem del tot un succés, diem que és probable. De la mateixa manera, defineix l'equiprobabilitat de dos successos partint de la nostra falta de coneixement respecte aquests i les seves causes. Laplace defensa que la teoria de la probabilitat s'ha d'aplicar a les ciències de la naturalesa i també a la societat, argumentant que la predicció que proporciona ens permet evitar els canvis sobtats deguts a la nostra ignorància com a humans.

Amb això, Laplace aplica molts dels seus resultats de teoria de probabilitat a aspectes quotidians, considerant que els principals problemes socials són de naturalesa probabilística. Tanmateix, mai justifica aquestes aplicacions socials, i és per això que és criticat per molts d'altres matemàtics més moderns en aquest sentit.

Laplace creia que els fenòmens que observem i la veritable naturalesa de les coses no coincideixen, que els nostres sentits provoquen aquesta distorsió fenomenològica, i que la ciència té la missió de corregir-ho, comparant i experimentant constantment.

Finalment, també sostenia que la llei de la gravitació universal explica tots els fenòmens naturals que ocorren dins els límits del sistema solar.

2.2 Robert Adrain

Robert Adrain (Irlanda, 1775- New Brunswick, 1843) va ser un matemàtic irlandès exiliat gran part de la seva vida als Estats Units. Va deixar Irlanda després de liderar els insurgents republicans en la rebel·lió armada contra les tropes britàniques de 1798, anant a viure a Nova Jersey, i més tard a Pennsylvania. Allà, va exercir com a professor en diverses universitats, com Princeton, Columbia o Pensilvania. També va escriure en diverses revistes i publicacions de matemàtiques. Entre els seus treballs més importants, trobem la formulació del mètode dels mínims quadrats, aplicat en demostracions de la llei de distribució d'errors aleatoris. Aquesta llei la va obtenir com a generalització d'un problema particular que estava resolent, i ho va fer de manera simultània i independent a l'obtenció de la mateixa llei de distribució per part de Gauss (anomenada també distribució normal). Adrain també va aplicar les matemàtiques a camps com la física, l'astronomia i la geografia. La majoria dels seus treballs es centraven en la forma de la Terra, i és per això que se'l considera un matemàtic aplicat. El 1812 va ser escollit membre de la *American Philosophical Society*, i un any després de la *American Academy of Arts and Sciences.*

2.2.1 *The Analyst, or Mathematical Museum*

Analitzem l'article de Adrain de 1808 titulat *The Analyst, or Mathematical Museum* [1] [Vol. 1, pp. 93-109], doncs és aquí on obté (en dos apartats) la llei de distribució normal per errors aleatoris de diverses observacions. Adrain es proposa resoldre el següent problema: Sigui AB el valor real d'una quantitat, per exemple una distància, la mesura física de la qual és Ab amb un error bB. Quina és la probabilitat que cometem l'error bB en mesurar el segment AB?

Per respondre-la, Adrain pren AB, BC, ... successives distàncies amb mesures respectives Ab, bc, ..., i un error total Cc. Assumeix com a obvi que els errors en les mesures de AB i BC són proporcionals a les llargàries de cada segment, això és Ab i bc. Denotant $Ab = a$, $bc = b$, $Cc = c$, i els errors de les mesures Ab i bc com x i y, respectivament, obtenim que l'equació que dona la màxima probabilitat d'obtenir aquests errors és $\frac{x}{a} = \frac{y}{b}$. Anomenem X i Y les probabilitats de cometre els errors x i y en les distàncies mesurades a i b. Clarament, la probabilitat d'obtenir els dos errors mútuament és XY. Volem doncs obtenir X i Y amb la condició que XY sigui màxim, i per fer-ho prenem les funcions $f(x) = \ln X$, $g(y) = \ln Y$. Per tant el màxim de XY es correspon amb el màxim de $f(x) + g(y)$. Derivem i queda

$$f'(x)x' + g'(y)y' = 0 \qquad f'(x)x' = -g'(y)y'$$

2.2.1 En la probabilitat màxima d'obtenir els errors x i y, clarament $x + y = const.$, d'on $x' + y' = 0$ i $x' = -y'$. Introduint aquesta condició a l'equació 2.2.1 veiem que $f'(x) = g'(y)$. Aquesta equació ha de ser equivalent a $\frac{x}{a} = \frac{y}{b}$, la condició inicial de màxima probabilitat. Això dona lloc a l'expressió

$$f'(x) = \frac{mx}{a} \qquad i \qquad g'(y) = \frac{my}{b}$$

Considerant la primera equació, integrem $f'(x)$ i queda

$$\ln X = \frac{mx^2}{2a} + k; \qquad X = e^{\left(\frac{mx^2}{2a}+k\right)}$$

La funció $u = e^{\left(\frac{mx^2}{2a}+k\right)}$ és l'equació general de la corba de probabilitat [p. 94], i ens serveix per a comparar les probabilitats dels errors comesos en cada mesura.

Ara, s'exposa breument la segona obtenció de la llei de distribució normal en aquest article d'Adrain. Pren de nou la mesura d'un segment AB, amb la mateixa probabilitat d'error en la mesura en direcció horitzontal i azimutal. Considera que aquesta probabilitat es troba definint un cercle al voltant del punt B. Amb això, fa els càlculs corresponents i obté que les probabilitats dels errors en les mesures són

$$e^{\left(\frac{nx^2}{2}+k\right)} \qquad i \qquad e^{\left(\frac{ny^2}{2}+k\right)}$$

on x i y són els errors en les mesures horitzontal i azimutal, respectivament.

Per tant, aquest article de 1808 d'Adrain obté per primera vegada la llei de distribució normal d'errors aleatoris. En aquella mateixa època, Gauss també l'obté, i de fet és qui més reconeixement històric rep i a qui se li atribueix en gran mesura aquesta fita matemàtica.

2.3 Johann Carl Friedrich Gauss

Johann Carl Friedrich Gauss (1777-1855) va ser un matemàtic i físic alemany, que va fer grans contribucions en molts àmbits de la ciència. És considerat un dels matemàtics més importants de la història. Nascut a Brunswick, Alemanya, en el si d'una família rural, va destacar des de ben petit en les matemàtiques, en particular en l'aritmètica, esdevenint així un nen prodigi. Amb tant sols 21 anys, ja havia completat una de les seves obres mestres, *Disquisitiones arithmeticae*, que constitueix la consolidació de la teoria de nombres que avui dia encara s'estudia extensivament.

Gauss va ser el primer en demostrar rigorosament el teorema fonamental de l'àlgebra, en la seva tesi doctoral del 1799. El 1809 va ser nombrat director de l'observatori de Göttingen, on va fer importants treballs en el camp de l'astronomia, entre els quals es troba un mètode per a calcular l'òrbita d'un planeta i refinar-la posteriorment. Aquest es troba en la seva obra, del mateix any 1809, titulada *Theoria motus corporum coelestium in sectionibus conicis Solem ambientium* [10].

El 1835, va formular el conegut Teorema de la divergència de Gauss, una de les seves contribucions més importants a l'electromagnetisme, i de la qual deriven dues de les quatre Equacions de Maxwell. Gauss defensava un mètode de treball molt rigorós, obtenint tots els resultats i desenvolupaments matemàtics a partir de només alguns axiomes. La seva metodologia i estil van tenir un gran impacte en les matemàtiques, que ha arribat fins als nostres dies.

2.3.1 *Theoria motus corporum coelestium in sectionibus conicis Solem ambientium*

És en aquest article, publicat el 1809 [10], on apareixen els primers treballs de Gauss en teoria de la probabilitat. En la part final, apareix la seva llei dels errors obtinguts en les observacions: la famosa llei normal de distribució d'errors o distribució Gaussiana.

Aquest tractament probabilístic d'errors deriva dels seus treballs en astronomia i geodèsia, doncs aquests li van crear la necessitat de desenvolupar mètodes per a tractar els resultats observacionals que obtenia. En astronomia i geodèsia, es fan nombroses observacions en diverses ubicacions i usant diferents instruments, cosa que dóna lloc a molt d'error en determinar el valor observat de manera més exacta. Aquesta problemàtica va ser el detonant per a Gauss per desenvolupar la teoria d'errors, així com per confirmar el mètode dels mínims quadrats, prenent com a axioma el principi de la mitjana aritmètica.

L'interès científic per a minimitzar els errors d'una mesura fent moltes observacions d'aquesta ja venia d'uns segles abans: Tycho Brahe (1546-1601), a finals del segle XVI, ja havia usat aquesta tècnica.

Per altra banda, el mètode dels mínims quadrats ja havia estat desenvolupat i presentat l'any 1806 per Legendre (1752-1833), el qual nota que pot ser molt útil en aplicacions observacionals de la física i l'astronomia. De fet, Gauss fa menció al treball previ de Legendre quan ell mateix exposa aquest mètode en l'article de 1809.

En aquest article, Gauss es planteja el següent problema: donades diverses mesures equidistants d'una mateixa quantitat, cadascuna d'elles amb un error Δ, i amb densitat de probabilitat $\varphi(\Delta)$, volem trobar la forma de la funció $\varphi(\Delta)$, amb la hipòtesi que el valor més probable de la quantitat mesurada és el que correspon a la mitjana aritmètica de les observacions preses. Equivalentment, que φ té un màxim per error nul, i.e. $\Delta = 0$, i s'anul·la pel valor de Δ més gran. En paraules del propi Gauss:

$\varphi(\Delta)$, therefore, would appropiately be referred to the class of discontinuous functions, and if we undertake to substitute any analytical function in the place of it for practical purposes, this must be of such a form that it may converge to zero on both sides, asymptotically, as it were, from $\Delta = 0$. [[13], p. 93]

D'aquesta manera, Gauss arriba a la conclusió que la funció buscada és

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

on h s'anomena "la mesura de precisió de les observacions". Clarament, aquesta és la llei normal de distribució d'errors.

Gauss extreu un corol·lari interessant d'aquest resultat: la densitat de probabilitat d'una mostra poblacional és màxima quan la suma de quadrats de les diferències entre els valors observats i els reals es minimitza. És així com introdueix el principi dels mínims quadrats. En conseqüència, Gauss afirma que aquest principi s'ha de prendre com a axioma, i que la mitjana aritmètica de les observacions s'ha de considerar com el valor més probable de la quantitat mesurada.

Tornant a la llei normal de distribució d'errors, Gauss s'adona que ha comès un lleu error en plantejar-la: els errors sempre estan acotats en uns certs límits, i per tant la probabilitat de que els excedeixin hauria de ser zero, però la forma de $\varphi(\Delta)$ trobada sempre dona algun resultat. Per tant, aquesta admet errors de qualsevol magnitud, i això no hauria de ser així. De fet, això implicava que totes les observacions es podien considerar en l'anàlisi dels errors, per molts grans que fossin aquests. De retruc, fins a mitjans del segle XIX no es va establir un criteri per a descartar observacions en funció de la magnitud dels errors, considerant-se fins a aquell moment la llei normal de Gauss com un principi universal.

2.3.2 *Bestimmung der Genauigkeit der Beobachtungen*

Aquest article de Gauss, publicat el 1816 i traduït com *Determination of the Accuracy of the Observations* [11], esdevé la seva segona obra més important en teoria de la probabilitat. Aquí, Gauss estima qualitativament el valor h de la funció $\varphi(\Delta)$ trobada en l'article anterior.

Primer introdueix la funció d'error

$$\theta(t) = \frac{2}{\sqrt{\pi}} \int_0^t e^{-t^2} dt$$

remarcant el valor de ρ tal que $\theta(\rho) = 0.5$ en una taula amb diversos valors de la funció en qüestió. També anomena $r = \frac{\rho}{h}$ com "l'error probable de $\theta(ht)$".

A continuació, Gauss fa el següent plantejament amb l'objectiu de trobar el valor de h:

"Assume now that in performing m observations the errors $\alpha, \beta, \gamma, \delta, \ldots$ occurred and let us investigate what conclusions can be drawn concerning the values of h and r" [65, p. 122].

Gauss deriva la següent funció de densitat per les observacions experimentals fetes (notem que la notació usada és diferent a la de l'article original, per conveniència):

$$f = K_n h^n e^{-h^2(\alpha^2+\beta^2+\gamma^2+\delta^2+\ldots)}$$

amb K_n una constant. A continuació, deriva la funció f respecte h per a obtenir el valor de h que la maximitza:

$$f' = K_n[nh^{n-1}e^{-h^2(\alpha^2+\beta^2+\gamma^2+\delta^2+\ldots)} + h^n e^{-h^2(\alpha^2+\beta^2+\gamma^2+\delta^2+\ldots)}(-2h)(\alpha^2+\beta^2+\gamma^2+\delta^2+\ldots)] = K_n h^{n-1} e^{-h^2(\alpha^2+\beta^2+\gamma^2+\delta^2+\ldots)}[n - 2h^2(\alpha^2+\beta^2+\gamma^2+\delta^2+\ldots)]$$

Imposant que $f' = 0$ per la condició de màxim, obtenim que

$$h = \sqrt{\frac{n}{2(\alpha^2+\beta^2+\gamma^2+\delta^2+\ldots)}}$$

Conseqüentment, obté que

$$r = \frac{\rho}{h} = \rho\sqrt{\frac{2(\alpha^2+\beta^2+\gamma^2+\delta^2+\ldots)}{n}}$$

que és el valor més probable de r.

2.3.3 *Theoria combinationis observationum erroribus minimis obnoxiae*

Aquesta obra, publicada el 1821 i traduïda com *Theory of the Combination of Observations Least Subject to Errors* [12], és la tercera i segurament la més completa de Gauss en la teoria d'errors. De fet, aquí classifica els tipus d'errors en aleatoris i sistemàtics; els primers no segueixen cap regularitat, i els segons sí, podent-se així predir en certa forma.

Gauss estudia a consciència les lleis que regeixen la distribució aleatòria d'errors, partint de la funció de densitat d'errors $\varphi(x)$. La suposa simètrica, i.e. $\varphi(-x) = \varphi(x)$, amb un màxim a $x = 0$ i amb $\int_{-\infty}^{\infty} \varphi(x)dx = 1$.

Seguidament exposem el problema que considera Gauss:

Prenem les variables $y, x_1, x_2, \ldots, x_\sigma$ linealment dependents, és a dir tals que

$$y = \sum_{s=1}^{\sigma} a_s x_s$$

on les a_s són incògnites. Per a determinar-les, parteix dels valors y_r obtinguts experimentalment, tals que

$$y_r = \sum_{s=1}^{\sigma} a_s x_{sr} \qquad r = 1, 2, \ldots, N$$

Com que els valors y_r tenen associat un cert error experimental, en fer les mesures realment obtenim $\eta_r = y_r + \Delta_r$. D'aquesta manera, coneixent les variables x_{sr} i els valors obtinguts η_r, es vol trobar els valors aproximats α_s de les incògnites a_s. Per a fer-ho, Gauss parteix del mètode dels mínims quadrats, és a dir, imposa que s'ha de complir la condició següent

$$\sum_{r=1}^{N} \left(\eta_r - \sum_{s=1}^{\sigma} \alpha_s x_{sr} \right)^2 = min \tag{2.3}$$

D'aquesta forma, obté un sistema d'equacions a partir del qual pot trobar les α_s. Aquest sistema és:

$$\sum_{s=1}^{\sigma} \alpha_s \sum_{r=1}^{N} x_{sr} x_{ir} = \sum_{r=1}^{N} \eta_r x_{ir} \qquad i = 1, 2, \ldots, \sigma$$

Aleshores, l'equació 2.3 es minimitza, de fet s'anul·la quan el sistema d'equacions queda de la següent forma

$$\begin{aligned} \alpha_1 x_{11} + \alpha 2 x_{21} + \ldots + \alpha_\sigma x_{\sigma 1} &= \eta_1, \\ \alpha_1 x_{12} + \alpha 2 x_{22} + \ldots + \alpha_\sigma x_{\sigma 2} &= \eta_2, \\ \alpha_1 x_{1N} + \alpha 2 x_{2N} + \ldots + \alpha_\sigma x &= \eta_N \end{aligned}$$

I aquest és per tant el sistema que li permet trobar a Gauss els valors aproximats α_s de les incògnites a_s, amb la particularitat que no tenen biaix, i per tant l'esperança matemàtica $E(\alpha_s) = a_s$.

I fins aquí arriben les contribucions més destacables de Gauss en la teoria de la probabilitat.

2.4 Siméon Denis Poisson

Siméon Denis Poisson (1781-1840) va ser un matemàtic i físic francès, que va fer importants treballs en estadística, anàlisi complexa, equacions diferencials parcials, electricitat i magnetisme, termodinàmica, i més. Poisson va néixer al poble de Pithiviers, en el districte francès de Loiret, i el seu pare era oficial de l'armada. El 1798 va entrar a l'escola politècnica a Paris, on va destacar ja des del primer moment. El 1800, ja va publicar dues memòries que li van suposar el reconeixement d'il·lustres figures com Legendre o Laplace, que de fet esdevindrien mentors i amics del jove Poisson. Durant els anys posteriors, va esdevenir professor a l'escola politècnica a París, i el 1808 va anar a la institució *Bureau* *des Longitudes* com a astrònom. Un any més tard, va ser nomenat professor de mecànica a la facultat de ciència de Paris, i el 1812 examinador a l'escola militar de *Saint-Cyr*. Durant tota la seva vida va seguir lligat a aquestes institucions acadèmiques d'alguna forma, obtenint notables resultats i obres científiques. Com a reconeixement, el 1818 va ser escollit com a membre de la *Royal Society* de Londres, el 1822 membre honorari de la *American Academy of Arts and Sciences*, i el 1823 membre de la *Royal Swedish Academy of Sciences*.

Entre els seus articles més destacables en teoria de probabilitats, trobem *On the probability of mean results of observations* (1827),*Continuation of the memoir on mean results of observations* (1832), *Sur l'avantage du Banquier au jeu de Pharaon*, *On the probability of birth of boys and girls*. Tots ells s'inclouen en l'obra cabdal de Poisson en probabilitat: *Recherches sur la probabilité des jugements en matière criminelle et en matière civile*, (*Researches into the Probabilities of Judgements in Criminal and Civil Cases*) [31] publicat a Paris l'any 1837. És de fet aquí on s'hi presenta el famós teorema de Poisson.

2.4.1 *Researches into the Probabilities of Judgements in Criminal and Civil Cases*

En els dos primers capítols, Poisson introdueix conceptes elementals de probabilitats, com serien les variables aleatòries discretes, així com l'atzar i les probabilitats, les quals tracta de forma bastant subjectiva. En aquest sentit, també s'estudien les causes d'esdeveniments i la probabilitat que succeeixin. Cal destacar que un dels capítols consisteix exclusivament en les aplicacions dels resultats de probabilitat en les decisions dels tribunals jurídics, i és que Poisson defensava fermament que la teoria de la probabilitat tenia moltes aplicacions socials, en particular per a millorar la presa de decisions en els tribunals. Vol determinar la probabilitat de que es cometin errors en les sentències judicials, i per això estudia teoremes de pas al límit, arribant d'aquesta forma a la famosa llei dels grans nombres:

Es fan n experiments independents, que resulten en la ocurrència o no ocurrència de l'esdeveniment A. Suposa que en cada experiment la probabilitat d'ocurrència de l'esdeveniment A no és igual. Aleshores, si diem $\frac{m}{n}$ el quocient entre el nombre d'experiments on ocorre A i el nombre d'experiments totals, aquest quocient és arbitràriament pròxim a la mitjana aritmètica $\bar{p_n}$ de la probabilitat que ocorri A en un experiment independent. En notació actual, aquesta llei es representa amb el següent límit:

$$\lim_{n \to +\infty} \left(\left| \frac{m}{n} - \bar{p_n} \right| < \epsilon \right)$$

Aquest teorema es pot particularitzar al teorema de Bernouilli, en el cas que la probabilitat de l'esdeveniment A no varií en cada experiment, i.e. $\bar{p_n} = p_n = p$.

