INDEPENDÈNCIA ESTADÍSTICA EN PROBABILITAT, ANÀLISI I TEORIA DE NOMBRES

Mark Kac

Professor de Matemàtiques a la Universitat de Cornell

The Carus Mathematical Monographs, 12

publicat per

THE MATHEMATICAL ASSOCIATION OF AMERICA

1959

Traducció al català de Pelegrí Viader i Canals

Universitat Pompeu Fabra

Barcelona

Any de la traducció: 2003

AL MEU MESTRE
EL PROFESSOR HUGO STEINHAUS

PREFACI

A l'assemblea de la «Mathematical Asssociation of America» celebrada l'estiu de 1955, vaig tenir el privilegi d'impartir les «Hedrick Lectures». Em vaig alegrar en extrem quan, un temps més tard, el professor T. Rado, en nom del Comitè de la col·lecció «The Carus Mathematical Monographs», molt amablement em va invitar a ampliar les meves conferències fins a convertir-les en una monografia.

Més o menys al mateix temps, vaig ser honrat amb una invitació del Haverford College per a dictar una sèrie de conferències sota el «Philips Visitors Program». Aquesta invitació em va donar l'oportunitat d'assajar el projecte de monografia amb un públic «en directe». Aquest llibre és una versió lleugerament revisada de les conferències impartides en el Haverford College durant el tercer trimestre de l'any 1958.

L'objectiu principal de les Hedrick Lectures, i també l'objectiu d'aquesta versió augmentada, era mostrar que: a) observacions extremadament simples són, sovint, el punt de partida de teories molt riques i fructíferes, i b) molts desenvolupaments, aparentment no relacionats entre si, són, en realitat, variacions del mateix tema.

Aquest llibre tracta de la noció d'independència estadística, excepte en el darrer capítol on exposo una espectacular aplicació del teorema ergòdic a les fraccions continuades.

La noció d'independència es va originar en la teoria de la probabilitat i, durant molt de temps, va ser usada amb un grau elevat d'imprecisió, de tal manera que va aixecar sospites sobre la seva seriositat com a teoria matemàtica.

Ara sabem com definir la independència estadística en termes molt generals i abstractes. Però la moda moderna cap a la generalització i l'abstracció tendeix, no només a soterrar la simplicitat de la idea subjacent, sinó també a enfosquir la possibilitat d'aplicar les idees probabilistes fora de la teoria de la probabilitat.

En les pàgines que segueixen, he intentat rescatar la independència estadística del destí de l'oblit abstracte mostrant com, en la seva forma més simple, apareix en diversos contextos de manera transversal en diferents àrees de les matemàtiques.

Quant als prerequisits que el lector necessita, suposo que està familiaritzat amb la teoria de la mesura i la integració de Lebesgue, la teoria elemental de les integrals de Fourier i amb els rudiments de la teoria de nombres. Com que no vull suposar conegudes massa més coses, però tampoc vull fer la narració massa feixuga amb molts detalls tècnics, he omès les demostracions d'alguns resultats.

M'excuso per aquestes omissions i confio que el lector estarà suficientment interessat en el tema d'aquest llibre per a omplir els forats ell mateix. He afegit una petita bibliografia que no té cap pretensió de ser completa.

En el llibre també he inclòs uns quants problemes. Aquests problemes són, en general, difícils i el lector no s'hauria de desanimar si no els resol sense un esforç considerable.

Desitjo agrair al professor C. O. Oakley i a R. J. Wisner del Haverford College la seva esplèndida col·laboració i haver convertit la tasca de viatjar d'Ithaca a Haverford en un veritable plaer.

Vaig ser prou afortunat com per comptar entre el públic el professor H. Rademacher de la Universitat de Pennsylvania i el professor John Oxtoby del Bryn Mawr College. Les seves crítiques, suggeriments i el seu suport constant han estat, en veritat, d'una gran vàlua, i tinc amb ells un gran deute.

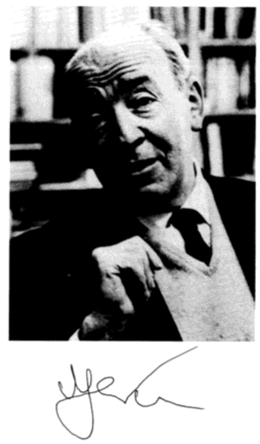
Els professors H. Widom i M. Schreiber, col·legues meus de Cornell, han llegit el manuscrit i són responsables d'un bon grapat de canvis i millores. És un plaer agrair-los la seva ajuda.

El meu agraïment, també, als estudiants de llicenciatura de Haverford i de Bryn Mawr que van fer de «conillets d'índies»; especialment a J. Reill que va compilar la bibliografia i va llegir les galerades del manuscrit.

I per acabar, però no per això amb menys grau, vull agrair a la senyora Axelsson de Haverford College i a la senyora Martin del Departament de Matemàtiques de Cornell la tasca, sovint impossible, de mecanografiar el manuscrit a partir dels meus apunts i notes gairebé il·legibles.

Mark Kac

Ithaca, New York
Setembre de 1959



NOTA BIOGRÀFICA

Mark Kac (1914-1984) va néixer a Polònia a l'inici de la Primera Guerra Mundial. El seu pare era un acadèmic que mentre feia classes particulars per a mantenir la família, ensenyava geometria al jove Kac. En paraules del mateix Kac:

> [. . .] el meu pare es desesperava perquè [. . .] jo era molt dolent amb les taules de multiplicar. Que algú sabés demostrar teoremes de geometria elemental sense saber quant eren set per nou li semblava d'allò més estrany.

Kac va estudiar matemàtiques a la Universitat Jan Kasimir de Lvov (o Lwow) on va tenir Hugo Steinhaus de professor. L'any 1938 va aconseguir emigrar a Amèrica i va ser admès com a ajudant a la Johns Hopkins abans d'anar a Cornell on va estar vint-i-dos anys. Després es va traslladar a la Univsersitat de Rockefeller (NY) on s'hi va passar vint anys més abans d'anar a jubilar-se a la Universitat de Southern California.

Kac va ser un gran divulgador de les matemàtiques. El seu estil és clar i directe al mateix temps que profund. Són famosos articles seus com ≪Random Walk and the Theory of Brownian motion≫, (Am. Math. Month., 1947, 54:369–91), o bé ≪Can One Hear the Shape of a Drum?≫, (Am. Math. Month., 1968, 73:1–23).

Kac va ser dels primers a treballar les aplicacions de la probabilitat a la física estadística, juntament amb físics de la categoria de Richard Feynman.

D'allò, però, que estava més orgullós, en les seves pròpies paraules:

> [...] retrospectivament, d'allò que estic més satisfet és del que vaig fer en col·laboració amb Erdös [...] va ser la introducció dels mètodes probabilistes en la teoria de nombres. Per a dir-ho de manera poètica, els nombres primers juguen un joc d'atzar [vegeu el capítol 4].

Precisament el tema del llibret que teniu a les mans, que és una de les joies de la literatura matemàtica moderna. Recull moltes de les idees de Kac, Erdös, Khintchine, Davenport i altres sobre la probabilitat aplicada a l'anàlisi i la teoria de nombres. L'estil veureu que és allunyat del formalisme i proper a les idees que hi ha sota els grans teoremes i resultats. És una mostra tant de l'art divulgatiu com de la bellesa de les matemàtiques que interconnecten àrees aparentment allunyades.

Els capítols 1, 2 i 5 són força assequibles. La divulgació preval sobre les matemàtiques. Els capítols 3 i 4, el nucli del llibre, són més complicats de seguir per als no especialistes, però encara que no s'entenguin tots els detalls veureu que dóna gust d'anar veient com Kac va mostrant de manera admirable les idees mestres que s'amaguen sota els resultats que ens ofereix.

Pelegrí Viader
Universitat Pompeu Fabra

ÍNDEX

Capítol 1

Des de Vieta fins a la noció d'independència estadística

1.1. Una fórmula de Vieta

Comencem amb una mica de trigonometria molt senzilla. Escrivim

$$\begin{aligned}\sin x &= 2\sin\frac{x}{2}\,\cos\frac{x}{2} \\ &= 2^2\sin\frac{x}{4}\,\cos\frac{x}{4}\,\cos\frac{x}{2} \\ &= 2^3\sin\frac{x}{8}\,\cos\frac{x}{8}\,\cos\frac{x}{4}\,\cos\frac{x}{2} \\ &\vdots \\ &= 2^n\sin\frac{x}{2^n}\prod_{k=1}^{n}\cos\frac{x}{2^k}.\end{aligned} \tag{1.1}$$

Del càlcul elemental sabem que, per $x \neq 0$,

$$1 = \lim_{n\to\infty}\frac{\sin\frac{x}{2^n}}{\frac{x}{2^n}} = \frac{1}{x}\lim_{n\to\infty}2^n\sin\frac{x}{2^n},$$

i, per tant,

$$\lim_{n\to\infty}2^n\sin\frac{x}{2^n} = x. \tag{1.2}$$

Combinant (1.2) amb (1.1) obtenim

$$\frac{\sin x}{x} = \prod_{k=1}^{\infty}\cos\frac{x}{2^k}. \tag{1.3}$$

Un cas concret de (1.3) és especialment interessant. Posant $x = \pi/2$ obtenim

$$\begin{aligned} \frac{2}{\pi} &= \prod_{n=1}^{\infty} \cos \frac{\pi}{2^{n+1}} \qquad (1.4) \\ &= \frac{\sqrt{2}}{2} \frac{\sqrt{2+\sqrt{2}}}{2} \frac{\sqrt{2+\sqrt{2+\sqrt{2}}}}{2} \cdots, \end{aligned}$$

una fórmula clàssica deguda a Vieta.

1.2. Una altra mirada a la fórmula de Vieta

Fins aquí tot ha estat força directe i conegut. Mirem-nos ara (1.3) des d'un punt de vista diferent.

És sabut que tot nombre real t, $0 \le t \le 1$, es pot escriure de manera *única* de la forma

$$t = \frac{\varepsilon_1}{2} + \frac{\varepsilon_2}{2^2} + \cdots \qquad (1.5)$$

on cada ε és 0 o 1.

Això no és altra cosa que el conegut *desenvolupament binari* de t. Per tal d'assegurar-ne la unicitat, convindrem a escriure els desenvolupaments finits de manera que, a partir d'un cert punt, tots els dígits siguin 0. Així, per exemple, escriurem

$$\frac{3}{4} = \frac{1}{2} + \frac{1}{2^2} + \frac{0}{2^3} + \frac{0}{2^4} + \cdots$$

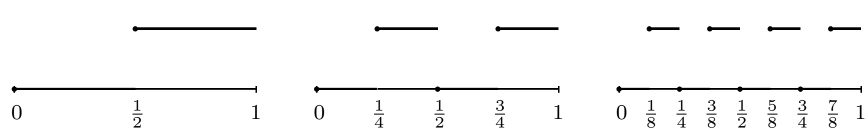
en comptes de

$$\frac{3}{4} = \frac{1}{2} + \frac{0}{2^2} + \frac{1}{2^3} + \frac{1}{2^4} + \cdots .$$

Els dígits ε_i són, evidentment, funcions de t. És, doncs, més adient escriure (1.5) de la forma

$$t = \frac{\varepsilon_1(t)}{2} + \frac{\varepsilon_2(t)}{2^2} + \frac{\varepsilon_3(t)}{2^3} + \cdots . \qquad (1.6)$$

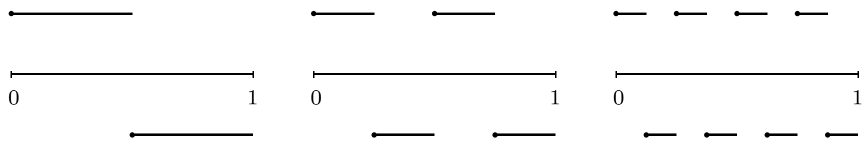
Amb el conveni que hem adoptat sobre els desenvolupaments finits, les gràfiques de $\varepsilon_1(t), \varepsilon_2(t), \varepsilon_3(t), \ldots$ són les següents:



És més convenient introduir les funcions $r_i(t)$ definides per les equacions

$$r_k(t) = 1 - 2\varepsilon_k(t), \quad k = 1, 2, 3, \cdots, \qquad (1.7)$$

que tenen les gràfiques següents:



Aquestes funcions, introduïdes i estudiades per primera vegada per H. Rademacher, es coneixen com a funcions de Rademacher. En termes de les funcions $r_k(t)$, podem reescriure (1.6) de la manera següent:

$$1 - 2t = \sum_{k=1}^{\infty} \frac{r_k(t)}{2^k}\,. \tag{1.8}$$

Observem ara que

$$\int_0^1 e^{ix(1-2t)}\,dt = \frac{\sin x}{x}$$

i que

$$\int_0^1 \exp\left(ix\,\frac{r_k(t)}{2^k}\right)\,dt = \cos\frac{x}{2^k}\,.$$

La fórmula (1.3) adopta la forma

$$\begin{aligned}\frac{\sin x}{x} &= \int_0^1 e^{ix(1-2t)}\,dt = \int_0^1 \exp\left(ix\sum_{k=1}^{\infty}\frac{r_k(t)}{2^k}\right)\,dt \\ &= \prod_{k=1}^{\infty}\cos\frac{x}{2^k} = \prod_{k=1}^{\infty}\int_0^1 \exp\left(ix\,\frac{r_k(t)}{2^k}\right)\,dt\end{aligned}$$

i, en particular, tenim

$$\int_0^1 \prod_{k=1}^{\infty} \exp\left(ix\,\frac{r_k(t)}{2^k}\right)\,dt = \prod_{k=1}^{\infty}\int_0^1 \exp\left(ix\,\frac{r_k(t)}{2^k}\right)\,dt\,. \tag{1.9}$$

La integral d'un producte és igual al producte d'integrals!

1.3. Un accident o el principi d'alguna cosa més profunda?

Podem considerar (1.9) com un accident i no donar-li més importància? Certament no, almenys fins que no hàgim investigat l'assumpte amb més deteniment.

Fem una ullada a la funció

$$\sum_{k=1}^{n} c_k r_k(t)\,.$$

És una funció esglaonada, constant en cada un dels intervals

$$\left(\frac{s}{2^n}, \frac{s+1}{2^n}\right), \quad s = 0, 1, 2, \cdots, 2^n - 1.$$

Els valors que pren són de la forma

$$\pm c_1 \pm c_2 \pm \cdots \pm c_n \, .$$

A cada seqüència (de longitud n) de $+1$ i -1 li correspon un, i només un, interval $(s/2^n, (s+1)/2^n)$. Així

$$\int_0^1 \exp\left[i \sum_1^n c_k r_k(t)\right] dt = \frac{1}{2^n} \sum \exp\left(i \sum_1^n \pm c_k\right),$$

on el sumatori exterior es fa sobre *totes* les possibles seqüències (de longitud n) de $+1$ i -1.