En aquest mateix tractat, Poisson també obté la "llei dels petits nombres", com un cas particular de l'anterior. Ve motivada pel fet que, quan p_n s'allunya del valor mig $1/2$ ($p_n \in [0,1]$), la representació asimptòtica de la probabilitat de m/n, $P_{m,n}$, que té la forma $(2\pi)^{-\frac{1}{2}} e^{-\frac{x^2}{2}}$, esdevé molt menys precisa. Per tant, Poisson troba una aproximació pel límit de p molt petita, això és, $p_n \to 0$, per $n \to \infty$ (molts experiments). En aquest límit, la probabilitat que A ocorri m vegades és igual a

$$P_{m,n} = \frac{\lambda^m}{m!} e^{-\lambda}$$

on $\lambda = np_n$.

Més endavant, aquesta llei de Poisson va ser usada a nivell pràctic per calcular la probabilitat d'esdeveniments poc comuns, com per exemple les morts degut a la cossa d'un cavall en l'armada prussiana, el naixement de trigèmins, i d'altres.

Poisson creia que tots els successos de la naturalesa, així com de caire moral i humà, s'havien d'explicar amb la seva llei dels grans nombres. Era partidari d'aplicar-la als tribunals de justícia, i de fet va analitzar moltes decisions de tribunals per arribar a calcular la probabilitat exacta de que un ciutadà fos acusat o absolt, independentment del context i la subjectivitat humana. Aquesta visió li va provocar moltes crítiques per una part de la comunitat matemàtica de l'època, que consideraven que s'estava malmetent la reputació de la ciència matemàtica amb aquelles aplicacions. Aquesta crítica es va fer extensiva als resultats i la perspectiva de Laplace. De fet, tot això va provocar que en molts països occidentals s'estanquessin en els avenços de la teoria de la probabilitat i les seves aplicacions durant els següents anys, donant lloc a l'establiment i domini de l'escola russa a l'est, que tractarem a continuació.

Abans, emperò, fem una breu menció al famós estadístic, matemàtic i sociòleg belga, Adolphe Quételet (1796-1874).

2.5 Adolphe Quételet (1796-1874)

Nascut a Ghant, en aquell moment part de França, l'any 1796, Quételet va estudiar a la universitat d'aquesta mateixa ciutat, obtenint el doctorat en matemàtiques l'any 1819. El 1820 va esdevenir membre de la *Royal Academy*, i el 1839 de la *American Philosophical Society*, així com el 1850 va entrar a formar part de la *Royal Swedish Academy of Sciences*. El 1828 va fundar l'observatori de Brusel·les. Durant tota la seva carrera va fundar diverses revistes i societats d'estadística, amb l'objectiu sobretot d'aconseguir cooperació internacional entre estadístics.

En estadística, Quételet és famós per la introducció de mètodes a nivell social, partint de la teoria (injustificada i sense fonament teòric) de que la societat i el comportament humà es regeix per les lleis de la teoria de la probabilitat.

El resultat d'aquesta teoria s'anomena l'*average man* de Quételet. A partir d'una gran quantitat de dades estadístiques, obté aquest prototip ideal d'humà perfecte, que és etern i immutable. Està caracteritzat per la mitjana de moltes variables mesurades. Aquesta visió determinista li va causar també moltes crítiques i controvèrsia en la comunitat científica, sobretot per la poca justificació formal que aportava.

En resum, en aquella època era comú trobar publicacions científiques com la de Quételet, sobre aplicacions socials injustificades de la teoria de la probabilitat. Això va provocar que, a mitjan del segle XIX, aquesta teoria topés amb un carreró sense sortida, doncs no es podia establir formalment una relació entre les matemàtiques aplicades a la ciència i la teoria de la probabilitat; no estava clar en quines àrees es podia aplicar aquesta última. I per a desencallar aquesta situació, va ser necessari el desenvolupament i aportacions en probabilitat de l'escola russa (focalitzada a St. Petersburg), que presentem a continuació.

2.6 Teoria de la probabilitat a Rússia

2.6.1 *Vilnius University*

És en aquesta universitat, actualment situada a Lituània, i en aquell moment part de l'Imperi Rus, on es va ensenyar per primera vegada probabilitats a nivell acadèmic a Rússia. I va ser gràcies al professor de matemàtiques I.A. Snyadezkii, qui el 1808 va proposar la idea de complementar el desenvolupament matemàtic de la universitat amb els fonaments de la teoria de la probabilitat, sobre la qual l'any 1817 ja va impartir una xerrada. L'oficina central de la *Vilnius University* expressava la necessitat d'adoptar aquesta teoria en el seu pla d'estudis en una carta al ministeri d'educació, exemplificant la importància que té en la societat a través de resultats com els obtinguts per Laplace. A més, en aquesta mateixa carta recomanen al professor Sygmund Revkovskii com a candidat per impartir les classes, degut a la seva destacable trajectòria estudiant probabilitats. Així, el curs 1829-1830 Revkovskii va impartir per primera vegada a Rússia classes de probabilitat, i en la presentació de l'assignatura fa una definició interessant de la teoria de la probabilitat, la qual dóna els mètodes per a calcular l'"esperança", la mesura de la qual defineix com la probabilitat. El programa de l'assignatura inclou: probabilitats simples, compostes, variables discretes i contínues, la llei de Bernouilli, probabilitats futures, i aplicacions en naturalesa, filosofia i sociologia. En aquest sentit, es van estudiar aplicacions en la geodèsia, moralitat i política, així com taules de mortalitat.

Vilnius va ser doncs la primera universitat en impartir probabilitats a Rússia, l'any 1830. Aquell mateix any, el reconegut matemàtic Mikhaïl Ostrogradski (1801-1861) s'adreçava a l'Acadèmia de Ciències demanant que l'estudi de la teoria de la probabilitat es fes extensiu a totes les universitats i instituts de l'Imperi Rus, argumentant el benefici que podria aportar a les generacions futures del país.

2.6.2 Universitat de Moscou

Tot i els esforços de la comunitat científica, la introducció de les probabilitats en les institucions acadèmiques russes va ser lenta, i fins el 1850 no es va començar a impartir classes d'aquesta branca matemàtica a la Universitat de Moscou, dirigides pel matemàtic i enginyer rus August Davidov (1823-1885). El pla d'estudis de la primera assignatura impartida per Davidov, el curs 1851-1852, incloïa els següents temes: definició de probabilitats *a priori* (probabilitat condicional, teorema de Bernouilli, esperança matemàtica, etc.), definició de probabilitats *a posteriori* (probabilitat d'esdeveniments futurs, de les causes d'aquests, etc.), aplicacions de la teoria de la probabilitat en l'estadística (taules de mortalitat, longevitat, etc.), i probabilitats en les assegurances.

Davidov va fer importants treballs en probabilitats entre els anys 1854 i 1857, i cal destacar el seu article titulat *An application of probability theory to medicine*, on aplica mètodes estadístics per a determinar símptomes de malalties i altres problemes de salut. Per exemple, pren mostres de 200 casos d'una malaltia, on un símptoma en concret apareix 130 vegades. Aleshores, Davidov es pregunta si aquest símptoma es pot considerar una característica pròpia d'aquesta malaltia. Per respondre aquesta pregunta elabora una taula, a partir de la qual obté que la probabilitat de repetició d'aquest símptoma es troba entre el límit inferior $\frac{56}{100}$ i el límit superior $\frac{74}{100}$. Com que ambdós límits es troben per sobre de $\frac{1}{2}$, Davidov conclou que efectivament aquest símptoma en concret és característic d'aquesta malaltia.

Abans d'aquests treballs de Davidov, emperò, ja s'havien publicat les primeres obres en teoria de probabilitats a la Universitat de Moscou. Els responsables d'aquestes van ser els matemàtics russos Nikolai Brashman (1796-1886) i Nikolai Zernov (1804-1862). El primer va ser mentor del mateix Davidov, i va escriure el tractat *Solutions of problems in the calculus of probabilities* (1835). Pel que fa a Zernov, va ser el primer professor rus en obtenir un doctorat en matemàtiques, i el 1843 va impartir una important xerrada a la Universitat de Moscou sobre *The theory of probability with special applications to mortality and insurance.* De fet, més endavant dedicarem un subapartat sencer a la figura de Zernov i la seva obra.

2.6.3 Universitat de Sant Petersburg

Aquesta és la universitat on apareixeran les figures més importants de l'escola russa, sobretot durant la segona meitat del segle XIX (Chebyshev, Màrkov, Liapunov, etc.). Durant la primera meitat d'aquest segle, entre els seus predecessors trobem el matemàtic V.A. Ankudovich, que va impartir el primer curs en probabilitats a aquesta institució. Després d'ell, aquestes classes van ser impartides durant la dècada del 1850 per Víktor Buniakovski (1804-1889), reconegut matemàtic rus, sobretot per la conjectura que porta el seu nom.

A continuació, desenvolupem alguns dels matemàtics més importants i influents en el desenvolupament de la teoria de la probabilitat a Rússia durant la primera meitat del segle XIX. Ells serien els predecessors de la famosa escola russa, també anomenada escola de Sant Petersburg (en referència a la universitat que va liderar i centralitzar els treballs en probabilitats aquella època).

2.6.4 Nikolai Lobatxevski

Nikolai Lobatxevski (1792-1856) va ser un matemàtic i geòmetra rus, autor de les primeres obres importants en teoria de la probabilitat en aquell imperi. Va néixer a Nizhny Novgorod, ciutat de l'imperi rus, i quan tenia set anys el seu pare va morir, motiu pel qual la família va anar a viure a Kazan, la capital de la República de Tatarstan (Rússia). Allà, Lobatxevski va estudiar a la Universitat de Kazan, graduant-se en matemàtiques i física l'any 1811. El 1822, va esdevenir professor, i el 1827 rector d'aquesta mateixa universitat. Tota la seva vida hi va estar vinculat, excel·lint sobretot en el desenvolupament d'una geometria no euclidiana, que de fet porta el seu nom.

Són de fet els seus treballs en geometria que el van portar a interessar-se per les probabilitats i teoria d'errors. Lobatxevski era partidari del l'experimentació i treball empíric per a entendre les propietats de l'entorn i adquirir nocions espacials adequades. Per exemple, als seus articles hi trobem el càlcul de la (extremadament) petita desviació angular respecte 180º, valor resultant de la suma dels angles en el triangle format per la Terra, el Sol i l'estrella Sirius. Com que aquesta desviació és tant petita, va despertar en Lobatxevski l'interés per l'estimació d'errors en les observacions, per així obtenir-les amb la màxima exactitud possible. En aquest sentit, en el seu famós tractat *New elements of geometry with a complete theory of parallels* (1835-38) [22], expressa la importància d'introduir la teoria d'errors en càlculs usant dades de triangles amb un cert grau d'exactitud, i ho fa d'aquesta forma:

"Thus, the solution will be fully adequate if and only if in addition to the precision of

the calculations we also present the probability of the occurrences of the errors" [[22] pp. 397-398]

Per això, en aquest mateix tractat soluciona dos problemes amb l'objectiu de determinar la llei de distribució d'una suma de variables aleatòries independents idènticament distribuïdes. El primer problema és el següent:

Volem trobar la distribució de la suma $\mu_r = \lambda_1 + \lambda_2 + \ldots + \lambda_r$, on λ_i és una variable independent i $P(\lambda_i \in [-a, a]) = \frac{1}{2a+1}$. A més, μ_r és un enter m tal que $m \in [-ra, ra]$. Amb això, Lobatxevski calcula la possibilitat de que $\mu_r = m$, amb m fixat (notem que fa servir lletres de l'alfabet rus per a denotar la probabilitat, per tant en comptes de P usa la C):

$$C_r(m) = \sum (-1)^{\lambda} C_r^{\lambda} C_{(r-2\lambda)a+m+r-1-\lambda}^{r-1}$$

on el sumatori és sobre $\lambda \in [0, \frac{ra+m+1}{2a+1})$.

Pel que fa al segon problema considerat: volem obtenir la distribució de la mitjana aritmètica $\xi_r = (\delta_1 + \delta_2 + \ldots + \delta_r)/r$, on δ_i és una variable independent i uniformament distribuïda en l'interval $[-1, 1]$. Lobatxevski troba una equació per la probabilitat que aquest mitjana sigui menor o igual a un valor determinat, això és, $P(|\xi_r| \leq x)$:

$$C_r(x) = \left(1 - \frac{1}{r!2^{r-1}}\right) \sum_{0<\lambda<\frac{r-rx}{2}} (-1)^{\lambda} C_r^{\lambda} (r - rx - 2\lambda)^r$$

Per tant, tot i que Lobatxevski no va centrar-se gairebé mai en l'estudi de les probabilitats, els seus resultats pel que fa a formules de distribució de sumes de variables independents van ser molt importants per al desenvolupament de la teoria d'errors, i fins i tot avui en dia encara s'usen.

2.6.5 Mikhaïl Ostrogradski

Mikhaïl Ostrogradski (1801-1862) va ser un matemàtic, físic i mecànic de l'Imperi Rus, d'ascendència ucraïnesa. Va néixer al poble de Pashennaya (actualment situat a la província de Poltava, Ucraïna). Va estudiar a la universitat de Kharkov, de 1816 a 1820, per després estudiar a la Universitat de Paris de 1822 a 1826. Posteriorment, l'any 1828 va tornar a l'Imperi Rus, establint-se a Sant Petersburg, i esdevenint membre de l'Acadèmia de Ciències de Rússia. Els seus principals treballs es troben en l'àmbit de les matemàtiques aplicades. És destacable el fet que, l'any 1826, va demostrar per primera vegada el famós Teorema de la Divergència, que de fet s'anomena també Teorema de Gauss-Ostrogradski.

Pel que fa a les seves contribucions a la teoria de la probabilitat, es centren sobretot en aplicacions pràctiques d'aquesta: pensions, judicis, apostes, assegurançes, etc.

Pel que fa a les pensions, Ostrogradski va rebre l'encàrrec de calcular i establir els fons d'inversió/pensions vitalícies destinats als acomiadaments de milers de militars de la marina russa, després de la pèrdua de la guerra de Crimea l'any 1856, i conseqüentment la pèrdua del dret d'operar al Mar Negre per part de Rússia. I així ho va fer, calculant i presentant la pensió mínima adequada de forma comprensible per al públic en general. Es pot trobar en el llibre *A proposal for establishing a Retirement Pension Fund in the Navy Department* (1868) [29].

Ostrogradski era partidari de la visió de Laplace respecte les probabilitats, considerant que aquestes s'apliquen a temes morals, socials i de visió filosòficament determinista. Per això, Ostrogradski publica articles on resol problemes probabilístics en el marc legal i jurídic. Un d'aquests el trobem en el seu article titulat *Extract from a memoir on the probability of judicial errors* (1838) [29]. Aquí, considera un tribunal amb diversos jutges, cadascun dels quals té una opinió lleugerament diferent. Llavors, Ostrogradski obté les equacions que donen la probabilitat de que el veredicte del tribunal sigui erroni, donat el nombre de jutges i coneixent els límits de la veracitat de cadascun d'ells. Entre les conclusions que aquest matemàtic rus extreu, destaquem la següent: si la probabilitat de que cada jutge prengui la decisió correcta és la mateixa, llavors la probabilitat d'un veredicte erroni només depèn de la majoria de vots, i no del nombre total de jutges.

Per altra banda, Ostrogradski resol un problema aplicable al control de qualitat de mostres amb molts objectes, usant un paral·lelisme amb una capsa plena de pilotes blanques i negres. Ho fa a l'article *On a problem concerning probabilities* (1848) [29]. En aquest, considera el següent problema amb la capsa plena de boles blanques i negres:

"An urn contains white and black balls, whose total number is known, but the number of balls of each color is unknown. A certain number of balls is drawn from the urn and, after counting the number of white and black balls among them, the balls are returned to the urn. It is required to determine the probability that the total number of white balls is contained within specified limits." [[29] p.215]

Llavors, exposa una aplicació pràctica d'aquest problema de la següent forma:

"To understand the importance of this problem, let us visualize the situation that a person is to receive a large number of objects subject to certain requirements and a certain amount of time is required to check these requirements. Military suppliers often face this problem. The balls in the urn represent the objects supplied, the white balls may, for example, correspond to the acceptable ones, satisfying certain requirements, and the black corresoponding to the unacceptable ones. The draw of a certain number of objects to determine their colors corresponds to checking a part of the received order to determine their quality. This part is five, six, or seven percent of the total number of objects taken at random. After drawing this part and computing the number of acceptable in it, we find the probability that the total number of the acceptable objects will stay within the limits set up in advance. These calculations are performed in the same manner as the calculations of white and black balls in a urn. By a suitable choice of the limits as well as of the number of objects to be checked, the required probability may approach certainty as closely as desired. Hence, if we solve the problem stated above, a supplier may utilize it to reduce almost twentifold the very tiresome mechanical labor, as for example in the case of sacks of flour or bolts of cloth." [[29] p. 215]

Per tant, per a resoldre aquest problema pràctic, Ostrogradski parteix d'una capsa amb un nombre determinat de boles blanques i negres, amb proporció desconeguda. Ara, prenem una mostra de l boles de la capsa, volem trobar la probabilitat que d'aquestes n'hi hagi n blanques i m negres. Ostrogradski la calcula de la següent forma:

"If l balls are drawn, there can be $0, 1, 2, 3, \ldots, n, \ldots, l-1, l$ white balls and, hence correspondingly $l, l-1, l-2, l-3, \ldots, m, \ldots, 1, 0$ black balls. Since all these different hypotheses, whose total number is $l+1$, are equally probable, the probability of each one of them, including the one we have in mind, is $\frac{1}{l+1}$." [[29] p.216]

Per tant, conclou que la probabilitat de que hi hagi n boles blanques i m boles negres entre les l boles extretes és igual a $\frac{1}{l+1}$. Procedeix llavors a trobar fórmules més genèriques, introduint la notació

$$[z]^k = z(z-1)(z-2)\dots(z-k+1)$$

Aleshores, calcula la mateixa probabilitat que abans però coneixent la distribució prèvia de boles blanques i negres a la caixa:

"Let there be x white balls and y black balls in the urn. Find the probability that among the l drawn balls there will be n white and m black ones." [[29] p.216]

Aquesta probabilitat resulta ser

$$\frac{[l]^l[x]^n[y]^m}{[n]^n[m]^m[s]^l}$$

essent s el nombre total de boles a la capsa.

En aquesta mateixa línia, es planteja ara que de les n boles extretes n'obtenim n blanques i m negres. Llavors, calcula la probabilitat que de les $s-l$ boles restants n'hi hagi x blanques i y negres, i el resultat és

$$\frac{[l+1]^{l+1}[x]^n[y]^m}{[n]^n[m]^m[s+1]^{l+1}}$$

en notació combinatòria actual s'expressa com

$$\frac{C_x^n C_y^m}{C_{s+1}^{l+1}}$$

Pel que fa a les aplicacions probabilístiques en assegurances, Ostrogradski publica l'any 1847 un article titulat *On insurance* [29]. En aquest, remarca que els estudis d'assegurances tenen el seu fonament en la teoria de la probabilitat. Per això, el primer que fa en aquest tractat és introduir la definició de probabilitat, que en la seva opinió és subjectiva a cada observador; està relacionada amb el grau d'ignorància de l'esdeveniment per part de l'observador. Ho expressa així:

"There is no probability in Nature. Everything that occurs in the universe is inevitable and beyond doubt. Probability is a consequence of human weakness; it relates to us, exists for us and can be only for us. A consideration of probability is an important and even necessary addition to those few truths which we know with relative certainty." [[29] p.240]

A nivell formal, defineix la probabilitat de forma clàssica (com ho feia Laplace), és a dir com el quocient de nombre de successos favorables entre nombre de successos totals. La seva visió determinista de la naturalesa s'exemplifica clarament amb el següent exemple referent a la probabilitat de que la bola extreta d'una capsa amb 5 boles (3 blanques i 2 negres) sigui blanca:

"There is no reason to assume that one ball will fall in our hands faster than another. By saying "no reason", we mean that we do not have a reason - the reason, however, exists, but is unknown to us...
[...]
We repeat that probability and equal possibility of the outcomes, as well as the measure of probability, exist only in our minds. For those who have total knowledge of all the phenomena, probability not only has no measure, but simply does not make sense." [[29] p.241]

Finalment, Ostrogradski estudia alguns problemes directament relacionats amb assegurances i loteries, usant l'esperança matemàtica. Després de resoldre numèricament alguns exemples pràctics, conclou que l'asseguradora mai perd diners i per tant el que li convé més al ciutadà corrent és pagar el mínim possible en assegurances. És també molt crític amb les empreses que organitzen les loteries, argumentant que són les úniques que treuen profit d'aquestes.