Ara

$$\frac{1}{2^n} \sum \exp\left(i \sum_1^n \pm c_k\right) = \prod_{k=1}^n \left(\frac{e^{ic_k} + e^{-ic_k}}{2}\right) = \prod_{k=1}^n \cos c_k,$$

i, consegüentment,

$$\int_0^1 \exp\left[i \sum_i^n c_k r_k(t)\right] dt = \prod_{k=1}^n \cos c_k = \prod_{k=1}^n \int_0^1 e^{ic_k r_k(t)}\, dt. \tag{1.10}$$

Posant

$$c_k = \frac{x}{2^k}$$

obtenim

$$\int_0^1 \exp\left(ix \sum_1^n \frac{r_k(t)}{2^k}\right) dt = \prod_{k=1}^n \cos\frac{x}{2^k},$$

i, atès que

$$\lim_{n\to\infty} \sum_1^n \frac{r_k(t)}{2^k} = 1 - 2t$$

uniformement en $(0, 1)$, tenim

$$\begin{aligned} \frac{\sin x}{x} &= \int_0^1 e^{ix(1-2t)}\, dt = \lim_{n\to\infty} \int_0^1 \exp\left(ix \sum_1^n \frac{r_k(t)}{2^k}\right) dt \\ &= \lim_{n\to\infty} \prod_{k=1}^n \cos\frac{x}{2^k} = \prod_{k=1}^\infty \cos\frac{x}{2^k}\, . \end{aligned}$$

Hem obtingut, doncs, una demostració de la fórmula (1.3) diferent. És millor que la que hem donat en el § 1.1?

De ben segur és més complicada, però també és més instructiva perquè, d'alguna manera, connecta la fórmula de Vieta amb els dígits binaris.

Ara bé, quina és la propietat dels dígits binaris que fa que la demostració funcioni?

1.4. $\left(\frac{1}{2}\right)^n = \frac{1}{2}\cdots\frac{1}{2}$ (n vegades)

Considerem el conjunt de les t per les quals

$$r_1(t) = +1, \quad r_2(t) = -1, \quad r_3(t) = -1.$$

Una ullada a les gràfiques de r_1, r_2 i r_3 ens mostrarà que aquest conjunt (amb la possible excepció dels punts extrems) és simplement l'interval $(3/8, 4/8)$.

La longitud (o mesura) d'aquest interval és, òbviament, $1/8$, i

$$\frac{1}{8} = \frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2}\,.$$

Aquesta observació trivial es pot escriure de la manera següent:

$$\begin{aligned} & \mu\left\{r_1(t) = +1, r_2(t) = -1, r_3(t) = -1\right\} \\ = \; & \mu\left\{r_1(t) = +1\right\}\cdot\mu\left\{r_2(t) = -1\right\}\cdot\mu\left\{r_3(t) = -1\right\}, \end{aligned}$$

on μ representa la mesura (longitud) del conjunt definit entre claus.

El lector no ha de tenir cap dificultat per a generalitzar aquest resultat a un nombre arbitrari de r. Obtindrà el resultat següent: si $\delta_1, \cdots, \delta_n$ és una seqüència de $+1$ i -1, llavors

$$\begin{aligned} & \mu\left\{r_1(t) = \delta_1, \cdots, r_n(t) = \delta_n\right\} \\ = \; & \mu\left\{r_1(t) = \delta_1\right\}\cdot\mu\left\{r_2(t) = \delta_2\right\}\cdots\mu\left\{r_n(t) = \delta_n\right\}. \end{aligned} \tag{1.11}$$

Això pot semblar només una manera complicada d'escriure

$$\left(\frac{1}{2}\right)^n = \frac{1}{2}\times\frac{1}{2}\times\cdots\times\frac{1}{2} \text{ (n vegades)},$$

però, en realitat, vol dir molt més. Expressa una propietat profunda de les funcions $r_k(t)$ (i, per tant, dels dígits binaris) i és el punt de partida d'un desenvolupament ric i fructífer. Aquesta és la propietat que és l'essència de la demostració del § 1.3. En efecte, ara, (1.10) es pot demostrar de la manera següent:

$$
\begin{aligned}
&\int_0^1 \exp\left[i\sum_1^n c_k r_k(t)\right] dt \\
&\quad= \sum_{\delta_1,\cdots,\delta_n} \exp\left(i\sum_1^n c_k\delta_k\right)\mu\left\{r_1(t)=\delta_1,\cdots,r_n(t)=\delta_n\right\} \\
&\quad= \sum_{\delta_1,\cdots,\delta_n} \prod_1^n e^{ic_k\delta_k}\prod_1^n \mu\left\{r_k(t)=\delta_k\right\} \\
&\quad= \sum_{\delta_1,\cdots,\delta_n} \prod_{k=1}^n e^{ic_k\delta_k}\mu\left\{r_k(t)=\delta_k\right\} \\
&\quad= \prod_{k=1}^n \sum_{\delta_k} e^{ic_k\delta_k}\mu\left\{r_k(t)=\delta_k\right\} \\
&\quad= \prod_{k=1}^n \int_0^1 e^{ic_k r_k(t)}\, dt\,.
\end{aligned}
$$

1.5. Cara o creu?

La teoria elemental del llançament d'una moneda es fonamenta en dos supòsits:

a) la moneda no està trucada;

b) els llançaments successius són independents.

El primer supòsit significa que, en cada llançament individual, les alternatives C (cara) i X (creu) són equiprobables, és a dir, que cada una té assignada «probabilitat» 1/2. El segon, s'utilitza per a justificar la *regla de la multiplicació de les probabilitats*. Aquesta regla (formulada de manera una mica imprecisa) és la següent: si els esdeveniments $A_1, \cdots, A_n$ són *independents*, llavors la probabilitat de l'ocurrència de tots alhora és el producte de les probabilitats de cada ocurrència individual. En altres paraules:

$$
\begin{aligned}
&\text{Prob.}\,\{A_1 \;\text{ i }\; A_2 \;\text{ i }\; A_3 \;\cdots\; \text{ i }\; A_n\} \\
&\quad= \text{Prob.}\,\{A_1\}\cdot\text{Prob.}\,\{A_2\}\cdots\text{Prob.}\,\{A_n\}\,.
\end{aligned} \tag{1.12}
$$

La regla, aplicada als llançaments independents d'una moneda no trucada, ens diu que la probabilitat associada a qualsevol seqüència (de longitud n) de C i X (per exemple, $CCXX\cdots X$) és

$$\frac{1}{2}\times\frac{1}{2}\times\cdots\times\frac{1}{2}=\frac{1}{2^n}\,.$$

Això ens recorda molt el § 1.4 i potser podrem, doncs, usar les funcions $r_k(t)$ com un *model* per al llançament d'una moneda. Per a aconseguir-ho, farem servir el diccionari de termes següent:

Símbol C	$+1$
Símbol X	-1
k–èsim llançament $(k = 1, 2, \cdots)$	$r_k(t)$ $(k = 1, 2, \cdots)$
Esdeveniment	Conjunt de t
Probabilitat d'un esdeveniment	Mesura del corresponent conjunt de t.

Per a veure com s'aplica aquest diccionari, considerem el problema següent: quina és la probabilitat que en n llançaments independents d'una moneda no trucada, exactament l resultats siguin cara? Usant el diccionari traduirem el problema de manera que ara dirà: quina és la mesura del conjunt de t pels quals exactament l dels n nombres $r_1(t), r_2(t), \cdots, r_n(t)$ són iguals a $+1$? Podem resoldre aquest problema (sense recórrer a l'habitual ús de combinacions) mitjançant un recurs que trobarem sovint (amb diferents disfresses) en allò que segueix.

Primer de tot, observem que la condició que exactament l valors entre $r_1(t), \cdots, r_n(t)$ siguin iguals a $+1$ és equivalent a la condició

$$r_1(t) + r_2(t) + \cdots + r_n(t) = 2l - n\,. \tag{1.13}$$

Després, observem que, per m enter, tenim

$$\frac{1}{2\pi}\int_0^{2\pi} e^{imx}\,dx = \begin{cases} 1, & m = 0 \\ 0, & m \neq 0, \end{cases} \tag{1.14}$$

i, en conseqüència,

$$\phi(t) = \frac{1}{2\pi}\int_0^{2\pi} e^{ix[r_1(t)+\cdots+r_n(t)-(2l-n)]}\,dx \tag{1.15}$$

és igual a 1 si (1.13) es satisfà, i és igual a 0 en cas contrari. Així

$$\begin{aligned} \mu\left\{r_1(t) + \cdots + r_n(t) = 2l - n\right\} &= \int_0^1 \phi(t)\,dt \\ &= \int_0^1 \frac{1}{2\pi}\int_0^{2\pi} e^{ix[r_1(t)+\cdots+r_n(t)-(2l-n)]}\,dx\,dt \\ &= \frac{1}{2\pi}\int_0^{2\pi} e^{-i(2l-n)x}\left(\int_0^1 e^{ix[r_1(t)+\cdots+r_n(t)]}\,dt\right)dx. \end{aligned}$$

(L'últim pas involucra l'intercanvi d'ordre d'integració. Això, generalment, es justifica apel·lant a un teorema general de Fubini. En el nostre cas la justificació és trivial atès que $r_1(t) + \cdots + r_n(t)$ és una funció esglaonada.)

Ara, fent servir (1.10) juntament amb $c_1 = c_2 = \cdots = c_n = x$ obtindrem

$$\mu\{r_1(t) + \cdots + r_n(t) = 2l - n\} = \frac{1}{2\pi}\int_0^{2\pi} e^{-i(2l-n)x}\cos^n x\,dx\,. \qquad (1.16)$$

Finalment, deixem com exercici la demostració que

$$\mu\{r_1(t) + \cdots + r_n(t) = 2l - n\} = \frac{1}{2^n}\binom{n}{l}\,. \qquad (1.17)$$

1.6. Independència i «independència»

La noció d'independència, malgrat la seva centralitat en la teoria de la probabilitat, no és purament una noció matemàtica. La regla de multiplicació de les probabilitats d'esdeveniments independents és un intent de formalitzar aquesta noció i de construir un càlcul al seu voltant. Hom es veu naturalment inclinat a considerar esdeveniments que semblen no relacionats com a independents l'un de l'altre. Així, un físic que considera esdeveniments que tenen lloc en dues mostres d'un gas remotament distants l'una de l'altra, les considerarà independents (com podria ser d'una altra manera si una mostra és, diguem a Bismarck, ND, i l'altra és a Washington DC?) i, amb la consciència ben tranquil·la, no tindrà cap problema a invocar la regla de la multiplicació de les probabilitats.

Malauradament, fent això hom pot (de manera innocent i involuntària) crear la impressió que allò de què estem parlant aquí és una *implicació estrictament lògica*.

Allò de què realment estem parlant és d'una *definició* d'independència i d'una creença (nascuda de l'experiència i l'experimentació, per tal d'assegurar-se) que la definició és aplicable a una situació concreta.

Així, existeix la independència en un sentit imprecís i intuïtiu, i existeix una «independència» en un sentit molt més restringit, però ben definit, que diu que la regla de multiplicació de les probabilitats és aplicable.

Van ser les nocions imprecises i intuïtives les que varen subministrar, durant molt de temps, la motivavió principal i la força motriu de la teoria de la probabilitat.

Però mentre s'anava creant un formalisme força impressionant, els matemàtics (amb molt poques excepcions) romanien aliens perquè no tenien clar quins eren els objectes als quals tot aquell formalisme s'aplicava.*

Llavors, el 1909, E. Borel va fer l'observació que els dígits binaris $\varepsilon_k(t)$, [o, equivalentment, les funcions de Rademacher, $r_k(t)$] eren «independents» [vegeu (1.11)]. Finlament, hom tenia objectes ben definits als quals la teoria de la

*Imagineu-vos que un llibre sobre equacions diferencials escrit només en termes de masses, forces, acceleracions i coses semblants caigués en mans d'algú que no ha sentit a parlar mai de mecànica. Aquest hipotètic lector podria ben bé perdre's completament el seu ric contingut purament matemàtic.

probabilitat dels esdeveniments independents es podia aplicar, sense por de quedar embolicat amb monedes, esdeveniments, llançaments i experiments.

La publicació de la, ara ja clàssica, memòria de Borel «Sur les probabilités dénombrables et leurs applications arithmétiques» marca el principi de la teoria de la probabilitat moderna. En el proper capítol discutirem algunes de les línies sobre les quals aquesta teoria s'ha anat desenvolupant.

PROBLEMES

1. Escriviu el desenvolupament ternari de t, $0 \leq t \leq 1$ de la manera següent:
$$t = \frac{\eta_1(t)}{3} + \frac{\eta_2(t)}{3^2} + \frac{\eta_3(t)}{3^3} + \cdots$$
(cada η_k pot prendre els valors 0, 1 i 2). Demostreu que les η són independents.

2. Demostreu que
$$\frac{\sin x}{x} = \prod_{k=1}^{\infty} \frac{1 + 2\cos \dfrac{2x}{3^k}}{3},$$
i generalitzeu el resultat.

3. Demostreu que si $k_1 < k_2 < \cdots < k_s$, llavors
$$\int_0^1 r_{k_1}(t) r_{k_2}(t) \cdots r_{k_s}(t)\, dt = 0.$$

4. Escrivim $2n$ (un enter positiu parell) en el sistema binari:
$$2n = 2^{n_1} + 2^{n_2} + \cdots + 2^{n_k}, \quad 1 \leq n_1 < n_2 < \cdots < n_k,$$
i definim les funcions $\omega_n(t)$ (funcions de Walsh-Kaczmarz) de la manera següent:
$$\begin{aligned} \omega_0(t) &= 1 \\ \omega_n(t) &= r_{n_1}(t) \cdots r_{n_k}(t), \quad n \geq 1. \end{aligned}$$
Demostreu que
   (a) $\int_0^1 \omega_m(t)\omega_n(t)\, dt = \delta_{m,n}$.
   (b) Si $f(t)$ és integrable i
$$\int_0^1 f(t)\omega_n(t)\, dt = 0, \qquad n = 0, 1, 2, \cdots$$
   llavors, $f(t) = 0$ gairebé pertot.