Tot i que és cert que Ostrogradski va cometre l'error de caure en el subjectivisme filosòfic i en les aplicacions injustificades de les probabilitats en alguns àmbits socials, això no treu que les seves aportacions van ser de cabdal importància pel desenvolupament de la teoria de la probabilitat a Rússia al segle XIX.

2.6.6 Nikolai Zernov

Nikolai Zernov (1804-1862) va ser un matemàtic de l'Imperi Rus, professor de la Universitat de Moscou, i com hem dit un dels pioners en introduir la teoria de la probabilitat en aquesta institució acadèmica. La seva obra més destacable, i que tractem a continuació, és *Probability Theory* (1843) [34].

En aquesta, Zernov comença fent un breu esquema històric de la teoria de la probabilitat. Llavors, exposa la seva visió determinista, en consonància amb el corrent filosòfic de l'època. Zernov era partidari de l'aplicació de les probabilitats a tots els àmbits socials i en les ciències naturals. Ho justifica per la importància que prenen en els estudis demogràfics, en les assegurances, en l'exactitud de les observacions astronòmiques, en la decisió dels tribunals per a obtenir un veredicte just, i fins i tot en estudis històrics i de lògica. En aquest sentit, presenta algunes aplicacions sense fonament ni justificació teòrica, com serien les jurídiques o històriques.

Seguidament, es defineixen alguns conceptes bàsics de probabilitats: la definició clàssica ($p = \frac{m}{n}$), on m nombre de casos favorables i n nombre de casos totals; propietats com la de $p+q=1$; suma i multiplicació de probabilitats; i probabilitat relativa. Aquesta última es presenta amb el següent exemple: tenim una capsa amb 3 boles vermelles, 2 blanques i 1 negra. Llavors, les probabilitats de treure'n cada tipus són $P_{vermell} = \frac{3}{6}$, $P_{blanc} = \frac{2}{6}$ i $P_{negre} = \frac{1}{6}$. Ara, si suposem només la probabilitat relativa d'obtenir una bola blanca o negra, sense considerar les vermelles, queda que $P_{blanc} = \frac{2}{3}$ i que $P_{negre} = \frac{1}{3}$.

A continuació, Zernov exposa el teorema de Bernouilli, i la importància que té aquest en les aplicacions de la probabilitat a la naturalesa. Igual que Ostrogradski, analitza el funcionament de les loteries, concloent també que només són altament profitoses a nivell econòmic pels organitzadors.

És remarcable l'enfocament que dóna Zernov a l'esperança moral, redefinint els conceptes de capital i valor. Això és, una quantitat de diners determinada s'ha de valorar no només per ella mateixa, sinó també per l'individu que n'és propietari. En aquest sentit, s'expressa així: *"Physical strength, any type of ability, and even the opportunity to beg or acquire debts is capital or in itself property."* [[34] p.22]

És partidari d'introduir i diferenciar l'esperança moral de l'esperança matemàtica, tal com predica el següent exemple: *"A person has 100 rubles; let the probability of his acquiring or losing 50 ruble equal to $\frac{1}{2}$. The ratio of the acquired sum to the future capital is $\frac{50}{150} = \frac{1}{3}$, and the lost sum to the present capital is $\frac{50}{100} = \frac{1}{2}$. It follows from here that the loss of 50 rubles results in a greater tangible loss, although form the point of view of mathematical expectation the game is fair."* [[34] p.25]

Finalment, Zernov aplica l'estadística a problemes demogràfics i d'assegurances. Pel que fa als primers, pren dades de mortalitat en nounats a Rússia l'any 1834: d'un total de 1.000.000 de nounats, 540.762 (més d'un 54%) van morir abans de fer els 5 anys. Aquestes dades preocupen i escandalitzen a Zernov; conclou que la societat russa en general no sap cuidar els nounats, que hi ha molta incultura i supersticions perjudicials per aquests, i que això és el pitjor enemic de l'Imperi.

En general, en aquest tractat Zernov es va basar en exposar i redefinir resultats probabilístics ja existents fins la data. Per tant, tot i la important introducció que fa de la teoria de la probabilitat a Moscou i Rússia en general en aquella època, Zernov no és considerat una figura remarcable en el desenvolupament d'aquesta teoria al segle XIX. En canvi, sí que ho és (i molt) l'il·lustre matemàtic rus Víktor Buniakovski (1804-1889), que introduim a continuació.

2.6.7 Víktor Buniakovski

Víktor Buniakovski (1804-1889) va néixer al poble de Bar, a l'Imperi Rus (actualment situat a la província de Vinnytsia, a Ucraïna). El seu pare, coronel de cavalleria, va morir l'any 1809. Conseqüentment, Buniakovski va ser inicialment educat en matemàtiques a casa pel general Tormasov, amic del seu pare. Gràcies a Tormasov, el jove Buniakovski va assistir a la Universitat de Paris (Sorbonna), on va estudiar matemàtiques. Allà, va ser alumne de Laplace i Poisson. Després de graduar-se l'any 1824, va seguir investigant a Paris sota la supervisió de Cauchy, amb qui va obtenir el doctorat.

L'any 1826, va retornar a Rússia, establint-se a Sant Petersburg com a professor i investigador, places que ocuparia ja tota la seva vida. Entre les institucions acadèmiques on va exercir, la més remarcable és la Universitat de Sant Petersburg, de 1846 a 1880. Buniakovski va fer importants contribucions en teoria de nombres (conjectura amb el seu nom), teoria de probabilitats, i en aplicacions matemàtiques a la física. Se li atribueix també el descobriment de la desigualtat de Cauchy-Schwarz.

El 1828 va entrar a formar part de l'Acadèmia de Ciències de Rússia, esdevenint vicepresident d'aquesta l'any 1864. També va ser vicepresident de l'Acadèmia de Ciències de Sant Petersburg del 1864 al 1889, d'on van sortir alguns dels matemàtics més importants de la història de Rússia (Chebyshev, Màrkov, etc.). Buniakovski és considerat com un gran innovador en probabilitat al context històric de la probabilitat a l'Imperi Rus.

La seva obra més important en probabilitats és *Foundations of the Mathematical Theory of Probabilities* (1846) [5]. Procedim a exposar-ne el contingut i problemes més importants.

En la introducció, Buniakovski remarca que va haver de desenvolupar una terminologia adequada per explicar la teoria de la probabilitat, doncs en aquell moment a Rússia no hi havia cap conveni formal establert. Ho va fer molt exitosament. Igual que Ostrogradski i Zernov, la seva recerca es basava en gran mesura en la *Théorie Analytique des Probabilitiés* [19] de Laplace. Conseqüentment, Bunyakovski defensa fermament el determinisme filosòfic, rebutjant per tant l'atzar com a concepte.

Considera la teoria de la probabilitat com una part de les matemàtiques aplicades (a diferència d'altres autors que ja la començaven a considerar una branca de les matemàtiques com a tal), i la defineix amb les següents frases:

"The Analysis of Probabilities examines and evaluates numerically events which depend

on causes not only completely unknown to us, but also those with due to our ignorance, are not subject to any suppositions" [[5] p.1]

"The likelihood, under different circumstances, may be smaller or larger, hence it, as any mathematical quantity, can be measured. This measure in the mathematical sense is called probability and its calculus which deals with its precise determination is the Analysis of Probabilities." [[5] p.3]

Per acabar la introducció, Buniakovski defineix la seva concepció de certesa moral. Això és, quan acceptem un fet com a cert, però no podem argumentar-lo adequadament. Per tant, cal combinar i contrastar la certesa moral amb la certesa matemàtica.

A continuació, Buniakovski defineix la probabilitat formalment, com la proporció entre casos equiprobables. Per a clarificar les seves afirmacions, presenta alguns problemes, entre els quals en destaquem dos.

Problema 1: volem trobar quantes tirades d'un dau són necessàries de manera que la probabilitat d'obtenir un resultat concret (p.e. sis) sigui $\frac{1}{2}$. Obté que per quatre tirades aquesta probabilitat és major que $\frac{1}{2}$. És a dir, en les quatre tirades, que obtinguem un sis és més probable que que no l'obtinguem.

Problema 2: volem trobar ara el nombre de tirades de dos daus necessàries perquè la probabilitat d'obtenir dos sisos simultàniament sigui $\frac{1}{2}$. Ara obté que la solució és $24,6$ tirades. Per tant, si fem 24 tirades del dau, que obtinguem un 12 és menys probable que que no l'obtinguem. I si fem 25 tirades, és més probable que surti un 12 que que no surti.

Més endavant, es presenta el teorema de Bernouilli (a partir del qual Laplace va formular el Teorema de Moivre-Laplace, com s'ha explicat en un apartat anterior). Buniakovski el presenta textualment de la següent forma:

"If the number of repetitions of a trial which leads to one of two simple events A or B is increased indefinitely, the ratio between the number of occurrences of this event steadily approaches the ratio of their simple probabilities and, finally, with an appropiate number of trials, differs from this ratio arbitrarily little." [[5] p.36]

Recordem que Buniakovski sempre exemplificava els teoremes i resultats amb problemes i exemples pràctics, i el teorema de Bernouilli no és una excepció:

Considerem m experiments, cadascun dels quals té resultat A o B, amb probabilitats $p = \frac{a}{a+b}$ i $q = 1 - p = \frac{b}{a+b}$ respectivament. Com que es considera una probabilitat composta, el cas més probable és $A^x B^{x'}$, on el quocient x/x' difereix arbitràriament poc de $a/b = p/(1-p)$ i $x + x' = 1$. Aleshores, Buniakovski presenta dos problemes:

Problema 1: Calculem la probabilitat P de que, en els m experiments (considerant m molt gran), A s'observarà en un rang de $[x-l, x+l]$ vegades, i per tant B s'observarà en un rang de $[x'-l, x'+l]$ vegades. La solució és

$$P = \frac{2}{\sqrt{\pi}} \int_0^t e^{-t^2} dt + \frac{\sqrt{m}}{\sqrt{2xx'm}} e^{-t^2}$$

on $t = l\sqrt{\frac{m}{2xx'}}$. Llavors, calculant la integral, Buniakovski obté que per $t \to \infty$, $P \to 1$.

Problema 2: sigui p la probabilitat de A desconeguda, però coneixem els cops que ocorre A en els m experiments (i), volem trobar la probabilitat P' de que p estigui continguda en uns límits determinats. Buniakovski defineix aquests límits com

$$\frac{i}{m} \mp \frac{t\sqrt{[2i(m-i)]}}{m^{\frac{3}{2}}} \qquad on \qquad t = \frac{l\sqrt{m}}{\sqrt{[2i(m-i)]}}$$

i expressa la probabilitat buscada

$$P' = \frac{2}{\sqrt{\pi}} \int_0^t e^{-t^2} dt + \frac{\sqrt{m}}{\sqrt{\pi[2i(m-i)]}} e^{-t^2}$$

En el capítol 3, es tracta l'esperança matemàtica, que Buniakovski introdueix a través dels jocs d'atzar, com és el cas de la loteria. Estudiant les probabilitats d'aquesta última, conclou que no és un joc just per als compradors.

Al capítol 4, es tracta l'esperança moral; i el següent capítol, anomenat *On the probabilities of human life*, es tracten problemes demogràfics. Per exemple, Buniakovski analitza la mortalitat infantil amb estadístiques de l'any 1842 a Moscou. Observa que el fet de viure a la ciutat implica un augment de la probabilitat de morir en els nens respecte els de zones rurals, i ho argumenta amb les males condicions de salubritat a la ciutat.

A continuació, trobem el capítol titulat *On the most favorable results of observations*, on s'estudia l'estimació d'errors observacionals i la distribució d'aquests. També s'hi afirma que el valor més probable d'un conjunt d'observacions és la mitjana aritmètica d'aquestes. Finalment, Buniakovski presenta la llei normal com a distribució més adequada pels errors de les observacions.

Tot i la importància que va adquirir Buniakovski a la comunitat científica i estadística del moment, degut als seus articles de caire formal i comprensible, cal mencionar que també va rebre algunes crítiques. Aquestes es fonamentaven en algunes aplicacions injustificades de les probabilitats a la justícia i als tribunals, com era comú en molts autors contemporanis.

Per concloure aquesta secció de les probabilitats a Rússia durant la primera meitat del segle XIX, podem afirmar que Ostrogradski i Buniakovski van ser els millors representants i les figures més influents en el desenvolupament de la teoria de la probabilitat. Tanmateix, ambdós usaven encara metodologies antigues i que començaven a ser obsoletes per a fer un salt endavant en aquesta branca de les matemàtiques. I és aquí on prenen importància els protagonistes en teoria de la probabilitat durant la segona meitat del segle XIX a Rússia.

3 Segona meitat del segle XIX

En aquest període temporal, ens centrarem sobretot en l'estudi dels matemàtics que van establir l'escola Russa, també anomenada escola de Sant Petersburg (degut a la universitat on es va localitzar tota la investigació i talent).

3.1 Pafnuti Chebyshev

Pafnuti Chebyshev (1821-1894) va ser un matemàtic rus, considerat el pare de les matemàtiques a Rússia. Va fer grans contribucions en la probabilitat, l'estadística, mecànica, i teoria de nombres. Ell va liderar l'establiment de l'escola russa de matemàtiques a Sant Petersburg, i el conseqüent reconeixement mundial d'aquesta. El treball en probabilitats i matemàtiques fet per Lobatxevski i Ostrogradski va ser el detonant per a començar aquesta revolució liderada per Chebyshev. Aquest va destacar no només com a investigador, sinó també com a docent i mentor, doncs sota la seva supervisió van sortir grans matemàtics, com per exemple A. A. Màrkov, o bé A. M. Liapunov.

Chebyshev va néixer al poble d'Okatovo, a la província de Kaluga (Rússia), en el si d'una família noble amb nou germans. Primerament, Chebyshev va ser educat a casa per la seva mare i la seva cosina. La seva infantesa i adolescència van estar marcades per un problema psicomotriu (posició de Tredelenburg) que li impedia caminar amb normalitat i fer moltes activitats amb els nens de la seva edat, motiu pel qual es va veure forçat a centrar-se en l'estudi de les matemàtiques. El 1832, la família es va mudar a Moscou, on Chebyshev va seguir amb l'educació a casa, però ara de la mà de reconeguts professors del moment.

El 1837, Chebyshev va començar matemàtiques a la Universitat de Moscou, on va tenir com a professors a Zernov i Brashman, que van exercir una important influència sobre ell. El 1847 va esdevenir professor associat a la Universitat de Sant Petersburg, on va conèixer a Buniakovski, el qual el va incitar a repassar els treballs de Leonhard Euler. El 1849 va defensar la seva tesi doctoral, titulada *The Theory of Congruences.* Això li va permetre escalar posicions a dins la universitat fins a esdevenir professor ordinari, fins al 1872, quan va esdevenir professor emèrit. Deu anys més tard, es va retirar per dedicar-se únicament a la investigació. Per altra banda, des del 1858 era membre de l'Acadèmia Imperial Russa de Ciències, i va rebre molts d'altres reconeixements degut als seus treballs. Entre ells, destaca que l'any 1893 va esdevenir membre honorable de la Societat Matemàtica de Sant Petersburg. En probabilitats, la seva contribució més important es coneix com la desigualtat de Chebyshev, que permet provar la llei feble dels grans nombres, i que presentarem detalladament en aquest capítol.

A diferència dels seus predecessors russos, Chebyshev va destacar pel seu enfocament més materialista, menys determinista, en les matemàtiques. És a dir, evitava les aplicacions sense fonamentar en ciències socials, i va adoptar una estructura metodològica més rigorosa. Aquesta metodologia seria la que prevaldria al llarg de l'escola Russa que Chebyshev va establir. Com hem comentat, la teoria de la probabilitat s'havia estancat en l'enfocament clàssic, necessitava una visió completament renovada. I és això el que va proporcionar Chebyshev, resolent alguns problemes clàssics de forma rigorosa i donant peu a nous problemes que resoldrien els seus aprenents, amb l'objectiu de sortir d'aquest estancament matemàtic. Tots aquests problemes que es plantejava tenien un objectiu pràctic, doncs la pràctica era un factor determinant en els seus estudis. Per a resoldre'ls, usava eines matemàtiques relativament simples, i afirmava que una solució aproximada es pot obtenir només si n'acotem els errors.

Fins llavors, les majors fites en teoria de la probabilitat eren la llei dels grans nombres i el teorema de Moivre-Laplace, exposats ja en els apartats de Laplace i Poisson. Tanmateix, no estaven ben determinats els límits d'aplicació d'aquests teoremes, doncs Laplace i Poisson, entre d'altres, els havien aplicat injustificadament en alguns àmbits com els tribunals, judicis, o problemes socials. Això va provocar una pèrdua de rigorositat de la teoria de la probabilitat entre la comunitat matemàtica, sorgint així la necessitat de crear una metodologia i limitació dels postulats més concises, clarificant les aplicacions pràctiques de cada teorema. I en això va excel·lir Chebyshev. Procedim a analitzar la seva obra i teoremes més importants, recollits en els diversos volums del llibre *Complete collected works.* [6], [7], [8]

L'any 1846 va publicar la seva tesi de màster a la Universitat de Moscou, titulada *An essay on elementary analysis of probability theory.* Durant els seus anys com a professor a la Universitat de Sant Petersburg, va impartir sobretot classes de teoria de nombres i teoria de probabilitats, i va investigar molt els teoremes del pas al límit en aquesta última. En la tesi de màster, Chebyshev exposa el seu principi elemental: obtenir límits per totes les aproximacions que feia. Això ho aplica per la demostració del teorema de Bernouilli en aquest mateix article. Abans, però, introdueix diverses definicions elementals de probabilitats:

"If out of a given number of different events under certain conditions one event should take place and there is no particular reason to expect that any one of the given events is preferred over the others, we designate these events as equipossible. [...] The ration of the number of equipossible cases favorable for the given event to the number of all the equiprobable cases. [...] A science on probabilities, known under the name of probability theory, deals with the determination of probabilities of an event based on the given connection of this event with events whose probabilities are known." [[8], pp. 28-29]

Veiem que, fins aquí, usava unes terminologies molt semblants als seus precedents clàssics. Llavors, també afegeix el següent enunciat:

Sigui F un esdeveniment que només ocorre conjuntament amb un dels μ esdeveniments $E_1, E_2, \dots, E_\mu$, i sabent que les probabilitats de F després que ocorrin cadascun d'aquests esdeveniments són $P_1, P_2, \dots, P_\mu$, respectivament. També es coneixen les probabilitats de $E_1, E_2, \dots, E_\mu$, que són $p_1, p_2, \dots, p_\mu$. Aleshores, la probabilitat de que E_1 ocorri conjuntament amb F és

$$\frac{P_1 p_1}{P_1 p_1 + P_2 p_2 + \cdots + P_\mu p_\mu}$$

o bé, usant notació actual de probabilitats condicionades seria:

$$\frac{P(F|E_1)P(E_1)}{P(F|E_1)P(E_1) + P(F|E_2)P(E_2) + \cdots + P(F|E_\mu)P(E_\mu)}$$

Al següent capítol, obté algunes fórmules rellevants com la probabilitat que, en μ successos, l'esdeveniment E_1' ocorre m_1 vegades; E_2' m_2 vegades;...;$E_{\sigma-1}'$ $m_{\sigma-1}$ vegades, mentre que l'esdeveniment oposat a $E_1', E_2', \dots, E_{\sigma-1}'$ ocorre $\mu - m_1 - m_2 - \cdots - m_{\sigma-1}$ vegades:

$$\frac{\mu! p_1^{m_1} p_2^{m_2} \cdots p_{\sigma-1}^{m_{\sigma-1}} (1 - p_1 - p_2 - \cdots - p_{\sigma-1})^{\mu - m_1 - m_2 - \cdots - m_{\sigma-1}}}{m_1! m_2! \cdots m_{\sigma-1}! (\mu - m_1 - m_2 - \cdots - m_{\sigma-1})!}$$

on $p_1, p_2, \dots, p_{\sigma-1}$ són les probabilitats dels esdeveniments $E_1', E_2', \dots, E_{\sigma-1}'$.