(c) $\displaystyle\int_0^1 \int_0^1 \left| \sum_{k=0}^{2^n} \omega_k(t)\omega_k(s) \right| dt\, ds = 1.$

5. Fent servir la fórmula

$$|\, z \,| = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{1 - \cos zx}{x^2}\, dx$$

demostreu, en primer lloc, que

$$\int_0^1 \left| \sum_1^n r_k(t) \right| dt = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{1 - \cos^n x}{x^2}\, dx > \frac{1}{\pi} \int_{-1/\sqrt{n}}^{1/\sqrt{n}} \frac{1 - \cos^n x}{x^2}\, dx$$

i, finalment, demostreu que

$$\int_0^1 \left| \sum_1^n r_k(t) \right| dt > A\sqrt{n}$$

amb

$$A = \frac{1}{\pi} \int_{-1}^{1} \frac{1 - e^{-y^2/2}}{y^2}\, dy.$$

Nota: la desigualtat de Schwarz combinada amb el resultat del problema 3 per $s = 2$ ens proporciona

$$\int_0^1 \left| \sum_1^n r_k(t) \right| dt \le \sqrt{n}.$$

Capítol 2

Borel i més enllà

2.1. «Lleis dels grans nombres»

Tothom ha sentit dir que és molt difícil fer-se ric només jugant un joc d'atzar sense fer trampes. El que es diu en aquestes ocasions, i d'altres similars, és que «la llei de les mitjanes s'encarrega que això no passi». Què és aquesta «llei de les mitjanes»? És alguna mena de llei física, o és purament una afirmació matemàtica? En gran mesura, és això últim, encara que s'ha de dir que la concordança amb l'evidència empírica és notablement bona. Oblidem-nos de l'evidència empírica i concentrem-nos en les qüestions matemàtiques. Suposem que llenço una moneda no trucada, amb un guany d'1$ cada vegada que surt C i amb una pèrdua d'1$ cada vegada que surt X. Què puc dir sobre la meva fortuna després de n llançaments? Fent servir el diccionari del § 1.4, podem representar aquesta fortuna mitjançant

$$r_1(t) + r_2(t) + \cdots + r_n(t). \tag{2.1}$$

La qüestió que realment interessa el jugador és quina és la probabilitat que, després de n llançaments, la seva fortuna sobrepassi un valor determinat, A_n. Novament amb el nostre diccionari, això és quivalent a preguntar-se per la mesura del conjunt de les t per les quals

$$r_1(t) + r_2(t) + \cdots + r_n(t) > A_n. \tag{2.2}$$

Si és veritat que és difícil fer-se ric jugant a aquest joc, llavors, per un valor d'A_n «suficientment gran» la mesura del conjunt definit per (2.2) hauria de ser «petita». (De manera similar, també hauria de ser poc probable que les pèrdues fossin superiors a A_n.) Podem fer tot això més precís demostrant el teorema següent.

Per tot $\epsilon > 0$,

$$\lim_{n\to\infty} \mu\{|\, r_1(t) + \cdots + r_n(t)\,| > \epsilon n\} = 0. \tag{2.3}$$

Un atac evident es pot basar en la fórmula (1.17) del capítol 1. De fet, tenim

$$\begin{aligned}
&\mu\{\,|\,r_1(t)+\cdots+r_n(t)\,|>\epsilon n\,\}\\
&\quad=\sum_{|\,2l-n\,|>\epsilon n}\mu\{\,r_1(t)+\cdots+r_n(t)=2l-n\,\}\\
&\quad=\sum_{|\,2l-n\,|>\epsilon n}\frac{1}{2^n}\binom{n}{l},
\end{aligned}$$

i només hem de demostrar que, per tot $\epsilon>0$,

$$\lim_{n\to\infty}\sum_{|\,2l-n\,|>\epsilon n}\frac{1}{2^n}\binom{n}{l}=0. \tag{2.4}$$

Intenteu-ho! No és difícil però tampoc és massa fàcil, sobretot si seguiu la inclinació natural de fer servir la fórmula de Stirling. Si us en sortiu, haureu redescobert, essencialment, la demostració orginal de Bernoulli. Hi ha, però, un camí millor i més fàcil degut a Tchebyscheff.

Simplement escriviu

$$\begin{aligned}
&\int_0^1 (r_1(t)+\cdots+r_n(t))^2\,dt &(2.5)\\
&\quad\geq\int_{|\,r_1(t)+\cdots+r_n(t)\,|>\epsilon n}(r_1(t)+\cdots+r_n(t))^2\,dt\\
&\quad>\epsilon^2n^2\mu\{\,|\,r_1(t)+\cdots+r_n(t)\,|>\epsilon n\,\}. &(2.6)
\end{aligned}$$

Si heu treballat el problema 3 al final del capítol 1, tindreu

$$\int_0^1 (r_1(t)+\cdots+r_n(t))^2\,dt=n \tag{2.7}$$

i, per tant, usant (2.5),

$$\mu\{\,|\,r_1(t)+\cdots+r_n(t)>\epsilon n\,|\,\}<\frac{1}{\epsilon^2n}. \tag{2.8}$$

Queda, doncs, demostrat (2.3) amb «espai de sobres».

Recordeu aquest bonic esquema de Tchebysheff; el trobarem novament!

L'afirmació (2.3) dóna cos a l'exemple més senzill d'allò que es coneix com *la llei feble dels grans nombres*. L'adjectiu feble no té un sentit pejoratiu; només s'utilitza per a distingir aquesta llei d'una altra llei dels grans nombres, a la qual ens referim habitualment com la *llei forta*. Forta no té cap sentit de lloa; simplement es fa servir perquè en el joc de cara o creu implica la llei feble i, per tant, és més forta en el sentit lògic de la paraula.

Ambdues lleis han estat a bastament generalitzades i, en les seves versions més recents, cap d'aquestes implica l'altra. Aquestes són, però, qüestions tècniques que aquí no vénen a tomb. El contingut matemàtic de la llei feble dels grans nombres és relativament magre. En la forma (2.4) és un teorema divertit sobre nombres combinatoris. Podria ser aquesta, doncs, una formulació d'aquella misteriosa «llei de les mitjanes» a la qual ens referiem més amunt? Em temo que sí. Això és, essencialment, tot el que podem esperar d'una teoria purament matemàtica.

2.2. Borel i els «nombres normals»

Borel va descobrir una altra llei dels grans nombres. Va demostrar que, per a gairebé tot t (és a dir, per a qualsevol t excepte els d'un conjunt de mesura 0) tenim

$$\lim_{n\to\infty} \frac{r_1(t) + r_2(t) + \cdots + r_n(t)}{n} = 0. \tag{2.9}$$

La demostració és senzilla. Es basa en un teorema ben conegut de la teoria de la mesura i la integració de Lebesgue. El teorema en qüestió és el següent:

Si $\{f_n(t)\}$ és una successió de funcions *no negatives* integrables Lebesgue, llavors la convergència de

$$\sum_{n=1}^{\infty} \int_0^1 f_n(t)\,dt \tag{2.10}$$

implica la convergència *gairebé pertot* de la sèrie

$$\sum_{n=1}^{\infty} f_n(t)\,dt. \tag{2.11}$$

Posem

$$f_n(t) = \left(\frac{r_1(t) + \cdots + r_n(t)}{n}\right)^4 \tag{2.12}$$

i considerem

$$\int_0^1 \left(\frac{r_1(t) + \cdots + r_n(t)}{n}\right)^4 dt.$$

Fent servir el resultat del problema 3 del final del capítol 1, trobem fàcilment que

$$\int_0^1 \left(\frac{r_1(t) + \cdots + r_n(t)}{n}\right)^4 dt = \frac{n + \dfrac{4!}{2!2!}\dbinom{n}{2}}{n^4},$$

i, per tant,

$$\sum_{n=1}^{\infty} \int_0^1 f_n(t)\,dt < \infty.$$

Es dedueix que

$$\sum_{n=1}^{\infty}\left(\frac{r_1(t)+\cdots+r_n(t)}{n}\right)^4$$

convergeix gairebé pertot i, *a fortiori*,

$$\lim_{n\to\infty}\left(\frac{r_1(t)+\cdots+r_n(t)}{n}\right)^4=0$$

gairebé pertot. Això demostra (2.9).