A continuació Chebyshev usa la fórmula de Stirling: $x! \approx \sqrt{2\pi}x^{x+\frac{1}{2}}e^{-x}$, per a obtenir la probabilitat de que, en μ experiments, un esdeveniment de probabilitat p ocorri m vegades:

$$P_{\mu,m} = \frac{1}{\sqrt{2\pi p(1-p)\mu}} e^{\frac{-Z^2}{2p(1-p)\mu}}$$

on $m = p\mu + Z$. Aquesta expressió es correspon amb la distribució normal, i com que l'esdeveniment només té una sola probabilitat (p), veiem que es tracta de la fórmula de James Bernouilli. De fet, Chebyshev prova aquesta de forma una mica més general, tal com s'exposa a continuació.

Considerem la probabilitat de que, en μ experiments, l'esdeveniment E (amb probabilitat p) ocorri almenys m vegades i com a molt $m+s$ vegades. Chebyshev l'obté i la denota $\prod_{\mu,m}^{m+s}$:

$$\prod_{\mu,m}^{m+s} = \frac{x}{\sqrt{\pi}(s+1)} \sum_{k=0}^{s} e^{-\left(X + x\frac{k}{s+1}\right)^2}$$

on

$$X = \frac{Z}{\sqrt{2p(1-p)\mu}}, \qquad x = \frac{s+1}{\sqrt{2p(1-p)\mu}}$$

Per valors de $s+1$ grans, cal usar computació per a calcular el sumatori. Així, Chebyshev estudia aquest cas introduint la funció $T(Z)$ (i una taula de valors corresponent):

$$T(Z) = \frac{1}{\sqrt{2\pi}10^{13}} \sum_{k=0}^{10^{15}-1} e^{-\left(Z + \frac{k}{10^{13}}\right)^2}$$

de manera que la probabilitat anterior per valors grans de $s+1$ és

$$\prod_{\mu,m}^{m+s} = T(x) - T(X+x)$$

Llavors, amb això Chebyshev calcula la probabilitat que un esdeveniment amb probabilitat p ocorri d vegades en un total de μ experiments, amb $d \in [p\mu - \sqrt{2p(1-p)\mu}, p\mu + \sqrt{2p(1-p)\mu}]$:

$$\prod_{\mu,p\mu-\sqrt{2p(1-p)\mu}}^{p\mu+\sqrt{2p(1-p)\mu}} = 1 - 2T(1) = 0.8427008$$

i calcula el mateix reduint els intervals de d en un factor 2 i 5, respectivament:

si $d \in [p\mu - 2\sqrt{2p(1-p)\mu}, p\mu + 2\sqrt{2p(1-p)\mu}]$, queda

$$\prod_{\mu,p\mu-2\sqrt{2p(1-p)\mu}}^{p\mu+2\sqrt{2p(1-p)\mu}} = 1 - 2T(2) = 0.9953224$$

i si $d \in [p\mu - 5\sqrt{2p(1-p)\mu}, p\mu + 5\sqrt{2p(1-p)\mu}]$, queda

$$\prod_{\mu,p\mu-5\sqrt{2p(1-p)\mu}}^{p\mu+5\sqrt{2p(1-p)\mu}} = 1 - 2T(5) = 0.999999999998$$

Veiem, doncs, que si l'interval del nombre de successos en que ocorre l'esdeveniment (d) es fa petit, la probabilitat de que l'esdeveniment ocorri les d vegades augmenta. Una forma alternativa de llegir aquests resultats és la següent: són les probabilitats de que, en μ experiments, el quocient del nombre de vegades que ocorre un esdeveniment entre el total μ difereixi de $p - \frac{1}{2\mu}$ per menys que

$$\frac{\sqrt{2p(1-p)\mu}}{\sqrt{\mu}-\frac{1}{2\mu}}, \qquad \frac{2\sqrt{2p(1-p)\mu}}{\sqrt{\mu}-\frac{1}{2\mu}}, \qquad \frac{5\sqrt{2p(1-p)\mu}}{\sqrt{\mu}-\frac{1}{2\mu}}$$

D'aquesta forma, Chebyshev obté una "prova"pel teorema de James Bernouilli, i així ho expressa:

"It is therefore extremely probable that, as the number of cases increases, the ratio of the number of repetitions of an event to the number of all the cases will deviate very little from the probability of the event in these cases. This is the meaning of J. Bernouilli's theorem." [[8], p. 53]

Per tant, a més de provar el teorema de Bernouilli, Chebyshev va donar fites pels errors comesos en aplicar-lo. De fet, fitar les aproximacions era un dels seus principals objectius.

Més tard, Chebyshev va fer la primera demostració general del teorema de Poisson per a un nombre finit de probabilitats, i la va presentar a l'article *Elementary proof of a general proposition in probability theory*, de l'any 1846. Com a cas particular del teorema de Poisson, s'obté el teorema de Bernouilli. Això és, quan les diferents probabilitats de l'esdeveniment E en els μ diferents experiments, $p_1, p_2, \ldots, p_\mu$, valen totes igual: p.

Chebyshev afirma que la demostració que va donar Poisson al seu propi teorema és incompleta, doncs no va acotar els errors. Per tant, Chebyshev procedeix a fer una prova formal, amb els següents passos ordenats.

Primer, defineix P_m la probabilitat de que l'esdeveniment E ocorri com a mínim m vegades, en un total de μ experiments; i suposa que P_m és simètrica respecte a p_k i hi manté una dependència lineal, on p_k és la probabilitat de E en l'experiment k-èssim, on $k = 1, \ldots, \mu$. Com a exemple, si prenem p_1 i p_2, queda

$$P_m = U + V(p_1 + p_2) + W p_1 p_2$$

on U, V, W no depenen de p_1, p_2. Així, si $a := p_1 + p_2 \in (0, 1]$, llavors P_m assoleix el seu màxim per $p_1 = p_2 = a/2$ o per $p_1 = 0$, evidenciant doncs un comportament simètric relatiu a p_1 i p_2.

Seguidament, Chebyshev presenta el primer dels tres teoremes que confeccionaran la demostració del teorema de Poisson.

Teorema 3.1. *La màxima probabilitat P_m, considerant que $p_1 + p_2 + \ldots + p_\mu = s$, s'obté pels valors*

$$p_1 = 0, p_2 = 0, \ldots, p_\rho = 0, p_{\rho+1} = 1, p_{\rho+2} = 1, \ldots, p_{\rho+\sigma} = 1,$$
$$p_{\rho+\sigma+1} = \frac{s-\sigma}{\mu-\rho-\sigma}, p_{\rho+\sigma+2} = \frac{s-\sigma}{\mu-\rho-\sigma}, \ldots, p_\mu = \frac{s-\sigma}{\mu-\rho-\sigma}$$

on ρ i σ valors fixats.

Amb aquest resultat, Chebyshev obté la següent desigualtat per $m > s+1$:

$$P_m < \frac{\mu!}{m!(\mu-m)!}\left(\frac{s}{\mu}\right)^m \left(\frac{\mu-s}{\mu}\right)^{\mu-m+1} \frac{m}{m-s} \tag{3.1}$$

Que pel cas particular de Bernouilli, i.e. $p_1 = p_2 = \cdots = p_\mu = s/\mu$, queda com $P_m < \rho_m \frac{m}{m-s}$, on ρ_m és la probabilitat que E ocorre m vegades en el total d'experiments fets.

Fent una sèrie de treballs analítics, Chebyshev transforma la desigualtat (3.1) en

$$P_m < \frac{1}{2(m-s)}\sqrt{\frac{m(\mu-m)}{\mu}}\left(\frac{s}{m}\right)^m\left(\frac{\mu-s}{\mu-m}\right)^{\mu-m+1}$$

la qual dona peu al segon teorema:

Teorema 3.2. *Siguin $p_1, p_2, \ldots, p_\mu$ les probabilitats de l'esdeveniment E en cadascun dels μ experiments, complint $p_1 + p_2 + \ldots + p_\mu = s$, llavors la probabilitat que E no ocorri més de n vegades ve acotada superiorment per la magnitud*

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

on $n < s - 1$.

Finalment, Chebyshev obté el tercer teorema com a conseqüència dels altres dos:

Teorema 3.3. *"However, repetitions of the event E may yield only one of the following cases: either the event E occurs at least m times, or no more than n times, or, finally, more than n, but less than m times. Therefore, the probability of the latter case is determined by subtracting the probabilities of the first two cases from unity." [[6], p. 21]*

Finalment, a partir dels anteriors teoremes derivem el següent resultat: Siguin $p_1 + p_2 + \ldots + p_\mu$ les probabilitats de l'esdeveniment E en cadascun dels μ experiments fets, de manera que $p_1 + p_2 + \ldots + p_\mu = s$. Aleshores la probabilitat que el nombre de repeticions de E sigui menor que m i major que n té per cota inferior

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

per $m > s + 1$ i $n < s - 1$.

D'aquesta manera, Chebyshev va obtenir la demostració del teorema de Poisson per esdeveniments independents. En aquest sentit, cal remarcar que Chebyshev mai va fer referència explícita al fet que els esdeveniments fossin independents, sinó que segurament ja ho donava per suposat, doncs els seus resultats són vàlids només per aquest cas. Mai fa referència a esdeveniments dependents, per tant s'intueix que era partidari d'aplicar les probabilitats només per esdeveniments independents.

A continuació, exposem la famosa desigualtat de Chebyshev, que el mateix presenta a través de diversos teoremes en l'article *On mean values*, l'any 1866. D'aquesta desigualtat, se'n deriven els teoremes de Bernouilli i Poisson, com a casos particulars. Chebyshev comença l'article amb els següents teoremes:

Teorema 3.4. *Siguin $a, b, c, \ldots$ les esperances matemàtiques de les variables $x, y, z, \ldots$, i $a_1, b_1, c_1, \ldots$ les esperances dels seus quadrats $x^2, y^2, z^2, \ldots$; llavors la probabilitat que la suma $x + y + z + \ldots$ estigui inclosa en els límits*

$$a + b + c + \cdots - \alpha[a_1 + b_1 + c_1 + \cdots - a^2 - b^2 - c^2 - \cdots]^{1/2}$$

i

$$a + b + c + \cdots + \alpha[a_1 + b_1 + c_1 + \cdots - a^2 - b^2 - c^2 - \cdots]^{1/2}$$

serà sempre major que $1 - 1/\alpha^2$, independentment del valor de α. [[6], p. 431]

A continuació, Chebyshev defineix $\alpha = \frac{\sqrt{N}}{t}$, on N és el nombre de variables $x, y, z, \ldots$. Amb això, enuncia el teorema que procedeix.

Teorema 3.5. *Siguin $a, b, c, \ldots$ les esperances matemàtiques de les variables $x, y, z, \ldots$, i $a_1, b_1, c_1, \ldots$ les esperances dels seus quadrats $x^2, y^2, z^2, \ldots$; llavors la probabilitat de que la diferència entre la mitjana aritmètica de les variables i la mitjana aritmètica de les esperances respectives no excedeixi*

$$\frac{1}{t}\sqrt{\frac{a_1 + b_1 + c_1 + \cdots}{N} - \frac{a^2 + b^2 + c^2 - \cdots}{N}}$$

serà sempre major que $1 - t^2/N$, independentment del valor de t. [[6], p. 435]

Aquesta última és la desigualtat de Chebyshev, que en nomenclatura actual, si prenem una sola variable aleatòria x, s'expressa com

$$P(|x - \bar{x}| \leq \epsilon) \geq 1 - [D(x)/\epsilon^2] \tag{3.2}$$

on $\bar{x} = E(x)$, $D(x) = \bar{x^2} - (\bar{x})^2$ és la variància finita de x, i ϵ és un nombre positiu arbitrari.

A partir d'aquesta desigualtat, Chebyshev obté la llei dels grans nombres a través del següent teorema:

Teorema 3.6. *Si les esperances matemàtiques de les N variables $U_1, U_2, U_3, \ldots$ i els seus quadrats $U_1^2, U_2^2, U_3^2, \ldots$, són finites, la probabilitat de que la diferència entre la mitjana aritmètica d'aquestes N variables i la mitjana aritmètica de les esperances respectives sigui menor que una quantitat donada, esdevé 1 quan N tendeix a infinit. [[6], p. 436]*

El reescrivim de forma més analítica: prenent x en comptes de U per les N variables, com que les esperances i els seus quadrats estan acotades, tenim que $D(x_i) < c, i = 1, \ldots, N$, per c fixada. Llavors, la variància de la mitjana aritmètica queda acotada com

$$D(\tilde{x}) = \frac{D(x_1) + D(x_2 + \cdots + D(x_N)}{N^2} < \frac{Nc}{N^2} = \frac{c}{N}$$

Aleshores, si apliquem això a la desigualtat de Chebyshev (3.2), queda com

$$P\left(\left|\frac{x_1 + x_2 + \cdots + x_N}{N} - \frac{\bar{x}_1 + \bar{x}_2 + \cdots + \bar{x}_N}{N}\right| \leq \epsilon\right) \geq 1 - \frac{D(\tilde{x})}{\epsilon^2} > 1 - \frac{c}{N\epsilon^2}$$

i fent el límit per $N \to \infty$ queda

$$\lim_{N\to\infty} P\left(\left|\frac{x_1 + x_2 + \cdots + x_N}{N} - \frac{\bar{x}_1 + \bar{x}_2 + \cdots + \bar{x}_N}{N}\right| \leq \epsilon\right) = 1$$

Aquest resultat és l'anomenat teorema de Chebyshev o llei dels grans nombres de Chebyshev, doncs estableix que per un nombre prou gran d'experiments, es pot afirmar que la mitjana aritmètica de les variables difereix arbitràriament poc d'un valor fixat (la mitjana de les seves esperances matemàtiques).

Veiem ara com d'aquesta llei dels grans nombres s'obtenen els teoremes de Poisson i Bernouilli com a casos particulars:

Fem N experiments, i l'esdeveniment E ocorre amb probabilitat $p_1, p_2, \ldots, p_N$ en cadascun d'ells. Sigui la variable x_i el nombre de vegades que succeeix E en l'experiment i-èssim. Llavors, clarament tenim que

$$E(x_i) = 1p_i + 0q_i, \qquad D(x_i) = p_i q_i \leq 1/4,$$
$$\tilde{x} = \frac{x_1 + x_2 + \cdots + x_N}{N} = \frac{m}{N}, \qquad \tilde{p} = \frac{p_1 + p_2 + \cdots + p_N}{N}$$

i per tant les variables x_i compleixen la desigualtat de Chebyshev (3.2):

$$\lim_{N\to\infty} P(|(m/N) - \tilde{p}| \leq \epsilon) = 1$$

I això és de fet el teorema de Poisson per N variables aleatòries independents. Aleshores, pel cas particular que $p_i = p \ \ \forall i = 1, \ldots, N$, tenim que $\tilde{p} = p$, i per tant s'obté el teorema de Bernouilli:

$$\lim_{N\to\infty} P(|(m/N) - p| \leq \epsilon) = 1$$

Seguidament, presentem el segon gran problema que va tractar Chebyshev: el teorema del límit central. El presenta al seu article de 1887 titulat *On two theorems concerning probabilities*. El títol fa referència als dos problemes cabdals en probabilitat: la llei dels grans nombres i el teorema del límit central. Aquest últim l'exposa com segueix:

Teorema 3.7. *Considerem les n variables $u_1, u_2, \ldots, u_n$, amb esperances matemàtiques nul·les, i amb les esperances de totes les seves potències acotades superiorment; això és, si denominem l'esperança de la variable i-èssima a la potència μ-èssima com a_i^μ, les condicions són:*

- $a_i^1 = 0$
- $a_i^2, a_i^3, \ldots, a_i^\mu, \ldots$ *acotats superiorment*

Aleshores, es compleix que

$$\lim_{n\to\infty} P\left(t \leq \frac{u_1 + u_2 + \cdots + u_n}{\sqrt{2(a_1^2 + a_2^2 + \cdots + a_n^2)}} \leq t'\right) = \frac{1}{\sqrt{\pi}} \int_t^{t'} e^{-x^2} dx$$

Més endavant, Kolmogórov va analitzar aquest teorema, notant que cal afegir una condició per evitar que sigui incorrecte. Això és, el fet que

$$\frac{1}{q_n^2} = \frac{a_1^2 + a_2^2 + \cdots + a_n^2}{n}$$

pot tendir a 0 quan $n \to \infty$. Per evitar-ho, Kolmogórov ho modifica aplicant la condició que aquest quocient s'acosta al valor fix positiu $1/q^2$ quan $n \to \infty$. A partir d'això, Kolmogórov va poder provar aquest teorema de forma rigorosa.

Finalment, una altra gran contribució de Chebyshev en la teoria de la probabilitat va ser el desenvolupament del mètode dels moments. De fet, l'usa extensivament en molts dels seus articles, doncs és molt útil per la llei dels grans nombres i el teorema del límit central. No entrarem en l'anàlisi dels moments que va fer Chebyshev, però sí que ho descrivim en línies generals: siguin els nombres $C_i, i = 1, \ldots, k$, l'objectiu és trobar una funció $\varphi(x)$ tal que $C_k = \int_a^b x^k d\varphi(x)$. El valor C_k s'anomena moment d'ordre k. Sigui $f(x)$ la derivada de $\varphi(x)$, es pot reescriure com $C_k = \int_a^b x^k f(x)dx$. En notació moderna, es diu que sigui f la funció de densitat de la variable aleatòria x, el moment d'ordre k de x és l'esperança matemàtica de x^k, i.e. $C_k = M_k = E(x^k)$, on E fa referència a esperança. En altres paraules, l'esperança és el moment d'odre 1 de la variable x.

En resum, les més importants contribucions de Chebyshev en la teoria de la probabilitat van ser la llei dels grans nombres i el teorema de límit central, ambdós aplicats a la suma de variables aleatòries independents. Tanmateix, no va arribar a provar-los per altres tipus de variables aleatòries, com serien les dependents. Va deixar per tant aquest exercici per les generacions posteriors, definint d'aquesta manera un camí a seguir pels seus alumnes. El posterior desenvolupament de la teoria de la probabilitat depenia de l'excel·lència i ambició d'aquests.

Chebyshev va donar restriccions a l'aplicabilitat dels principals teoremes en probabilitat. Així, va establir per primera vegada un significat matemàtic clar a les variables aleatòries, formalitzant i descrivint adequadament les distribucions que segueixen. En conseqüència, a partir d'aquell moment, es podia discernir si per unes variables aleatòries donades es podien aplicar teoremes de pas al límit o no.

Com ja hem dit, Chebyshev era partidari del materialisme, i per això la seva teoria de la probabilitat està molt enfocada a aplicacions pràctiques en diferents àmbits. Amb tot, Chebyshev és considerat el creador i "pare"de l'escola russa de teoria de la probabilitat. El també reconegut matemàtic rus A. Ya. Khinchin (1894-1959), posteriorment exposa aquest fet amb les següents paraules:

"the only country [Russia] in which the mathematical foundations of probability theory were cultivated with the seriousness it deserved, in view of its prominent role in the natural sciences and engineering. It is entirely due to the works of Chebyshev that the Russian school of probability theory attained this exceptional position."

3.2 Andrei Màrkov

Andrei Màrkov (1856-1922) va ser un matemàtic rus, considerat el més famós deixeble i seguidor de Chebyshev. Gràcies a ell, la teoria de la probabilitat promoguda per Chebyshev es va extendre a tota la comunitat matemàtica. Màrkov va destacar pels seus estudis en teoremes de pas al límit en variables independents i també dependents, les últimes de les quals va tractar amb les anomenades "cadenes de Màrkov". Nascut a Ryazan, Rússia, Màrkov va assistir a l'escola primària de Sant Petersburg, on no va destacar gaire a nivell acadèmic, ans el contrari: alguns professors el recordaven com un estudiant rebel. Emperò, les matemàtiques sí que se li donaven bé, i de fet va estudiar-les posteriorment a la Universitat de Sant Petersburg. Allà, va tenir com a professors a reconeguts matemàtics, entre els quals destaca Chebyshev. El 1878, un cop va acabar el grau universitari, li van oferir esdevenir professor associat a la mateixa Universitat de Sant Petersburg. Així, aquell mateix any va començar a preparar-se per les proves i plaça de professor. L'any 1880, Màrkov va presentar la seva tesi de màster, començant també a exercir com a professor. Entre les classes que va impartir, hi trobem la de teoria de la probabilitat, en la qual va succeir a Chebyshev després que aquest es retirés l'any 1882.

L'any 1884, va presentar la seva tesi doctoral, i poc després (1886) va esdevenir professor extraordinari. El 1890, va ser escollit membre de l'Acadèmia de Ciències de Rússia. Va seguir tota la seva vida investigant i donant classes a la Universitat de Sant Petersburg. De fet, l'any 1905 va esdevenir professor emèrit, i per tant ja es podia retirar, però va seguir impartint teoria de la probabilitat i càlcul diferencial fins als seus últims dies.

Màrkov va destacar molt per la seva qualitat com a docent, doncs la seva metodologia era rigorosa però no es basava en donar molta informació i "saturar" als seus estudiants, sinó donar unes eines i bases sòlides sobre les quals els alumnes poguessin desenvolupar correctament la teoria de la probabilitat. Màrkov volia promoure així el pensament crític, donant peu a futurs matemàtics que ajudessin al creixement de les matemàtiques. En diversos textos hi ha constància de la preparació que feia per impartir les seves classes, així com de l'agudesa i agilitat mental amb la que transmetia els coneixements. Tota aquesta qualitat com a pedagog es plasma perfectament en el seu llibre *The Calculus of Probabilities* [26], de 1913. Comencem analitzant aquesta obra per la importància que té en els seus estudis, tot i que després es tractaran altres obres més específiques, on Màrkov va provar els teoremes de pas al límit excel·lentment. El llibre [26] és de caire educatiu, i s'hi presenten les demostracions amb un rigor i exhaustivitat impecables. A més, s'hi presenten moltes aplicacions i investigacions dels resultats de la teoria de la probabilitat. Els continguts principals són els següents:

1. Al primer capítol, s'introdueixen les nocions bàsiques: definició clàssica de probabilitat, definició d'equiprobabilitat, regles d'addició i multiplicació, etc.

2. Al segon capítol, s'obté la fórmula de Bernouilli: $P_{m,n} = C_n^m p^m q^{n-m}$. De fet, Màrkov dóna tres demostracions diferents al teorema de Bernouilli, essent una d'elles la original del propi Bernouilli. Cal remarcar que també posa en dubte la certesa absoluta de la definició de probabilitat associada a aquest teorema: $\lim_{n\to\infty} m/n = p$. Finalment, Màrkov també dedueix la fórmula de Stirling i el teorema de Moivre-Laplace.

3. Al tercer capítol, es tracta sobretot la llei dels grans nombres, que Màrkov considera com una generalització del teorema de Bernouilli. També s'hi presenten algunes propietats de l'esperança matemàtica, així com la demostració de la desigualtat i el teorema de Chebyshev.

4. Els capítols següents tracten diverses aplicacions de la teoria de la probabilitat: loteries, jocs d'atzar, assegurances,...; així com probabilitats de fraccions racionals irreductibles i probabilitats futures.

5. Addicionalment, apareixen alguns dels articles més importants de Màrkov. Per tant, en el seu conjunt, aquest llibre representa un recull perfectament estructurat de les contribucions més importants de Màrkov en teoria de la probabilitat, és la seva obra cabdal.

Passem ara als teoremes de pas al límit. La primera demostració que fa Màrkov del teorema del límit central es troba en unes correspondències amb A. V. Vasiliev, l'any 1898. Presenta el teorema a partir d'esperances matemàtiques i el mètode dels moments, com segueix:

Teorema 3.8. *Siguin les variables $x_1, x_2, \ldots, x_n$ amb esperança nul·la, i esperança de x_n^k finita $\forall k \geq 2$ fins i tot quan $n \to \infty$. Expressat en termes de moments: el moment d'ordre u $M(x_i) = 0 \quad \forall i = 1, \ldots, n$, i per cada enter $k, k \geq 2$, el moment k-èssim de x_n ($M(x_n^k)$) existeix i és finit, per qualsevol n prou gran. Llavors, sigui*

$$Y_n = \frac{x_1, x_2, \cdots, x_n}{\sqrt{n}}$$

Màrkov prova que amb aquestes condicions es compleix que

$$\lim_{n\to\infty} (M(Y_n^m) - A_m[M(Y_n^2)]^{m/2}) = 0$$

on

$$A_m = \frac{2^{m/2}}{\sqrt{pi}} \int_{-\infty}^{\infty} t^m e^{-t^2} dt$$

per m un enter qualsevol.

Notem que aquest enunciat del teorema del límit central difereix en certa manera del de Chebyshev, però essencialment és el mateix. Màrkov ho fa així per a evitar petits errors que havia trobat en el plantejament de Chebyshev, aconseguint així provar-lo més rigorosament.

Aquell mateix any, el 1898, Màrkov prova de nou el teorema del límit central en l'article *On the roots of the equation* $e^{x^2}\partial^m e^{-x^2}/\partial x^m = 0$ [24], aquesta vegada enunciant-lo d'una forma alternativa que segurament ens és més familiar:

Teorema 3.9. *La probabilitat que la suma de variables aleatòries independents $u_1 + u_2 + \cdots + u_n$ estigui continguda entre els límits*

$$\alpha\sqrt{2(a_1 + a_2 + \cdots + a_n)} \qquad i \qquad \beta\sqrt{2(a_1 + a_2 + \cdots + a_n)}$$

on $a_1 + a_2 + \ldots + a_n$ són les esperances matemàtiques de les variables $u_1^2, u_2^2, \ldots, u_n^2$ i α i β són dos valors arbitraris, quan $n \to \infty$ s'acosta al valor límit

$$\frac{1}{\sqrt{\pi}} \int_{\alpha}^{\beta} e^{-x^2} dx$$

sempre que les variables independents $u_1 + u_2 + \ldots + u_n$ compleixin les següents condicions:

1. *les esperances matemàtiques de $u_1 + u_2 + \ldots + u_n$ són nul·les.*
2. *les esperances de $u_n^2, u_n^3, u_n^4, \ldots$ són finites.*
3. *l'esperança de u_k^2 no s'anul·la per k indefinidament gran.*

D'aquesta forma, veiem que Màrkov va enunciar i provar el teorema del límit central amb les mateixes hipòtesis que Chebyshev, tot i que prenent com a fil conductor els moments en comptes de les esperances. Per altra banda, Màrkov va fer referència explícita a la independència de les seves variables, cosa que Chebyshev sempre va donar per suposat.

Aquest mateix teorema del límit central va ser plantejat i demostrat pel també matemàtic rus A.M. Liapunov (1857-1918), en dos articles de 1900 i 1901, de forma més genèrica i usant un mètode diferent. Concretament, Liapunov el va provar usant el mètode de les funcions característiques, que ell mateix va desenvolupar. D'aquesta forma, al no usar el mètode de moments va reduir molt el nombre de restriccions/condicions que requeria el plantejament de Màrkov. Això és perquè les funcions característiques existeixen per qualsevol variable aleatòria, mentre que el moment (i.e. l'esperança) no existeix en qualsevol cas. De fet, el que coneixem actualment com a teorema del límit central és el que va plantejar Liapunov en aquell moment. S'exposa en el següent subapartat, dedicat a Liapunov, juntament amb el mètode de les funcions característiques.

Davant d'aquests resultats, Màrkov va reconèixer que Liapunov havia aconseguit una formulació més genèrica del teorema del límit central, i que usant esperances matemàtiques no es podia afirmar l'existència d'algunes en particular per als casos estudiats per Liapunov. Tanmateix, Màrkov no va donar per perduda la seva metodologia, i va treballar fins aconseguir provar el teorema del límit central (plantejat per Liapunov) usant moments. El plantejament corresponent va ser el següent:

Teorema 3.10. *Sigui $z_1, z_2, \ldots, z_k, \ldots, z_n, \ldots$ una seqüència infinita de variables aleatòries independents, de manera que per cada k existeixen les següents quantitats (moments)*

$$a_k = M(z_k) \qquad b_k = M(z_k - a_k)^2$$

i també

$$b_k^{2+\delta} = M|z_k - a_k|^{2+\delta}$$

on δ és un valor positiu. Per últim, suposem que el quocient

$$\frac{b_1^{2+\delta} + b_2^{2+\delta} + \cdots + b_n^{2+\delta}}{(b_1 + b_2 + \cdots + b_n)^{1+\frac{\delta}{2}}}$$

tendeix a zero quan $n \to \infty$. Fins aquí, tenim les condicions de Liapunov expressades en forma de moments. Llavors, Màrkov observa que permeten obtenir el teorema del límit central. Això és, siguin dos valors t_1 i t_2 tals que $t_2 > t_1$, es compleix que la probabilitat de la desigualtat

$$t_1 < \frac{z_1 + z_2 + \cdots + z_n - a_1 - a_2 - \cdots - a_n}{\sqrt{2(b_1 + b_2 + \cdots + b_n)}} < t_2$$

tendeix al valor límit

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

quan $n \to \infty$. [[25], pp. 322-323]

La demostració d'aquest teorema es va publicar en l'article de Màrkov de 1912, titulat *The theorem on the limit of probability for the cases of A.M. Lyapunov.* Per a demostrar el TLC de Liapunov exitosament, Màrkov aplica el mètode dels moments a través de variables aleatòries truncades, les quals es defineixen així:

$$x_k = \left\{ \begin{array}{ll} z_k - a_k & \textit{per } |z_k - a_k| < N \\ 0 & \textit{per } |z_k - a_k| \geq N \end{array} \right\}$$

on z_k són les variables aleatòries en qüestió i $a_k = M(z_k)$. L'interessant d'aquestes variables truncades x_k és que tenen moments de qualsevol ordre, permetent així provar el teorema de Liapunov sense problemes. A més, la suma de les variables truncades x_k és equivalent a la suma de les variables originals z_k per N prou gran.

L'altra gran contribució de Màrkov en la teoria de la probabilitat va ser l'estudi de variables aleatòries dependents, concretament per als dos grans problemes d'aquesta teoria: la llei dels grans nombres i el teorema del límit central.

Pel que fa a la llei dels grans nombres, és necessari fer un petit parèntesi en forma d'aclariment respecte l'origen del terme. Tot i que J. Bernouilli va ser el primer en provar-la com a cas concret, no li va donar el nom ell. Va ser Poisson qui va inventar aquest terme, generalitzant el cas de Bernouilli per probabilitats diferents en cada variable aleatòria. Finalment, Chebyshev obté una major generalització que Poisson, però no s'hi refereix com a llei dels grans nombres.

Amb això, es pot concloure que els tres teoremes (Bernouilli, Poisson i Chebyshev) són realment diferents formulacions de la llei dels grans nombres, essent la última la més general. Pel que fa a Màrkov, la interpreta igual que els seus predecessors, expressant-la així:

"in view of which it can be asserted with probability as close to certainty as desired that the arithmetic mean of several variables, provided the number of these variables is large, will deviate in an arbitrarily small amount from the arithmetic mean of their mathematical expectations". [[25], p. 341]

A més, Màrkov considera que la llei dels grans nombres de Chebyshev no és només vàlida per variables independents, estenent-la així al cas de variables dependents amb variàncies acotades ($D(x) \leq c$). Ho fa en l'article *Extension of the law of large numbers to dependent variables*, de 1907. Comença provant-la per variables independents, i.e. siguin les variables $x_1, x_2, \ldots, x_n$ tals que

$$\lim_{n\to\infty} \frac{1}{n^2} \sum_{k=1}^{n} D(x_k) = 0$$

llavors es compleix

$$\lim_{n\to\infty} P\left(\left|\frac{x_1, x_2, \cdots, x_n}{n} - \frac{M(x_1), M(x_2), \cdots, M(x_n)}{n}\right| < \epsilon\right) = 1.$$

I a continuació demostra que aquest resultat es pot aplicar també a la suma $S_n = \sum_{i=1}^{n} x_i$, amb les condicions que $D(x_i) < c$ i que la dependència entre les variables és tal que l'augment d'una d'elles comporta la disminució de l'esperança matemàtica de les altres. Aquesta última condició també es pot interpretar a la inversa: la disminució de l'esperança d'una de les variables comporta l'augment de la suma de la resta de variables.

Més endavant, Màrkov va introduir les famoses "cadenes de Màrkov" per a tractar seqüències de variables dependents, i veure que també aquestes seguien la llei dels grans nombres. Els dos tipus principals de cadenes són la simple i la composta. La simple es basa en una seqüència de variables aleatòries, tals que la probabilitat d'un resultat concret de l'experiment $n+1$-èssim (variable $n+1$) depèn únicament del resultat de l'experiment n-èssim (variable n). Pel que fa a la cadena composta, aquesta probabilitat depèn dels resultats de tots els n experiments previs. Màrkov usa cadenes per comprovar la validesa de la llei dels grans nombres per variables dependents, així com aplica el TLC a la suma de variables aleatòries formant cadenes simples i compostes. Aquest últim el tracta exhaustivament en l'article *Investigation of a general case of trials connected in* *a chain* (1910): prova el TLC per la suma $S_n = \sum_{i=1}^{n} x_i$, on x_i formen una cadena i admeten només dos valors amb probabilitats de transició

$$p_i' = P(x_i = 1|x_{i-1} = 1), \qquad p_i'' = P(x_i = 1|x_{i-1} = 0)$$

que compleixen $p_0 < p_i' < 1-p_0, p_0 < p_i'' < 1-p_0$, essent $p_0 > 0$ una constant independent de n.

A part de les simples i compostes, Màrkov va introduir diversos tipus de cadenes, entre les quals destaquen les cadenes de Markov-Bruns. El cas més simple d'aquest tipus de cadena és el següent: prenem una seqüència $Y_1, Y_2, \ldots, Y_n, \ldots$ de variables independents Y_n, tals que $P(Y_n = 1) = \alpha$ i $P(Y_n = 0) = \beta$ $(\alpha + \beta = 1)$. Ara, la cadena de Markov-Bruns és la seqüència de variables dependents $x_1, x_2, \ldots, x_n, \ldots$ tal que $x_n = Y_n Y_{n+1}, n = 1, 2, \ldots$.

Les cadenes de Màrkov van ser de gran importància per l'aplicació de la teoria de la probabilitat a les ciències naturals i l'enginyeria. S'usen per sistemes que es mouen d'un estat a un altre només per uns instants de temps determinats $t_1, t_2, \ldots, t_k, \ldots$ i amb unes probabilitats determinades. Això és, les probabilitats condicionals de que el sistema es trobi en l'estat ω_j en l'instant t_{k+1} si en l'instant t_k es trobava en l'estat ω_i. Un exemple d'aquests sistemes és el model de l'àtom de Bohr, on l'electró es pot localitzar només en les òrbites admissibles, i per tant les transicions només poden ser unes determinades.

La generalització de les cadenes de Màrkov s'anomenen "processos de Màrkov", en els quals l'estat inicial del sistema determina la distribució de probabilitat dels altres possibles estats en instants de temps posteriors. Conseqüentment, les cadenes de Màrkov són casos particulars en que el sistema evoluciona de forma escalonada pels instants de temps determinats. Entre els camps de la física en que s'aplica, destaca la teoria del moviment Brownià (que es tracta breument en l'annex, dedicat a la física estadística).

Les investigacions de Màrkov en àmbits més aplicats no van consistir en les ciències naturals i enginyeria, sinó en el llenguatge rus. Concretament, va estudiar l'intercanvi de vocals i consonants. Això és, la probabilitat que una lletra escollida aleatòriament en un text escrit rus fos una vocal. Va concloure que aquesta probabilitat era 0.423.

3.3 Aleksandr Liapunov

Aleksandr Liapunov (1857-1918) va ser un matemàtic rus, reconegut pel desenvolupament de la teoria d'estabilitat en sistemes dinàmics (funcions de Liapunov), així com per les nombroses contribucions en matemàtica física i teoria de la probabilitat, en la qual destaquem el TLC usant funcions característiques.

Liapunov va néixer a la ciutat russa Yaroslavl, fill de Mikhaïl Liapunov, un reconegut astrònom que treballava a l'observatori de la Universitat de Kazan. El seu pare el va educar a casa en diversos àmbits científics, però va morir de cop quan Aleksandr era jove. Més endavant, Aleksandr va estudiar al departament de matemàtiques i física de la Universitat de Sant Petersburg, on va conèixer a Màrkov. Al departament de matemàtiques, hi impartia classes Chebyshev en aquell moment. Liapunov es va graduar l'any 1880, i quatre anys més tard va obtenir el màster en matemàtica aplicada gràcies a culminar una complexa tesis sobre cossos celestes. En aquell moment, ja començava a estudiar la teoria de la probabilitat, i a destacar notablement pels seus avenços i qualitat d'investigació.

L'any 1885 va esdevenir professor a la Universitat de Járkov (actualment Ucraïna), on va impartir classes de mecànica teòrica, equacions diferencials i teoria de la probabilitat.

El 1892 va completar el seu doctorat, passant així a ser catedràtic en aquesta universitat. Més endavant, en morir Chebyshev l'any 1894, Liapunov va rebre el càrrec de professor principal en matemàtica aplicada a la Universitat de Sant Petersburg, on es dedicaria sobretot a investigar i a fer docència. Per les seves grans contribucions, el 1900 va ser escollit membre de l'Acadèmia Russa de Ciències.

Pel que fa a la teoria de la probabilitat, les investigacions de Liapunov van jugar un paper clau en el desenvolupament del Teorema del Límit Central, doncs és tal com el va plantejar ell que s'estudia avui dia. A finals del segle XIX/principis del XX, Liapunov descrivia molt acuradament la situació general i feia un resum de l'estudi de les probabilitats i del TLC en aquell moment:

Chebyshev havia provat en els seus estudis de teorema de pas al límit que es podia obtenir el TLC formulat per Laplace i Poisson. Aquest teorema era objecte de múltiples estudis, i Chebyshev va ser el primer en obtenir una demostració rigorosa per condicions més generals que les de Laplace i Poisson. Tanmateix, la prova de Chebyshev necessitava ésser complementada adequadament, cosa que Màrkov va aconseguir a la perfecció. Emperò, la demostració de Màrkov no acabava de convèncer a Liapunov, que considerava el mètode dels moments massa complicat i incòmode per treballar. És per això que Liapunov va investigar un mètode per provar el TLC de forma còmode i més general que els mencionats matemàtics. És així com va presentar els seus resultats usant el mètode de les funcions característiques, independent de qualsevol teoria prèvia.