Si recordem que

$$r_k(t)=1-2\,\varepsilon_k(t),$$

llavors (2.9) és equivalent a dir que, per a gairebé tot t,

$$\lim_{n\to\infty}\frac{\varepsilon_1(t)+\cdots+\varepsilon(t)}{n}=\frac{1}{2}\,. \tag{2.13}$$

En altres paraules, gairebé tot nombre t té (asimptòticament!) el mateix nombre de zeros i uns en el seu desenvolupament en binari! Aquest és el contingut aritmètic del teorema de Borel. Probabilísticament parlant, què ens diu el teorema? Usant el nostre diccionari, arribem al següent enunciat: si llencem indefinidament una moneda no trucada i els llançaments són independents llavors, amb probabilitat 1, la *freqüència* amb la qual les cares (creus) apareixen és 1/2 (òbviament, en el límit). Aquest enunciat satisfà la nostra intuïció d'allò que la llei de les mitjanes hauria de dir i ens reafirma en la validesa del nostre diccionari.

El lector s'haurà adonat, sens dubte, que no hi ha res sagrat al voltant de la base 2. Si g és un enter més gran que 1, podem escriure

$$t=\frac{\omega_1(t)}{g}+\frac{\omega_2(t)}{g^2}+\cdots \qquad 0\le t\le 1, \tag{2.14}$$

on cada dígit $\omega(t)$ pot ara prendre els valors $0,1,\cdots,g-1$. Deixem al lector la demostració que, per a gairebé tot t $(0\le t\le 1)$

$$\lim_{n\to\infty}\frac{F_n^{(k)}(t)}{n}=\frac{1}{g}\,, \tag{2.15}$$

on $F_n^{(k)}(t)$ denota el nombre de vegades que el dígit k, $0\le k\le g-1$, apareix entre els primers n valors de les ω. (Vegeu el problema 1 de la pàgina 15.) A partir del fet que la unió numerable de conjunts de mesura 0 té mesura 0, es dedueix que gairebé tot nombre t, $0\le t\le 1$, és tal que en qualsevol sistema de numeració (és a dir, per a *qualsevol* $g>1$) cada dígit permissible apareix amb la freqüència adequada (justament!). En altres paraules, gairebé tot nombre és «normal»!

Com acostuma a passar, és més fàcil demostrar que una gran majoria d'objectes posseix una determinada propietat que *exhibir-ne* encara que sigui només un. El cas que ara ens ocupa no és cap excepció. És força difícil exhibir un nombre normal! L'exemple més senzill és el nombre (escrit en notació decimal)

$$0.1234567891011121314151617181920 21\cdots,$$

on, després del punt decimal escrivim tots els enters positius en successió. La demostració de la normalitat d'aquest nombre no és, en absolut, trivial.

PROBLEMES

1. Demostreu (2.15) veient primer que les ω són independents i després generalitzant el resultat del problema 3 del capítol 1.

2. Sigui $f(t)$, $0 \leq t \leq 1$ una funció contínua. Demostreu que

$$\lim_{n\to\infty} \int_0^1 \cdots \int_0^1 f\left(\frac{x_1 + \cdots + x_n}{n}\right) dx_1\, dx_2 \cdots dx_n = f(1/2).$$

Indicació: demostreu, en primer lloc, imitant la demostració de Tchebysheff de (2.4), que el volum n-dimensional del conjunt definit per les desigualtats

$$\left|\frac{x_1 + \cdots + x_n}{n} - \frac{1}{2}\right| > \epsilon, \quad 0 \leq x_i \leq 1, \ \ i = 1, 2, \cdots, n$$

és menor que $1/12\,\epsilon^2 n$.

3. *La moneda trucada.* Definim $T_p(t)$, $0 < p < 1$, de la manera següent:

$$T_p(t) = \begin{cases} \dfrac{t}{p}, & 0 \leq t \leq p \\[2ex] \dfrac{t-p}{1-p}, & p < t \leq 1, \end{cases}$$

i també que

$$\varepsilon_p(t) = \begin{cases} 1, & 0 \leq t \leq p \\ 0, & p < t \leq 1. \end{cases}$$

Dibuixeu les funcions

$$\varepsilon_1^{(p)}(t) = \varepsilon_p(t), \quad \varepsilon_2^{(p)}(t) = \varepsilon_p(T_p(t)), \quad \varepsilon_3^{(p)}(t) = \varepsilon_p(T_p(T_p(t))), \cdots$$

i demostreu que són independents. Observeu que si $p = 1/2$, hom obté 1 menys els dígits binaris.

4. Demostreu que la mesura del conjunt en el qual

$$\varepsilon_1^{(p)}(t) + \cdots + \varepsilon_n^{(p)}(t) = l, \quad 0 \leq l \leq n$$

és igual a

$$\binom{n}{l} p^l (1-p)^{n-l}.$$

5. Expliqueu com es poden fer servir les funcions $\varepsilon_n^{(p)}(t)$ per a construir un model de llançaments independents d'una moneda trucada, on la probabilitat de C sigui p i la probabilitat de X sigui $q = 1 - p$.

6. Demostreu que si $f(t)$ és contínua llavors

$$\int_0^1 f\left(\frac{\varepsilon_1^{(p)}(t) + \cdots + \varepsilon_n^{(p)}(t)}{n}\right) dt$$
$$= \sum_{k=0}^{n} f\left(\frac{k}{n}\right) \binom{n}{k} p^k (1-p)^{n-k} = B_n(p).$$

[Els $B_n(p)$ són els coneguts polinomis de Bernstein.]

7. Amb l'ajut del «truquet» de Tchebyscheff estimeu la mesura del conjunt en el qual

$$\left| \frac{\varepsilon_1^{(p)}(t) + \cdots + \varepsilon_n^{(p)}(t)}{n} - p \right| > \epsilon$$

i demostreu que

$$\lim_{n \to \infty} B_n(p) = f(p)$$

uniformement a $0 \leq p \leq 1$ [definim $B_n(0) = f(0)$ i $B_n(1) = f(1)$.] (Aquesta és la demostració original de S. Bernstein del famós teorema de Weierstrass sobre aproximació de funcions contínues per polinomis.)

8. Suposem que $f(t)$ satisfà la condició de Lipschitz d'ordre 1, és a dir,

$$|\, f(t_1) - f(t_2)\,| < M\,|\, t_1 - t_2\,|, \quad 0 \leq t_1, t_2 \leq 1,$$

on M és una constant independent de t_1 i t_2. Demostreu que

$$|\, f(p) - B_n(p)\,| \leq \frac{M}{2} \frac{1}{\sqrt{n}}.$$

9. Sigui

$$f(t) = \left| t - \frac{1}{2} \right|, \qquad 0 \leq t \leq 1.$$

Observeu que $f(t)$ satisfà la condició de Lipschitz d'ordre 1. Feu servir el resultat del problema 7 del capítol 1 per a estimar inferiorment

$$|\, f(1/2) - B_n(1/2)\,|\,,$$

i demostreu, d'aquesta manera, que l'ordre $\sqrt{n}$ en l'estimació del problema 8 de més amunt, és la millor possible.