La funció característica d'una variable aleatòria x és l'esperança matemàtica de la variable e^{itx}:

$$\varphi_x(t) = E(e^{itx}) = \int_{-\infty}^{+\infty} e^{itx} dF(x)$$

on $F(x)$ és la funció de distribució de x, i la integral convergeix $\forall t \in \mathbb{R}$.

El mètode de les funcions característiques és més general que el dels moments, i ràpidament va prendre més importància que aquest, ja que les funcions característiques existeixen per qualsevol variable. En canvi, els moments de qualsevol ordre no sempre existeixen. Una altra propietat que converteix aquestes funcions en eines molt útils és el producte, això és: $\varphi_{x+y}(t) = \varphi_x(t)\varphi_y(t)$.

La primera generalització i prova del TLC que fa Liapunov es troba en el seu primer article de teoria de la probabilitat, de l'any 1900:

Teorema 3.11. *Considerem una seqüència infinita de variables aleatòries independents $x_1, x_2, x_3, \ldots$, que prenen valors reals. Llavors, assumint l'existència de l'esperança de les variables $x_i, x_i^2, |x_i|^3$ per $i = 1, 2, 3, \ldots$ i anomenant-la α_i, a_i, l_i respectivament. Llavors, definim $A = a_1 - \alpha_1^2 + a_2 - \alpha_2^2 + \cdots + a_n - \alpha_n$, i denotem per L^3 el màxim de les n quantitats $l_1, l_2, \ldots, l_n$. Aleshores, si el valor $\frac{L^2}{A} n^{2/3}$ tendeix a zero quan $n \to \infty$, la probabilitat de les desigualtats*

$$z_1\sqrt{2A} < x_1 - \alpha_1 + x_2 - \alpha_2 + \cdots + x_n - \alpha_n < z_2\sqrt{2A}$$

per valors arbitraris z_1 i z_2 ($z_2 > z_1$) tendeix al límit

$$\frac{1}{\sqrt{\pi}} \int_{z_1}^{z_2} e^{-z^2} dz$$

quan $n \to \infty$. [[21], pp. 184-185]

Veiem doncs, que l'enunciat de Liapunov del TLC només requereix l'existència dels moments/esperances dels tres primers ordres, i no de tots els ordres, com era el cas de la versió del teorema de Màrkov.

Liapunov encara va obtenir una forma més genèrica del TLC, reduint més les restriccions en les condicions inicials, i obtenint així el teorema definitiu de pas al límit que porta el seu nom. El trobem en l'article de 1901 titulat *Nouvelle forme du théorème sur la limite de probabilité*, i és el següent:

Teorema 3.12. *Sigui δ un nombre positiu i d_i l'esperança matemàtica de la variable $|x_i - \alpha_i|^{2+\delta}$. Llavors, donat que existeix un valor δ tal que el quocient*

$$\frac{d_1 + d_2 + \cdots + d_n}{\sqrt{a_1 + a_2 + \cdots + a_n}} = \frac{\sum_{i=1}^{n} M(x_i - \alpha_i)^{2+\delta}}{\left[\sum_{i=1}^{n} D(x_i)\right]^{1+\delta/2}}$$

tendeix a zero quan $n \to \infty$, la probabilitat de les desigualtats

$$z_1 < \frac{x_1 - \alpha_1 + x_2 - \alpha_2 + \cdots + x_n - \alpha_n}{\sqrt{2(a_1 + a_2 + \cdots + a_n}} < z_2$$

tendeix al límit

$$\frac{1}{\sqrt{\pi}} \int_{z_1}^{z_2} e^{-z^2} dz$$

quan $n \to \infty$. Pel que fa a la notació: $x_1, x_2, x_3, \ldots$ és una seqüència infinita de variables aleatòries independents; $\alpha_1, \alpha_2, \alpha_3, \ldots$ són les seves esperances matemàtiques; $a_1, a_2, a_3, \ldots$ són les esperances de les variables $(x_1 - \alpha_1)^2, (x_2 - \alpha_2)^2, (x_3 - \alpha_3)^2, \ldots$; i z_1, z_2 són nombres donats tals que $z_2 > z_1$. [[21], pp. 223-224]

Per tant, la diferència substancial entre les condicions de Màrkov i Liapunov és la següent: Màrkov requereix que $\forall p > 2$

$$\lim_{n \to \infty} \frac{\sum_{i=1}^{n} M(x_i^p)}{\left[\sum_{i=1}^{n} D(x_i)\right]^{p/2}} = 0$$

mentre que Liapunov només existeixi un sol valor $\delta > 0$ tal que

$$\lim_{n \to \infty} \frac{\sum_{i=1}^{n} M(x_i - \alpha_i)^{2+\delta}}{\left[\sum_{i=1}^{n} D(x_i)\right]^{1+\delta/2}} = 0$$

En resum, degut a la rigorositat i eficàcia del plantejament i prova, el teorema del límit central de Liapunov ha esdevingut fonamental en la teoria de la probabilitat, i permet explicar el perquè moltes variables aleatòries segueixen distribucions normals. Això té múltiples aplicacions pràctiques: errors en mesures, índexs de qualitat, mides i pesos de productes, i moltes altres quantitats físiques que són propenses a patir canvis aleatoris.

Pel que fa a l'aplicació del teorema de Liapunov a variables dependents, no va ser fins al 1927 que S. N. Bernstein ho va aconseguir demostrar.

L'any 1947, durant una conferència a la Universitat de Moscou, Kolmogórov va fer una exposició amb el nom de *The role of Russian science in the development of probability theory.* En aquesta, va destacar la cabdal importància de Chebyshev, Markov i Liapunov per al desenvolupament de la teoria de la probabilitat. Va afirmar que el reconeixement i avaluació de la comunitat internacional no els va arribar fins als anys vint o trenta del segle XX, degut en part al fet que l'escola de Sant Petersburg treballava de forma independent i diferent a l'estadística predominant a nivell mundial. El fet que, durant la segona meitat del segle XIX, les investigacions en física a Rússia estiguessin molt poc avançades, va afectar a les aplicacions de teoria de probabilitats que es van estudiar en l'escola matemàtica de Sant Petersburg. Tot això, emperò, no treu importància ni validesa a l'enorme contribució en probabilitats i estadística dels mencionats matemàtics russos.

4 Inicis del segle XX

Gràcies a les contribucions de l'escola russa de Sant Petersburg, la teoria de la probabilitat va avançar molt significativament, també en camps més aplicats com la física estadística. Tanmateix, això va contrastar amb el fet que alguns matemàtics apliquessin les probabilitats a àmbits sense cap fonament científic i amb el subjectivisme a l'ordre del dia, com seria la política o la justícia. Aquesta diversificació pseudocientífica de la teoria de la probabilitat va provocar la pèrdua de la rigorositat i validesa d'aquesta teoria en les matemàtiques, a més de l'aparició de variades interpretacions de la mateixa. Això donaria peu a la necessitat d'una axiomatització de la teoria de la probabilitat, per així convertir-la en una branca de les matemàtiques amb la rigorositat que li correspon. En aquest apartat del projecte ens centrarem en presentar l'evolució i establiment de la mencionada axiomatització.

4.1 Pavel Nekrasov

Abans d'entrar en els detalls i protagonistes d'aquesta axiomatització, repassem breument un dels principals responsables de l'aplicació de les probabilitats a àmbits pseudocientífics i polítics: el matemàtic rus Pavel Nekrasov (1853-1924). Nekrasov va estudiar matemàtiques a la Universitat de Moscou, entrant l'any 1874. L'any 1885, després d'obtenir el doctorat, va esdevenir professor associat a aquesta mateixa universitat, i 5 anys més tard professor titular. El 1893 va ascendir al càrrec de rector, que mantindria durant força anys. Des de 1898, es va començar a dedicar exclusivament a tasques administratives, de gestió i de perfil polític: va ser nomenat superintendent del Districte Educatiu de Moscou, i més endavant (el 1905) va esdevenir membre del Consell Científic del Ministeri d'Educació Pública de Rússia, a Sant Petersburg.

Nekrasov era un ferm defensor de l'idealisme filosòfic. Aquest defensa la intangibilitat de les coses (les idees, l'esperit, la confiança), és a dir allò que no podem veure, tocar o sentir; defensa que la realitat que percebem és una construcció de la nostra ment i és immaterial. S'oposa frontalment al materialisme (del qual Màrkov és partidari), que defensa la tangibilitat de les coses i la matèria que podem tocar, defensa que la nostra realitat és allò que percebem físicament. Aquesta visió idealista va portar a Nekrasov a tergiversar el vertader significat de les probabilitats en les matemàtiques, tal com es pot intuir en les seves paraules del següent fragment:

A slave feels a stationary dependence on his master, a criminal on the court and police, and so on, and the measure of this dependence may be given by means of probability theory. [...] Probability theory gives a numerical measure of stationary, as well as nonstationary effects of dependences. [[28], pp. 19-21]

Una de les seves obres cabdals és *Probability Theory* [27], publicada el 1896, on s'exposen les classes que va impartir en probabilitats a la Universitat de Moscou, entre d'altres. S'hi poden trobar moltes aplicacions injustificades i pseudocientífiques de la teoria de la probabilitat, que es confonen amb nocions bàsiques de la mateixa. Entre aquestes aplicacions injustificades trobem creences religioses i problemes socials relacionats amb el tsar. És per això que va ser una obra fortament criticada per prominents matemàtics de l'escola russa, com Màrkov i Liapunov. De fet, aquests dos, juntament amb més matemàtics, van crear una comissió en l'Acadèmia de Ciències per a contrarestar els efectes negatius de la doctrina de Nekrasov sobre la comunitat matemàtica russa. I és que Nekrasov, com a membre del Consell Educatiu de Rússia, volia introduir un pla d'estudi de les probabilitats tal com les entenia ell als instituts, amb el fi últim d'educar en l'obediència inqüestionable del totpoderós tsar. Evidentment, aquesta malversació de les probabilitats era molt mal vista per Màrkov i companyia, que van presentar un escrit amb afirmacions com la següent:

Nekrasov's views have long been known to mathematicians. However, as long as they were confined to specialized mathematical journals, they could be considered harmless. [...] It is this Comission's opinion that the above mentioned blunders and... abuse of mathematics with the preconceived aim of transforming this science into a tool for religious and political persuasion... will cause irreparable harm to education. [[9], p. 178]

Veient doncs com Nekrasov "enmascarava" les seves deduccions pseudocientífiques en la teoria de la probabilitat, cada vegada era més evident la necessitat de re-avaluar els fonaments lògics d'aquesta teoria, per assegurar-ne indubtablement la seva posició com a disciplina genuïnament matemàtica.

L'ambigüitat existent en la interpretació de la teoria de la probabilitat va provocar que grans matemàtics caiguessin en l'error de trobar aplicacions injustificades d'aquesta teoria, actuant indirectament en detriment de la mateixa. Entre aquests trobem el matemàtic i polític francès Emile Borel (1871-1956).

4.2 Émile Borel

Borel va néixer a la regió d'Occitània el 1871, per més endavant estudiar matemàtiques a l'Escola Normal Superior de Paris. Va ser un dels pioners de la teoria de la mesura, i de les aplicacions d'aquesta a la teoria de la probabilitat. Fins i tot va establir una connexió entre la geometria hiperbòlica i la relativitat espacial, amb un treball seu de 1913. El 1914 va publicar un important llibre, anomenat *Le Hasard* [4], on tracta primer les lleis fonamentals de la teoria de la probabilitat, per després aplicar-la a la física, biologia, i més ciències. Tot i la gran qualitat del llibre i les aportacions que fa, són també remarcables les interpretacions errònies i injustificades de problemes relacionats amb les probabilitats. Un exemple d'una interpretació totalment arbitrària de la probabilitat és el següent:

Imagine one thousand Parisians passing by a seven story immovable property; they all agree to call it a house; however, they refuse to call a stone structure serving as a shelter to two rabbits and three hens a house. Let us consider an average structure (of these two); here opinion may be divided; if 748 out of 1000 voters call this structure a house, it would therefore be correct to assert that the probability that this structure is a house is 0.748 and the opposite probability is 0.252 [[4], p. 88]

Aquest exemple denota l'ambigüitat i vaguetat de la definició de probabilitat en aquell moment. Entre les diverses aplicacions que va fer Borel en aquest llibre de la teoria de la probabilitat, en trobem a nivell social i moral. Per exemple, dona valors probabilístics a l'egoisme i l'altruisme entre veïns. Per tant, amb aquestes afirmacions d'un prominent matemàtic com és Borel, s'evidencia que calia establir una metodologia rigorosa en probabilitats. De fet, Borel n'era conscient, de la necessitat d'axiomatitzar aquesta branca matemàtica. I a continuació presentem un dels predecessors dels fundadors d'aquesta requerida axiomatització de la teoria de la probabilitat: el reconegut matemàtic i físic francès Henri Poincaré (1854-1912).

4.3 Henri Poincaré

Va néixer el 1854 a Nancy (França), en el si d'una família molt benestant i influent: el seu pare era professor universitari, el seu cunyat un important filòsof, i un dels cosins va ser president de França. Durant la seva infantesa, va ser educat a casa per la seva mare, degut a l'afectació que li va provocar contraure la diftèria. Posteriorment, en assistir a l'escola, ja va destacar en totes les matèries des de ben petit, i sobretot en matemàtiques.

Va estudiar matemàtiques a la prestigiosa *École Polytechnique* de Paris, entre 1873 i 1876. A continuació, va estudiar enginyeria de mines a l'*École des Mines*, graduant-se l'any 1879. L'any 1879 va obtenir el doctorat, dirigit per Charles Hermite, i que tractava sobre les equacions diferencials. Poc després, va començar a exercir com a professor a la Universitat de Caen, investigant principalment en l'àmbit de les matemàtiques. Pel que fa a la mineria, sempre va anar-hi treballant, fent algunes obres públiques i dirigint grups de miners. A partir del 1881 i per la resta de la seva carrera, va exercir com a professor a la Universitat de Paris (la Sorbona), on ocuparia diverses càtedres en física i matemàtiques. Algunes de les seves contribucions més significatives són: el problema dels tres cossos, investigacions en teoria de la relativitat i en la teoria de la gravetat (que va precedir a la relativitat general). El 1887 va ser escollit membre de l'Acadèmia de Ciències de França, i el 1906 n'esdevindria el president. Poincaré és descrit com un polímata, doncs a més de matemàtic i físic, va ser un destacat científic i filòsof.

Tot i ser només una petita part de la seva contribució en les matemàtiques, els treballs de Poincaré en teoria de la probabilitat són molt destacables. Els recull en el seu llibre *Calcul des Probabilités* [30], publicat el 1912, i que es considera un dels llibres més rigorosos i interessants en aquest àmbit a principis del segle XX. A continuació en discutim els principals punts.

Poincaré defineix els esdeveniments aleatoris de la següent forma:

If a minute cause which escapes our notice determines a considerable effect that we cannot miss, then we say that this effect is due to chance. If we had an exact knowledge of the laws of nature and the position of the universe at the initial moment, we could predict exactly the position of that same universe in a succeeding moment. [...] It may happen that small differences in the initial conditions produce very great ones in the final phenomena. A small error in the former will produce an enormous error in the latter. Prediction becomes impossible, and we have a fortuitous phenomenon. [[30], pp. 1-5; 150]

Poincaré presenta diversos exemples de successos aleatoris en que una petita variació en les condicions inicials dona lloc a diferències apreciables en el resultat: l'equilibri inestable d'un con aguantant-se pel seu vèrtex, fenòmens meteorològics, les col·locacions dels planetes petits en les constel·lacions, o el joc de la ruleta. Per altra banda, també investiga un segon tipus de successos aleatoris: aquells en que el resultat depèn de multitud de causes complexes. Alguns exemples en són la teoria cinètica de gasos, la distribució de les gotes de pluja en una superfície, errors aleatoris en les observacions, etc. Finalment, considera un tercer tipus de successos aleatoris, que són en essència combinació dels dos tipus anteriors.

Poincaré es qüestiona si l'atzar segueix unes lleis concretes, i arriba a la conclusió que sí, ja que la distribució de l'atzar és contínua en la naturalesa. Ho exemplifica amb l'exemple d'una ruleta: la probabilitat de que en la primera tirada s'obtingui un valor comprès entre a i $a + \epsilon$ és la mateixa que estigui comprès entre $a + \epsilon$ i $a + 2\epsilon$, per una ϵ molt petita. I aquesta és una propietat de les funcions analítiques contínues. Per altra banda, Poincaré defensa que la ciència és determinista a priori, i assigna a la probabilitat aquells àmbits que encara no han sigut investigats, on el nostre coneixement és incomplet. En la mateixa línia, defineix l'atzar com una mesura de la nostra ignorància. Pel que fa a la definició clàssica de probabilitat, Poincaré l'analitza amb les següents paraules:

> *How can we determine that all the cases are equally probable? Mathematical determination is not possible in this case; in each application we must put conditions and stipulate that we shall consider these particular cases as equiprobable. These assumptions are not completely arbitrary, but they may escape the mathematician, if he does not analyze them after they have been made.* [[30], p. 28]

Per tant, es pot concloure que Poincaré, en la seva contribució en la teoria de la probabilitat, demanava també un establiment de fonaments rigorosos per aquesta. La gran qualitat dels seus estudis en aquest àmbit es veu lleugerament afeblida per la visió d'idealisme filosòfic amb la que planteja molts dels problemes.

Amb els autors tractats, s'evidencia clarament la necessitat d'unificar criteris i establir una base en teoria de la probabilitat. L'evolució d'aquesta a principis del segle XX requeria un repàs dels fonaments lògics sobre la que es construïa, i a això se li suma el fet que la física estadística estava patint notables avenços. Calia doncs una clarificació i justificació de les nocions probabilístiques, i la teoria clàssica de Laplace era clarament insuficient i inadequada. Establint els axiomes, que postulen la base de la teoria, es permet que totes les altres proposicions derivin d'aquests. Així, l'axiomatització engloba la totalitat d'objectes matemàtics d'una teoria concreta, i de fet Lobatxevski va ser un dels promotors d'aquest mètode.

4.4 Serguei Bernstein

Serguei Bernstein (1880-1968) va desenvolupar la primera axiomatització detallada de la teoria de la probabilitat, establint així les primeres bases formals d'aquesta. Més endavant, seria substituïda per l'axiomatització plantejada per Andrei Kolmogórov. Bernstein va néixer a Odessa (abans Rússia, actualment Ucraïna) el 1880, i era d'origen jueu. Va defensar la seva tesi doctoral a la Universitat de Paris, tractant sobre equacions diferencials. A part de teoria de la probabilitat, també va contribuir notablement en geometria diferencial i teoria de l'aproximació.

La seva obra cabdal en probabilitats és la titulada *Probability Theory* [2], publicada el 1927. Va ser (i encara és) un llibre referent en aquesta teoria durant molts anys i en diversos àmbits científics, i és aquí on Bernstein presenta la detallada axiomatització de probabilitats, que presentem a continuació.

Primer de tot, Bernstein exposa que la ocurrència d'un esdeveniment no sempre es pot predir amb absoluta certesa. Sigui un succés pertanyent a la classe A, i un conjunt qualsevol de condicions α, llavors només en el cas que aquestes siguin poques i es puguin observar, podrem afirmar que el fenomen ocorre amb absoluta certesa. Si això no es compleix, com passa en general, estem davant d'un esdeveniment aleatori. És llavors quan Bernstein es planteja introduir un conjunt més petit de variables β, en comptes de α, de manera que l'esdeveniment A adquireixi ara una certa probabilitat.

També introdueix una notació bàsica: si l'esdeveniment a dóna lloc a l'esdeveniment A, a s'anomena un cas particular de A. Si A pot ocórrer sense el seu cas particular A_1, llavors A_1 s'anomena un cas particular de A en el sentit estricte. En cas contrari, A_1 és un cas particular de A en el sentit ampli.