10. Demostreu que, per a gairebé tot t,

$$\lim_{n\to\infty} \frac{\varepsilon_1^{(p)}(t) + \cdots + \varepsilon_n^{(p)}(t)}{n} = p.$$

11. Demostreu que existeix una funció creixent $\phi_p(x)$ tal que

$$\varepsilon_k^{(p)}(t) = \varepsilon_k(\phi_p(t)), \quad k = 1, 2, \cdots$$

(els ε_k són els dígits binaris). Demostreu, a més a més, que per a $p \neq 1/2$ la funció $\phi_p(t)$ és «singular», és a dir, cada conjunt E de mesura positiva conté un subconjunt E_1 que difereix de E en un conjunt de mesura 0 i tal que la imatge $\phi_p(E_1)$ és de mesura 0. [Vegeu Z. Lomnicki i S. Ulam, *Fund. Math.*, **23** (1934), 237–278, en particular les pàgines 268–269.]

12. Demostreu que per a tot $\epsilon > 0$ la sèrie

$$\sum_{n=1}^{\infty} \frac{1}{n^{2+\epsilon}} \exp\left\{ \frac{\sqrt{2\log n}}{\sqrt{n}}\, |\, r_1(t) + \cdots + r_n(t)\,| \right\}$$

convergeix gairebé pertot i que, en conseqüència,

$$\limsup_{n\to\infty} \frac{|\, r_1(t) + \cdots + r_n(t)\,|}{\sqrt{n\log n}} \leq \sqrt{2}$$

gairebé pertot. *Indicació*: Observeu que (ξ real)

$$\int_0^1 e^{\xi\,|\,r_1(t)+\cdots+r_n(t)\,|}\,dt$$
$$< \int_0^1 e^{\xi(r_1(t)+\cdots+r_n(t))}\,dt + \int_0^1 e^{-\xi(r_1(t)+\cdots+r_n(t))}\,dt = 2(\cosh\xi)^n.$$

Nota. El resultat

$$\limsup_{n\to\infty} \frac{|\, r_1(t) + \cdots + r_n(t)\,|}{\sqrt{n\log n}} \leq \sqrt{2}$$

va ser obtingut per primera vegada per Hardy i Littlewood el 1914 d'una manera força complicada. El resultat, més fort, que diu que

$$\limsup_{n\to\infty} \frac{|\, r_1(t) + \cdots + r_n(t)\,|}{\sqrt{n\log\log n}} \leq \sqrt{2}$$

gairebé pertot, va ser demostrat el 1922 per Khintchine. Aquest és considerablement més difícil d'obtenir.

2.3. «Cara o creu»: una formulació més abstracta

Un procediment universalment acceptat de les teories estadístiques (és a dir, les teories basades en la noció de probabilitat) es pot resumir de la manera següent.

Hom comença amb un conjunt Ω (l'espai mostral) amb mesura (probabilitat) igual a 1 per conveni. A Ω hi ha una col·lecció de subconjunts (*conjunts elementals* o *esdeveniments elementals*) amb mesures (probabilitats) donades a l'avançada. El problema consisteix a «estendre» aquesta mesura a la col·lecció de subconjunts de Ω més àmplia possible.

Les regles per a fer l'extensió són les següents:

1a. Si $A_1, A_2, \cdots$ són subconjunts de Ω (esdeveniments) disjunts (incompatibles) i són mesurables (és a dir, tenen assignada una mesura), llavors la seva unió $\bigcup_{k=1}^{\infty} A_k$ també és mesurable i

$$\mu\left\{\bigcup_{k=1}^{\infty} A_k\right\} = \sum_{k=1}^{\infty} \mu\{A_k\},$$

on $\mu\{\ \}$ és la mesura assignada al conjunt entre claus.

2a. Si A és mesurable, també ho és el seu complementari $\Omega - A$. (De la 1a i la 2a es desprèn que $\mu\{\Omega - A\} = 1 - \mu\{A\}$ i, en particular, atès que Ω és mesurable per hipòtesi, que la mesura del conjunt buit és 0.)

3a. Un subconjunt d'un conjunt de mesura 0 és mesurable.

Les funcions mesurables $f(\omega)$, $\omega \in \Omega$, definides a Ω s'anomenen *variables aleatòries* (una terminologia força espantosa i que indueix a error, però que, malauradament, està arrelada de manera irreversible).

Veiem de quina manera les cares i creus entren en aquest esquema.

L'espai mostral Ω és simplement el conjunt de totes les seqüències infinites de símbols C i X, és a dir, seqüències del tipus

$$\omega \ :\ CXCCXXX\cdots.$$

Quins són els esdeveniments elementals? Tradicionalment són els *conjunts cilíndrics*, és a dir, conjunts de seqüències en els quals un nombre finit de llocs concrets es mantenen fixats. Per exemple, el conjunt de seqüències amb tercer element X, setè C i onzè X, és un conjunt cilíndric. Quina mesura s'assigna a un d'aquests conjunts cilíndrics? Això depèn, òbviament, dels supòsits no matemàtics sobre llançaments d'una moneda i que hem de traduir a llenguatge matemàtic. Els llançaments independents d'una moneda no trucada es tradueixen a aquest llenguatge assignant a cada conjunt cilíndric la mesura

$$\left(\frac{1}{2}\right)^k$$

on k és el nombre de llocs concrets que s'han mantingut fixats. Es planteja ara l'important problema de demostrar la *unicitat de la mesura estesa*. En el nostre cas, això es pot fer molt fàcilment apel·lant a la unicitat de la mesura de Lebesgue. Aquest resultat diu que si es defineix una mesura μ a $(0,1)$ que satisfà les condicions 1a, 2a i 3a i la μ-mesura de cada interval és igual a la seva longitud, llavors μ és la mesura ordinària de Lebesgue. Si escrivim 1 en lloc de C i 0 en lloc de X, llavors a cada seqüència de símbols C i X correspon (de manera única, llevat d'un conjunt numerable de racionals diàdics) un nombre t, $0 \leq t \leq 1$, a saber, el nombre escrit en binari resultant de substituir les C i les X de la seqüència per uns i zeros. Aquesta aplicació també té la propietat que transforma conjunts cilíndrics en unions disjuntes d'intervals amb extrems racionals diàdics i, a més a més, la mesura que tenim assignada al conjunt cilíndric és igual a la mesura de Lebesgue (longitud) del conjunt en el qual s'aplica. Ara ja estem!

La unicitat de l'extensió també es pot demostrar sense apel·lar a l'aplicació. El teorema més general d'aquesta mena va ser demostrat per Kolmogoroff en el seu llibre *Grundbegiffe der Wahrscheinlichkeitsrechnung* de 1933.

Un cop fermament establerta la mesura a Ω, podem, de manera estàndard, construir una teoria d'integració que transcorri paral·lelament a l'habitual teoria de Lebesgue.