A continuació s'exposen els tres axiomes que desenvolupa Bernstein:

1. **L'axioma de comparació de probabilitats:** si a és un cas particular de A en el sentit estricte, llavors $P(a) < P(A)$; en canvi, si pels esdeveniments a_1 i A es compleix la desigualtat $P(a_1) < P(A)$, llavors $P(a_1) = P(a)$, on a és un cas particular de A en el sentit estricte. D'aquest primer axioma se'n poden deduir dos axiomes:

   - La probabilitat d'un esdeveniment segur és major que la d'un esdeveniment possible.
   - La probabilitat d'un esdeveniment possible és major que la d'un esdeveniment impossible.

   Amb aquest axioma, per tant, s'estableix un criteri per a la probabilitat de tots els esdeveniments segurs, que és igual i màxima, mentre que la dels esdeveniments impossibles és sempre igual i mínima.

2. **L'axioma d'esdeveniments incompatibles:** Siguin els esdeveniments A i A_1 incompatibles, i el mateix per B i B_1. A més, $P(A) = P(B)$ i $P(A_1) = P(B_1)$. Llavors, la probabilitat de C, que consisteix en l'ocurrència de l'esdeveniment A o el A_1, és igual a la probabilitat de C_1, que al seu torn consisteix en l'ocurrència de B o B_1. És a dir, $P(A \sqcup A_1) = P(B \sqcup B_1)$.

   Aquest axioma indica doncs que la probabilitat de que succeeixi un dels dos esdeveniments disjunts ve determinada per la probabilitat de cadascun d'ells separadament.

3. **Axioma de combinació d'esdeveniments:** Sigui α un cas particular de A, llavors la probabilitat de α sota condicions donades depèn només de la probabilitat de A sota aquestes mateixes condicions, així com de la probabilitat adquirida per α quan A ocorre. Això implica que si α_1 és un cas particular de A_1, llavors $P(\alpha) = P(\alpha_1)$ si $P(A) = P(A_1)$ amb les condicions donades, i.e. donat que la probabilitat de α després de que A ocorri és igual a la probabilitat de α_1 després que A_1 ocorri.

   Aquest axioma es pot expressar de forma alternativa: la probabilitat de la combinació d'esdeveniments A i B (sota les condicions donades) depèn només de la probabilitat de A (sota les mateixes condicions) i de la probabilitat adquirida per B després de que ocorri A.

   Per esdeveniments independents, aquest axioma afirma: siguin A i B independents, llavors la probabilitat de la combinació de A i B depèn només en les condicions inicials d'aquests dos. Això s'expressa com

   $$(A, B) = \Phi[(A), (B)_A] = \Phi[(B), (A)_B]$$

   on (A) és la probabilitat de A; $(B)_A$ és la probabilitat de B després que ocorri A; (A, B) és la probabilitat de la combinació de A i B; Φ és una funció fixada.

4.5 Richard von Mises

Els defectes i subjectivisme de la teoria clàssica de la probabilitat eren coneguts per molts matemàtics a principis del segle XX. Entre els diferents crítics amb aquesta teoria clàssica, hi trobem al reconegut matemàtic i físic austríac Richard von Mises (1883-1953). Va incidir constantment en les clares deficiències de la teoria clàssica de la probabilitat, creant al seu torn la famosa teoria freqüencial de la probabilitat (també anomenada probabilitat freqüencial). Aquesta es basa en la idea que el concepte de probabilitat només té sentit en parlar de fenòmens de masses/col·lectius. En la mateixa línia, von Mises defensava que la teoria de la probabilitat no era una disciplina matemàtica, sinó simplement una branca de la ciència que investigava fenòmens del món real, doncs a aquests sempre se'ls pot associar una component probabilística. En canvi, les matemàtiques no lidiaven amb aquest tipus de problemes realistes. Aquesta visió era característica dels defensors de la probabilitat freqüencial, i s'oposava al corrent dominant en teoria de la probabilitat, que la incloïa dins el regne de les matemàtiques. Amb els anys i posterior desenvolupament, aquesta última seria la visió correcta i acceptada.

Abans d'exposar l'axiomatització proposada per von Mises en el context de la probabilitat freqüencial, fem un breu repàs de la seva biografia. Von Mises va néixer a Lemberg, ciutat austro-hongaresa (actualment Lviv, Ucraïna), en el si d'una família benestant jueva. Va estudiar a la Universitat de Viena, graduant-se en física, matemàtiques i enginyeria. El 1905, encara com a estudiant, va publicar un important article sobre geometria de corbes, i l'any 1908 va rebre el doctorat. Això li va permetre començar a impartir classes d'enginyeria a la ciutat de Brno. El 1909 va esdevenir professor de matemàtiques a Estrasburg. Més endavant, durant la Primera Guerra Mundial, va combatre amb l'exèrcit austro-hongarès com a pilot de combat. Després d'aquesta guerra, va exercir com a professor i director a l'institut de matemàtiques aplicades de la Universitat de Berlin. Tanmateix, amb l'aixecament del partit nacional-socialista alemany, von Mises es va veure obligat a exiliar-se a Turquia, on esdevindria professor a la Universitat de Istanbul. El 1939 va acceptar una posició de professor als Estats Units, on el 1944 començaria a impartir classes en Aerodinàmica i Matemàtiques Aplicades a Harvard. Les seves contribucions omplen un rang ampli de ciències, enginyeria i matemàtiques.

Pel que fa a la teoria de la probabilitat freqüencial, von Mises la fonamenta en el concepte de col·lectivitat. Això és, una seqüència infinita K d'observacions similars, cadascuna de les quals determina un punt que pertany a un espai finit R. D'aquesta manera, només es pot parlar de probabilitat si existeix una col·lectivitat, i.e. un nombre definit d'esdeveniments. Partint d'això, aquesta col·lectivitat ha de complir dues condicions, els dos axiomes definits per von Mises:

1. El primer axioma postula l'existència del límit $\lim_{n\to\infty} \frac{m}{n} = P(S)$, on m és el nombre de casos favorables/esdeveniments en el total de les n observacions, i $S \subset R$ és un subconjunt. Per tant, cal que existeixi un límit de la freqüència m/n.

2. El segon axioma requereix que existeixi un límit anàleg a l'anterior, també amb valor $P(S)$, per qualsevol subseqüència K' de K. K' ha de complir la condició de que sempre es pot decidir si la n-èssima observació de K pertany a K' sense saber el resultat d'aquesta observació en particular.

Von Mises entenia aquests axiomes com a propietats de la col·lectivitat, i no com a axiomes d'una teoria matemàtica com a tal. A partir d'aquests dos axiomes, construeix els fonaments de la teoria de la probabilitat. Comença definint la probabilitat com el límit finit de la freqüència m/n. Segons von Mises, la probabilitat no és una propietat objectiva del fenòmen, sinó que els esdeveniments adquireixen probabilitat un cop s'ha realitzat l'experiment. D'aquesta manera, se li treu a la probabilitat el significat de valor numèric objectiu i característic de fenòmens reals. Aquesta visió de la teoria de la probabilitat va ser àmpliament criticada per molts matemàtics, al·legant que tal definició de probabilitat implicava una barreja d'elements teòrics i empírics, cosa que s'ha d'evitar en les axiomatitzacions. Algun exemple d'aquestes crítiques és la d'Andrei Kolmogórov:

...the limiting transition $m/n \to p$ cannot have real meaning. Moreover, the formulation of the stability of frequencies principle using this limiting process requires the availability of admissible methods for determining infinite sequences of trials which can be a mere mathematical fiction. [[18], pp. 274-275]

Tot i les importants reticències a la comunitat matemàtica, el mètode de probabilitat freqüencial de von Mises va ser àmpliament estès en les ciències, particularment en la física. Tot i que actualment, els seus postulats ja no s'usen enlloc. El problema principal de la teoria de la probabilitat freqüencial és que la seva interpretació no és prou abstracta i comporta un significat massa concret. I, tot i que sembli contradictori, com més abstracte és un sistema axiomàtic, més simple és. En canvi, si és molt concret i significatiu, es complica molt i costa obtenir deduccions de la teoria en qüestió. I aquest és el defecte fonamental de les teories de probabilitat freqüencial. S'exemplifica amb el fet que les majors contribucions en probabilitats les van fer matemàtics que mai van formar part de "l'escola freqüencial".

Igual que amb la probabilitat freqüencial, hi va haver altres intents d'establir les bases de la teoria de la probabilitat amb enfocaments alternatius. Entre ells cal destacar la interpretació subjectiva de la probabilitat establerta per Bruno de Finetti (1906-1985), així com el desenvolupament dels graus de probabilitat per part de Harold Jeffreys (1891-1989). De Finetti defensava que la probabilitat és una quantitat essencialment subjectiva (la defineix cadascú), i en conseqüència també ho són les nocions de dependència, independència i equiprobabilitat. Pel que fa a Jeffreys, la seva teoria atorga a cada proposició una certa probabilitat. Cap d'aquestes teories alternatives va arribar a tenir molta acceptació per part de la comunitat matemàtica.

4.6 Aleksandr Khintxin

Durant aquestes primeres dècades del segle XX, juntament amb els intents d'axiomatitzar la teoria de la probabilitat, hi va haver grans avenços en la mateixa. I analitzant aquesta tendència en desenvolupament, Kolmogórov va ser capaç de desenvolupar una axiomatització que esdevindria la base de la probabilitat moderna i donaria lloc al seu desenvolupament fins a l'actualitat.

Els treballs de Kolmogórov i Aleksandr Khintxin (1894-1959) en la llei dels grans nombres van donar lloc a l'escola de Moscou en teoria de la probabilitat. I a partir de la dècada de 1920, les investigacions en aquesta branca de les matemàtiques van entrar en contacte amb la teoria de conjunts i de funcions. Tot això assentaria els precedents per l'axiomatització de Kolmogórov l'any 1933.

Repassem ara breument la figura de Khintxin, reconegut matemàtic rus, i la seva contribució en la teoria de la probabilitat a través de la reformulació i generalització de la llei dels grans nombres (amb la llei del logaritme iterat). Va néixer el 1894 al poble rus de Kondrovo, i va estudiar matemàtiques a la Universitat de Moscou, graduant-se l'any 1916. Sis anys després, va esdevenir professor a aquesta mateixa universitat, i va mantenir aquesta posició durant la resta de la seva vida. El 1939 va ser escollit membre de l'Acadèmia de Ciències de la URSS. Entre les seves contribucions en teoria de la probabilitat, la més destacable és la formulació de la llei del logaritme iterat l'any 1923, generalitzant i culminant els avenços de la llei dels grans nombres. Abans de presentar la de Khintxin, fem un breu repàs de l'evolució d'aquesta llei fins al moment:

- Segons el teorema de Bernouilli, per $n \to \infty$ i $\forall \epsilon > 0$, tenim que

$$P(|m/n - p| < \epsilon) \to 1$$

- El 1909, E. Borel prova que per $p = 1/2$,

$$P(\lim_{n\to\infty} \frac{m - np}{n} = 0) = 1$$

  és a dir que la diferència $m - np$ per n gran és molt petita comparada amb n, i es pot afirmar amb absoluta certesa. En altres paraules, que $m - np = o(n)$.

- El 1917, F.P. Cantelli va generalitzar el resultat de Borel per $p \in (0, 1)$
- El 1913, F. Hausdorff va obtenir la següent aproximació pel teorema de Bernouilli: amb probabilitat 1, $m - np = o(n^{\epsilon+1/2})$, on $\epsilon > 0$ és un nombre arbitrari.
- El 1914, G.H. Hardy i J.E. Littlewood van provar que, amb probabilitat 1, $m - np = o(\sqrt{n \ln n})$.

Finalment, l'any 1924, Khintxin va formular el teorema del logaritme iterat:

Teorema 4.1. *Sigui p la probabilitat de l'esdeveniment A en cadascun dels n experiments independents, llavors el nombre m d'ocurrències de l'esdeveniment A en els n experiments satisfà que*

$$P(\lim_{n\to\infty} sup \frac{|m - np|}{\sqrt{2npq \ln(\ln n)}} = 1) = 1$$

on $q = 1 - p$. Per tant, la funció $\sqrt{2npq \ln(\ln n)}$ és la fita superior exacta de la variable aleatòria $|m - np|$.

Durant els anys posteriors, es van fer diversos treballs en la llei dels grans nombres per part de Khintxin i Kolmogórov, gràcies als quals la teoria de la probabilitat va avançar molt. Es va establir una relació entre aquesta teoria i la teoria de conjunts i funcions, cosa que va permetre aplicar correctament la llei dels grans nombres, després de molts anys d'intents fallits en aquest sentit.

4.7 Andrei Kolmogórov

Va ser Andrei Kolmogórov (1903-1987), il·lustre matemàtic rus, qui va obtenir les condicions d'aplicabilitat de la llei dels grans nombres, en el següent teorema de 1926:

Teorema 4.2. *Una seqüència de variables aleatòries independents $\xi_1, \xi_2, \ldots, \xi n$ obeeix la llei (feble) dels grans nombres si, i només si, es satisfan les següents relacions quan $n \to \infty$:*

$$(1) \qquad \sum_{k=1}^{n} \int_{|x|\geq n} dF_k(x) \to 0$$

$$(2) \qquad \frac{1}{n} \sum_{k=1}^{n} \int_{|x|<n} x dF_k(x) \to 0$$

$$(3) \qquad \frac{1}{n^2}\sum_{k=1}^{n}\int_{|x|<n} x^2 dF_k(x) \to 0$$

On $F_k(x)$ *denota* $P(\xi_k - E\xi_k < x)$.

Amb aquest teorema, Kolmogórov va resoldre completament i definitivament un dels problemes centrals de la teoria de la probabilitat. Dos anys més tard, el 1928, Khintxin va provar que si les variables $\xi_1, \xi_2, \ldots, \xi n$ són independents i a més idènticament distribuïdes, llavors l'existència de l'esperança $E\xi_n$ és condició necessària i suficient per aplicar la llei (feble) dels grans nombres. Cal diferenciar aquesta de la llei forta dels grans nombres, postulada per Borel el 1909:

$$P\left(\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}(\xi_k - E\xi_k) = 0\right) = 1$$

Borel va provar-la pel cas de Bernouilli i a més $p = 1/2$, i més tard Kolmogórov i Khintxin ho van fer de forma més general per variables independents i dependents, respectivament.

Abans d'analitzar en detall l'axiomatització feta per Kolmogórov, presentem breument la seva biografia: va néixer a la ciutat russa de Tambov el 1903. La seva mare va morir en el part, i el seu pare el va abandonar per morir poc després a la guerra civil russa. Va créixer doncs sota la tutela de la seva tieta. Ja des de ben petit, va destacar molt notablement a nivell acadèmic. L'any 1920 va començar a estudiar matemàtiques a la Universitat de Moscou, on va adquirir gran reputació pels seus primers estudis en teoria de conjunts, teoria de la probabilitat i teoria de series de Fourier (1921-1922). El 1929 va obtenir el doctorat en física i matemàtiques a la mateixa Universitat de Moscou. El 1931 va convertir-se en professor d'aquesta universitat. El 1933 va publicar el llibre *Foundations of the Theory of Probability* [17], on s'estableixen les bases modernes de la teoria de la probabilitat, igualant-se així amb altres disciplines matemàtiques, i gràcies al qual va esdevenir un dels majors experts del món en aquesta teoria. El 1939, va ser escollit membre de l'Acadèmia Russa de Ciències. Durant tota la seva vida, a part d'investigar en matemàtiques i física, va estar molt relacionat amb la pedagogia universitària i també de primària.

Nosaltres ens centrem en l'anàlisi de l'axiomatització que va fer Kolmogórov. Des dels anys vint, havia aconseguit introduir progressivament la teoria de funcions a la teoria de la probabilitat, i tot això va desembocar en el mencionat llibre *Foundations of the Theory of Probability* el 1933. Aquí, apareixen analogies constants entre mesura d'un conjunt i probabilitat d'un esdeveniment, entre integrals i l'esperança matemàtica, entre ortogonalitat de funcions i independència de variables, i moltes d'altres. Procedim a exposar breument aquesta axiomatització:

Suposem que es poden fer infinites observacions d'un experiment, obtenint diferents resultats que depenen de l'atzar. El conjunt de possibles resultats forma un conjunt E, que s'anomena "conjunt dels esdeveniments elementals", i és el concepte bàsic de l'axiomatització de Kolmogórov. S'anomenen esdeveniments aleatoris als elements d'una col·lecció F de subconjunts de E, on F té estructura de σ-àlgebra. En conseqüència, aquí una variable aleatòria es defineix com una funció d'esdeveniments elementals. Això xoca amb el que hi havia establert fins al moment, on la variable aleatòria era la noció elemental de la teoria de la probabilitat. A més, Kolmogórov només considera una part dels esdeveniments aleatoris possibles.

Amb la definició de E i els subconjunts respectius, s'aconsegueix l'abstracció necessària per a estudiar la probabilitat no només des d'un punt de vista probabilístic i amb aplicacions reals, sinó també amb moltes d'altres aplicacions i derivacions en les matemàtiques. Amb tot això, els axiomes que postula Kolmogórov són:

1. Siguin A i B dos esdeveniments aleatoris de la col·lecció F. Llavors, els esdeveniments $A \cup B$, $A \cap B$ i $\bar{A} \cap \bar{B}$, estan també continguts en F.

2. La col·lecció F conté al conjunt E.

3. A cada conjunt A de F se li assigna un nombre positiu $P(A)$. Aquest nombre s'anomena "probabilitat" de l'esdeveniment A.

4. $P(E) = 1$.

5. Si A i B no tenen elements en comú, llavors $P(A \cup B) = P(A) + P(B)$.

6. En el cas particular que F sigui una col·lecció infinita: sigui la seqüència decreixent d'esdeveniments de F
$$A_1 \supset A_2 \supset \cdots \supset A_n \supset \cdots$$
tal que la seva intersecció és nul·la, es compleix que
$$\lim_{n \to \infty} P(A_n) = 0.$$

Com hem dit, aquesta axiomatització va proporcionar a la teoria de la probabilitat l'abstracció necessària per allunyar-la de la concepció clàssica i també de la interpretació freqüencial ideada per von Mises. Així, la definició de probabilitat adquiria noció matemàtica en la seva màxima essència, i a la vegada permetia passar d'un sistema formal a un procés real i tangible. Des d'aquesta formalització de Kolmogórov, la teoria de la probabilitat ha anat evolucionant fins a donar forma a la que estudiem avui dia.

5 Conclusions

Després de fer aquest treball, es pot concloure que s'han complert els objectius inicials amb èxit: s'ha traçat una línia temporal a través dels principals matemàtics que van participar en el desenvolupament de la teoria de la probabilitat al llarg del segle XIX i principis del XX, i s'ha entès l'evolució d'aquesta teoria en el mencionat període temporal, sobretot pel que fa al Teorema del Límit Central (TLC) i la llei dels grans nombres.

Els treballs de Laplace en teoria de la probabilitat van suposar un punt d'inflexió en l'estudi d'aquesta teoria, formalitzant-la degudament per primera vegada a la història. Tot i usar definicions de caire bastant clàssic, amb el llibre *Théorie analytique des probabilités* [19], publicat el 1812, Laplace va donar peu al desenvolupament matemàtic i formal de les probabilitats en els anys posteriors. Entre les importants contribucions que va fer, cal destacar el Teorema del Límit Central. Més endavant, Poisson provaria la llei dels grans nombres, generalitzant així el resultat previ de Bernouilli.

S'ha vist també que l'escola russa va jugar un paper fonamental en "desencallar", a mitjan del segle XIX, la teoria de la probabilitat de la visió clàssica de Laplace amb la que seguia enfocada. Entre els predecessors de l'escola russa cal destacar notablement les figures de Ostrogradski i Buniakovski. Els seus treballs donarien forma a l'escola de Sant Petersburg, dominant durant la segona meitat del segle XX i liderada per Chebyshev, Màrkov i Liapunov. El primer va obtenir una generalització molt important de la llei dels grans nombres, culminant així el treball previ de Bernouilli i després Poisson. Màrkov va seguir en la línia d'investigació de Chebyshev, però en aquest cas amb variables dependents. Pel que fa a Liapunov, aquest va ser l'autor del TLC tal com el coneixem avui dia.

Finalment, als inicis del segle XX es va evidenciar la necessitat d'axiomatitzar i establir unes bases formals i unificades per la teoria de la probabilitat. Això és perquè el significat d'aquesta s'estava començant a malinterpretar i tergiversar amb aplicacions injustificades de tot tipus: a l'àmbit social, jurídic, polític, etc. Entre els predecessors d'aquesta axiomatització cal destacar a Poincaré o Bernstein; i de la mateixa manera és remarcable la proposició (amb els anys s'ha vist que sense èxit) de la probabilitat freqüencial per part de von Mises com a enfocament alternatiu a la probabilitat clàssica. El punt culminant d'aquest treball ha sigut l'anàlisi dels treballs en probabilitats dels matemàtics russos Khintxin i Kolmogórov, fundadors de l'escola de probabilitat de Moscou durant la dècada de 1920. El primer va ser l'autor d'una formulació molt genèrica de la llei dels grans nombres. El segon va postular, en el llibre *Foundations of the Theory of Probability* [17] de 1933, els axiomes que assentarien d'una vegada per totes la teoria de la probabilitat de forma rigorosa, al seu torn donant lloc al desenvolupament d'aquesta fins a l'actualitat.

En conclusió, s'ha vist que l'escola russa va ser la gran protagonista del desenvolupament de la teoria de la probabilitat als segles XIX i XX, juntament amb alguns membres de l'escola francesa com Laplace i Poincaré. Les investigacions dels autors tractats en aquest treball permeten entendre la importància de la cooperació internacional i el treball en equip per al reeixit desenvolupament de qualsevol branca de les matemàtiques i de la ciència en general. De la mateixa manera, s'ha vist que els dos principals problemes de la teoria de la probabilitat: el TLC i la llei dels grans nombres, van evolucionar molt durant els segles XIX i XX fins arribar a la formulació actual, sempre amb la característica que cada nova aportació solia generalitzar-ne més la seva demostració.

A La física estadística

L'objectiu d'aquest annex és exposar breument els inicis i primers desenvolupaments de la física estadística, liderats per Boltzmann i Gibbs durant la segona meitat del segle XIX. Exposem abans la motivació del desenvolupament d'aquesta branca de la física. El 1827, el botànic anglès Robert Brown va detectar el moviment d'unes partícules minúscules suspeses en l'aigua d'una de les seves observacions. Més endavant, es va descobrir que tota partícula prou petita suspesa en un fluid es mou constantment seguint un patró impredictible. Aquest moviment es va anomenar moviment brownià (sobre el qual A. Einstein en va construir una teoria el 1905). És causat per l'impacte de molècules en moviment (caòtic) sobre partícules en suspensió. Va ser gràcies a nocions probabilístiques que es va poder desenvolupar una teoria consistent sobre el moviment brownià. Principalment degut al gran número de molècules que hi participen i la naturalesa caòtica del seu moviment tèrmic. Es pot suposar amb certa precisió que les molècules d'un gas estan uniformement distribuïdes en un volum determinat. Per tant, en mitjana, les desviacions de la distribució uniforme són negligibles. Hi ha algunes desviacions d'aquest comportament, però com més notables són aquestes, menys freqüentment es succeeixen. Tots aquests resultats, així com altres resultats de física molecular, s'obtenen bàsicament usant mètodes estadístics. Alguns exemples de resultats són la velocitat de les molècules, la seva energia cinètica mitjana, i moltes altres quantitats. El problema principal de la física estadística és la determinació dels valors mitjans de diverses quantitats físiques i l'establiment de lleis connectant aquests valors. Les consideracions estadístiques conformen un dels mètodes més usats en física. Va ser durant la segona meitat del segle XIX que les àrees d'aplicació de la teoria de la probabilitat van evolucionar molt, passant de simples estudis observacionals (demografia, justícia, etc.) a complexos anàlisis en l'àmbit de la física. I això va ser principalment gràcies als treballs dels promintents científics que s'exposen a continuació.

A.1 Ludwig Boltzmann

Ludwig Boltzmann (1844-1906) va ser un dels físics teòrics més importants de la segona meitat del segle XIX, i un dels fundadors de la física moderna, degut sobretot a la creació i desenvolupament de la física estadística. Les seves contribucions més destacables van ser la interpretació cinetico-molecular de la segona llei de la termodinàmica, així com la derivació de la interpretació estadística de l'entropia. Boltzmann va néixer a Viena el 1844, en el si d'una família benestant. Va cursar els primers anys d'estudi a la ciutat de Lunz, per més endavant doctorar-se en física a la Universitat de Viena l'any 1866. El 1869 va exercir com a professor de física a Graz, i posteriorment a Heidelberg i Berlin. Durant aquells anys va treballar amb prominents físics com Kirchoff i Helmholtz. El 1873 va acceptar una posició de docent en matemàtiques a Viena, però el 1876 va tornar a Graz com a catedràtic. En aquell moment, Boltzmann ja era molt conegut en la comunitat científica pel desenvolupament de l'estadística de Maxwell-Boltzmann per les velocitats de les molècules d'un gas, de 1871. El 1894, va tornar a la Universitat de Viena per impartir física teòrica. Allà va coincidir amb el seu màxim rival científic: Ernst Mach, amb qui discernia respecte la naturalesa dels àtoms. En les seves obres destaca la pionera aplicació dels mètodes probabilístics a la mecànica, cosa que li va permetre fonamentar teòricament les lleis de la termodinàmica i traçar el camí per al desenvolupament posterior d'aquesta. L'informació que presentem a continuació està extreta del llibre de Boltzmann titulat *Lectures on Gas Theory* [3], publicat per la Universitat de Califòrnia Berkeley l'any 1964.

L'estat dels objectes físics es caracteritza per la temperatura, la pressió, la densitat, etc. A cada estat li corresponen diverses distribucions de molècules i àtoms. Com més distribucions diferents té un estat particular, més elevada és la probabilitat W de trobar el sistema en aquest estat. Igualment, com més petita és W, més improbable que trobem el sistema en l'estat corresponent. Tot i això, fins i tot els estats més improbables s'esdevenen alguna vegada. Per exemple, que totes les molècules d'un gas es trobin en una meitat del volum que el conté. És molt improbable, però possible. De fet, Boltzmann ho va calcular: per un volum de $1cm^3$, la probabilitat respectiva és $(1/10)^{10^{10}}$.

Si no hi ha forces externes, el sistema evoluciona d'estats menys probables a més probables, arribant finalment a l'estat més probable, anomenat estat d'equilibri. Per tant, la probabilitat tendeix a un valor maximal, i el mateix passa amb l'entropia, la qual sempre augmenta. Boltzmann introdueix la funció H com a anàleg de l'entropia, que juga un paper fonamental en la física estadística. També considera

$$f(\xi, \eta, \varepsilon, t) d\xi d\eta d\varepsilon = f d\omega$$

per denotar el nombre de molècules m a l'instant de temps t, que tenen les components de la velocitat en cadascun dels tres eixos entre els límits

$$\xi \quad i \quad \xi + d\xi, \qquad \eta \quad i \quad d\eta, \qquad \varepsilon \quad i \quad \varepsilon + d\varepsilon$$

respectivament. Si es coneix la funció f, es pot determinar la distribució de velocitats. Anàlogament, considerem

$$F_1(\xi_1, \eta_1, \varepsilon_1, t_1) d\xi_1 d\eta_1 d\varepsilon_1 = F_1 d\omega_1$$

amb les components de la velocitat entre els límits

$$\xi_1 \quad i \quad \xi_1 + d\xi_1, \qquad \eta_1 \quad i \quad d\eta_1, \qquad \varepsilon_1 \quad i \quad \varepsilon_1 + d\varepsilon_1$$

llavors obtenim que

$$H = \int f \ln f d\omega + \int F_1 \ln F_1 d\omega_1.$$

A continuació, Boltzmann prova que la funció H no pot augmentar amb el temps, i la interpreta com un anàleg de l'entropia (amb un signe negatiu). El 1877 va relacionar la funció H amb la probabilitat d'una distribució donada. Això el va portar a adonar-se de la naturalesa estadística de la segona llei de la termodinàmica, la qual afirma que l'entropia augmenta en qualsevol procés físic de la naturalesa. A partir d'aquí, va obtenir una expressió que relaciona l'entropia S amb la probabilitat W: $S = k \ln W$, on $k = 1,3810^{-23} J/K$ és la famosa constant de proporcionalitat de Boltzmann. D'aquesta forma, es va provar que l'entropia és una mesura de la probabilitat d'un estat particular del sistema donat. Segons Boltzmann, l'increment és el canvi més probable de l'entropia. També va estudiar els processos reversibles de l'univers, a més d'opinar sobre la possible aplicació de l'estadística en les ones electromagnètiques. En línies generals, les idees de Boltzmann van ser clau per al desenvolupament de la física moderna, en particular per a la física estadística i més endavant la física quàntica, que requereix moltes nocions probabilístiques.

A.2 James Clerk Maxwell

James Clerk Maxwell (1831-1879) va ser el predecessor immediat de Boltzmann. Maxwell va ser un famós matemàtic i físic escocès. Nascut a Edimburg, va destacar ja des de ben petit per la seva capacitat intel·lectual. Amb 16 anys, va començar a estudiar a la Universitat d'Edimburg, on es va graduar en filosofia natural, moral i de la ment. El 1850, va anar a la Universitat de Cambridge, però el mateix any es va traslladar al Trinity College de Londres, on el 1854 es va graduar en matemàtiques. Simultàniament a aquests estudis, Maxwell va obtenir diversos resultats en electromagnetisme i òptica, gràcies als quals va obtenir molt reconeixement i nombrosos premis. El 1856, se li va atorgar la càtedra en filosofia natural al Marischal College, a Aberdeen, i el 1860 va esdevenir professor al King's College de Londres, on va obtenir destacables resultats en l'elasticitat dels sòlids i geometria. El 1861, va ser escollit membre de la Royal Society.

Partint de la hipòtesi que les molècules són sòlids elàstics, Maxwell va elaborar una teoria cinètica de gasos, basada en els treballs de Daniel Bernouilli, i que va avançar molt gràcies a Rudolf Clausius. Maxwell va aconseguir així aplicar les matemàtiques a la física, investigant sistemes amb un nombre infinit de molècules. Distribuint les molècules en grups segons les seves velocitats, Maxwell va desenvolupar una teoria de la distribució de velocitats entre 1859 i 1866. Més endavant, Boltzmann la generalitzaria. De fet, aquesta llei de distribució de velocitats s'anomena distribució de Maxwell-Boltzmann:

Sigui $\varphi(x)dx$ la probabilitat que la projecció de la velocitat d'una molècula en l'eix x es trobi entre x i $x + dx$, i anàlogament per les probabilitats $\varphi(y)dy$ i $\varphi(z)dz$. La probabilitat que el vector des de l'origen representant la velocitat estigui contingut entre x, y, z i $x + dx, y + dy, z + dz$, és

$$\varphi(x)\varphi(y)\varphi(z) = f(x^2 + y^2 + z^2) = e^{-k^2(x^2+y^2+z^2}$$

Aquesta distribució va esdevenir un important resultat en la física estadística, però tot i així Maxwell no va destacar pels seus treballs en aquest àmbit. Els primers en introduir arguments consistents de l'estadística a la física van ser Boltzmann i Gibbs.

A.3 Josiah Willard Gibbs

Josiah Willard Gibbs (1839-1903) va ser un físic nord-americà que va contribuir molt notablement en la fonamentació teòrica de la termodinàmica i origen de la física estadística. Nascut a Connecticut, va estudiar a la Universitat de Yale, on va obtenir el doctorat en enginyeria. El 1871 va esdevenir professor de física matemàtica a aquesta mateixa universitat. La seva investigació es va centrar en la termodinàmica i en la teoria del càlcul vectorial usant quaternions. Seguint la línia de Maxwell i Boltzmann, Gibbs va estudiar problemes de mecànica estadística, amb aplicacions directes en la termodinàmica. De fet, durant la dècada de 1880 va començar a impartir classes en física estadística a Yale. Va recollir els seus resultats en la publicació del llibre *Basic Principles of Statistical Mechanics* [14], de 1902. Aquest va permetre a la física estadística adquirir la dimensió lògica que li corresponia, després dels estudis primaris duts a terme per Maxwell i Boltzmann.

Gibbs tractava les propietats dels sòlids com les d'una col·lectivitat (o bé *ensemble*), que consisteix en un gran nombre de partícules sotmeses a les lleis de la mecànica. Gibbs va estudiar les propietats de les col·lectivitats usant mètodes de teoria de la probabilitat i va estudiar el paper que juga aquesta en la física. Això permet dur a terme un anàlisi de les propietats macroscòpiques de les substàncies. Gibbs connecta aquestes propietats amb les propietats mitjanes (estadísticament) de les col·lectivitats. Usa àmpliament la idea de Boltzmann, interpretant l'entropia com l'estat probable del sistema. També planteja el problema referent a la contradicció entre la no-invertibilitat termodinàmica, conseqüència de la llei de l'augment de l'entropia, i la reversibilitat temporal associada als processos mecànics i les lleis del moviment que els regeixen. Avui dia, aquest problema encara no ha estat resolt.

La necessitat del desenvolupament de la física estadística radicava en el problema fonamental de que totes les observacions físiques estan sotmeses a un cert error. Gibbs era de l'opinió que la física no estudia els resultats dels esdeveniments que ocorren amb absoluta certesa, sinó aquells que ocorren amb una elevada probabilitat. Cal remarcar però que alguns dels treballs de Gibbs no es van poder completar per falta del desenvolupament formal de la teoria de la probabilitat.

En el seu llibre *Basic principles of statistical mechanics*, Gibbs descriu els objectius d'aplicar els mètodes estadístics en la física: *"The usual point of view in the study of mechanics is that where the attention is mainly directed to the changes which take place in the course of time in a given system"* [[14], p. VII, Preface]. Gibbs estudia els casos que difereixen infinitessimalment de l'estat real del sistema:

"We may imagine a great number of systems of the same nature, but differing in the configurations and velocities which they have at a given instant, and differing not merely infinitesimally, but it may be so as to embrace conceivable combination of configuration and velocities. And here we may set the problem, not to follow a particular system through its succession of configurations, but to determine how the whole number of systems will be distributed among the various conceivable configurations and velocities at any required time, when the distribution has been given for some on time. The fundamental equation for this inquiry is that which gives the rate of change of the number of systems which fall within any infinitisemal limits of configuration and velocity. Such inquires have been called by Maxwell statistical." [[14], pp. VII-VIII]

A continuació, considera la probabilitat de que un sistema arbitrari en la col·lectivitat estigui contingut entre uns límits donats i estableix el principi de probabilitat de fase:

"In the general case, the fundamental equation admits an integration, which gives a principle which may be variously expressed, according to the point of view from which it is regarded, as the conservation of probability of phase ... In other words, we combine the principle of conservation of probability of phase, which is exact, with those approximate relations, which it is customary to assume in the theory of errors." [[14], pp. X-XI]

Boltzmann i Gibbs creien que s'ha de considerar un nombre molt gran de masses gasoses idèntiques. Tanmateix, en aquestes masses idèntiques el moviment de les molècules no és el mateix. Investigant nombres suficientment grans d'aquestes masses i les seves propietats més comunes, Boltzmann i Gibbs van arribar a la interpretació i desenvolupament de la mecànica estadística.

Referències

[1] Adrain R.: *The Analyst, or Mathematical Museum*, 1808.

[2] Bernstein S.: *Probability Theory*, 4th ed. Moscow-Leningrad, 1946.

[3] Boltzmann L.: *Lectures on Gas Theory* (S.G. Brush translation), University of California Berkeley, 1964.

[4] Borel E.: *Le Hasard*, *Paris Librairie Félix Alcan*, París, 1914.

[5] Buniakovski V.Y.: *Foundations of the Mathematical Theory of Probabilities*, Sant Petersburg, 1846.

[6] Chebysev P. L.: *Complete Collected Works*, Vol. 2, Moscou-Leningrad, 1947.

[7] Chebysev P. L.: *Complete Collected Works*, Vol. 3, Moscou-Leningrad, 1948.

[8] Chebysev P. L.: *Complete Collected Works*, Vol. 5, Moscou-Leningrad, 1951.

[9] Màrkov A.A., Bobylev D.K., Krylov A.N., Steklov V.A.: *Istoriko-Matematicheskie Issledovaniya*, Vol. 1, 1948.

[10] Gauss C.F.: *Theoria motus corporum coelestium in sectionibus conicis Solem ambientium*, 1809.

[11] Gauss C.F.: *Bestimmung der Genauigkeit der Beobachtungen* (*Determination of the Accuracy of the Observations*), 1816.

[12] Gauss C.F.: *Theoria combinationis observationum erroribus minimis obnoxiae* (*Theory of the Combination of Observations Least Subject to Errors*), 1821.

[13] Gauss C.F.: *Selected Geodesic Works*, Vol. I, Moscou, 1957.

[14] Gibbs J.W.: *Basic Principles of Statistical Mechanics*, Moscou-Leningrad, 1946.

[15] Gorroochurn P.: *Classic Topics on the History of Modern Mathematical Statistics: From Laplace to More Recent Times*, Wiley Series, 2016.

[16] Hald A.: *A History of Mathematical Statistics, from 1750 to 1930*, Wiley Series, 1998.

[17] Kolmogórov A.: *Foundations of the Theory of Probability*, Chelsea publishing company, New York, 1956.

[18] Kolmogórov A., Fomin V.S.: *Mathematics: its content, methods and significance"*, vol. 2, Moscou, 1956.

[19] Laplace P. S.: *Théorie Analytique des Probabilités*, 1812.

[20] Laplace P. S.: *Essai philosophique sur les probabilités*, 1814.

[21] Liapunov A. M.: *Selected Works*, URSS Academy of Sciences, Moscou, 1948.

[22] Lobatxevski N. I.: *New elements of geometry with a complete story of parallels* (1835-38), *Lobachevsky-Complete Collected Works*, Vol. II Moscow-Leningrad, 1949.

[23] Maistrov L.E.: *Probability Theory. A Historical Sketch*, Academic Press, 1974.

[24] Màrkov A.A.: *On the roots of the equation* $e^{x^2}\partial^m e^{-x^2}/\partial x^m = 0$, URSS Academy of Sciences, Series 5, Chap. 9, 435-446, 1898.

[25] Màrkov A.A.: *Selected Works*, URSS Academy of Sciences, Moscou, 1951.

[26] Màrkov A.A.: *The Calculus of Probabilities*, 4a Edició, Moscou, 1924.

[27] Nekrasov P.: *Probability Theory*, rev. ed. Sant Petersburg, 1912.

[28] Nekrasov P.: *The Philosophy and Logic of the Science of Mass Manifestations of Human Activity*, Moscou, 1902.

[29] Ostrogradski M. V.: *Complete Collection of Works*, Vol. III, URSS Academy of Sciences, Kiev, 1961.

[30] Poincaré H.: *Calcul des Probabilités*, Paris, 1912.

[31] Poisson S.D.: *Researches into the Probabilities of Judgements in Criminal and Civil Cases*, Paris, 1837 (Translated by Oscar Sheynin, Berlin, 2013).

[32] Sheynin O.: *Theory of Probability. A Historical Essay*, Scholars' Press, 2017.

[33] Stigler S.M.: *The measurement of Uncertainty before 1900*, Harvard University Press, 1986.

[34] Zernov N.E.: *Probability Theory*, Moscou, 1843.