Sigui $\omega \in \Omega$, és a dir, sigui ω una seqüència de símbols C i X. Posem

$$X_k(\omega) = \begin{cases} +1, & \text{si el } k\text{–èsim element de } \omega \text{ és } C, \\ -1, & \text{si el } k\text{–èsim element de } \omega \text{ és } X. \end{cases}$$

Les funcions $X_k(\omega)$ són *variables aleatòries independents* en el sentit que

$$\begin{aligned} &\mu\,\{\,X_1(\omega) = \delta_1, X_2(\omega) = \delta_2, \cdots, X_n(\omega) = \delta_n\,\} \\ &\qquad\qquad = \frac{1}{2^n} = \prod_{k=1}^{n} \mu\,\{\,X_k(\omega) = \delta_k\,\} \end{aligned} \tag{2.16}$$

per a tota seqüència de δ_j, on cada δ_j és o bé $+1$ o bé -1. És clar que les $X_k(\omega)$ ens proporcionen un model de llançaments independents d'una moneda no trucada.

2.4. Abstracció, a quin preu?

Abstreure és, suposadament, limitar-se fins a allò més essencial. És deslliurar-se de característiques accidentals i concentrar l'atenció en les característiques crucials. De manera abstracta, la teoria de cares i creus (la moneda no trucada, els llançaments independents) és simplement l'estudi de les funcions $X_k(\omega)$ que tenen la propietat (2.16) definides en algun espai Ω (de mesura 1) en el qual hi ha donada una mesura μ que satisfà les condicions 1a, 2a i 3a de la secció precedent. Allò que Ω sigui realment és irrellevant, i hom només pot fer servir (2.16) i les rudimentàries propietats de la mesura 1a, 2a i 3a. Evidentment, hom s'ha de convèncer que no està treballant en un buit matemàtic, és a dir, que els objectes dels quals estem parlant es poden definir. Això s'aconsegueix prenent Ω com l'espai mostral i construint la mesura requerida, μ, tal com s'ha indicat en el § 2.3. El fet que la *materialització* de les $X_k(\omega)$ vingui donada per les funcions de Rademacher $r_k(t)$, és a dir, que puguem prendre com a Ω l'interval habitual $(0,1)$ amb la mesura ordinària de Lebesgue, es pot considerar com un fet accidental. Observem que, amb l'excepció de la demostració divertida de la fórmula de Vieta, en la qual hem fet servir una propietat molt concreta de les funcions de Rademacher, a saber, que

$$1 - 2t = \sum_{k=1}^{\infty} \frac{r_k(t)}{2^k},$$

no hem apel·lat a res més que a la propietat (2.16) i les propietats *generals* de la mesura. Tanmateix el preu que hom pot haver de pagar per l'abstracció desenfrenada és més gran, de fet, molt més gran. Perquè l'abstracció desenfrenada tendeix, també, a distreure l'atenció d'alguns trets que el punt de vista abstracte ha descartat. En tot aquest llibret es troben exemples de tot això. Comencem veient uns quants exemples del reialme que ja ens és familiar.

2.5. Exemple 1. Convergència de sèries amb signes aleatoris

Quina és la probabilitat que la sèrie

$$\sum_{k=1}^{\infty} \pm c_k \qquad (c_k \text{ real}),$$

convergeixi, sabent que els signes s'han triat cadascun de manera independent amb probabilitat $1/2$? El problema va ser formulat per primera vegada d'aquesta manera per H. Steinhaus l'any 1922 (i independentment per N. Wiener) a qui també devem la part més essencial del § 2.3. Steinhaus va fer notar que el problema és equivalent a trobar la mesura del conjunt dels t pels quals la sèrie

$$\sum_{1}^{\infty} c_k r_k(t) \tag{2.17}$$

convergeix. A l'època, aquesta qüestió ja havia estat contestada per Rademacher que havia demostrat que, si

$$\sum_{1}^{\infty} c_k^2 < \infty, \tag{2.18}$$

la sèrie (2.17) convergeix gairebé pertot. Podríem, evidentment, considerar la convergència de

$$\sum_1^\infty c_k X_k(\omega), \tag{2.19}$$

on les $X_k(\omega)$ tenen la propietat (2.16). De fet, la demostració de Kolmogoroff que ha trobat la generalització més gran possible del teorema de Rademacher només feia servir (2.16). Existeix, però, una demostració molt bonica deguda a Paley i Zygmund que fa un ús essencial de les funcions de Rademacher. És aquesta demostració la que reproduirem aquí per raons que es faran patents més endavant. La demostració es basa en dos teoremes, no del tot elementals però molt importants:

1. El teorema de Riesz-Fischer, que diu que si

   $$\sum a_k^2 < \infty$$

   i si $\phi_1(t), \phi_2(t), \cdots$ són ortonormals en un conjunt E, és a dir,

   $$\int_E \phi_i(t)\phi_j(t)\,dt = \delta_{i,j}, \tag{2.20}$$

   llavors existeix una funció $f(t) \in L^2$ (és a dir, tal que $\int_E f^2(t)\,dt < \infty$) tal que

   $$\lim_{n\to\infty} \int_E \left(f(t) - \sum_{k=1}^n \phi_k(t) \right)^2 dt = 0. \tag{2.21}$$

2. El teorema fonamental del càlcul, que en la seva versió «avançada» diu que si

   $$\int_0^1 |\, f(t)\,|\; dt < \infty, \tag{2.22}$$

   llavors, *per a gairebé tot* t_0,

   $$\lim_{n\to\infty} \frac{1}{\beta_m - \alpha_m} \int_{\alpha_m}^{\beta_m} f(t)\,dt = f(t_0) \tag{2.23}$$

   sempre que

   $$\alpha_m < t_0 < \beta_m \qquad \text{i} \qquad \lim_{m\to\infty} \alpha_m = \lim_{m\to\infty} \beta_m = t_0.$$

Ara, atès que sabem que les funcions de Rademacher són ortonormals a l'interval $(0,1)$

$$\int_0^1 r_i(t) r_j(t)\,dt = \delta_{i,j},$$

existeix una funció $f(t)$ (pel teorema de Riesz-Fischer de més amunt) tal que

$$\int_0^1 f^2(t)\,dt < \infty \tag{2.24}$$

i

$$\lim_{n\to\infty} \int_0^1 \left(f(t) - \sum_{k=1}^{n} c_k r_k(t) \right)^2 dt = 0. \tag{2.25}$$

(Recordem que hem suposat que

$$\sum_1^\infty c_k^2 < \infty.)$$

Sigui ara t_0 tal que (2.23) és vàlid [(2.24) implica (2.22)!], i sigui

$$\alpha_m = \frac{k_m}{2^m} < t_0 < \frac{k_m+1}{2^m} = \beta_m \tag{2.26}$$

(excloem la possibilitat que t_0 sigui un racional diàdic). Tenim

$$\begin{aligned} &\left| \int_{\alpha_m}^{\beta_m} \left(f(t) - \sum_1^n c_k r_k(t) \right) dt \right| \\ \leq\ &(\beta_m - \alpha_m)^{1/2} \left(\int_0^1 \left(f(t) - \sum_1^n c_k r_k(t) \right)^2 dt \right)^{1/2} \end{aligned}$$

i, per tant, per (2.25)

$$\int_{\alpha_m}^{\beta_m} f(t)\,dt = \sum_1^\infty c_k \int_{\alpha_m}^{\beta_m} r_k(t)\,dt. \tag{2.27}$$

Observem ara que

$$\int_{\alpha_m}^{\beta_m} r_k(t)\,dt = 0, \quad k > m \tag{2.28}$$

i

$$\int_{\alpha_m}^{\beta_m} r_k(t)\,dt = (\beta_m - \alpha_m)\, r_k(t_0), \quad k \leq m. \tag{2.29}$$

Així, (2.27) esdevé

$$\frac{1}{\beta_m - \alpha_m} \int_{\alpha_m}^{\beta_m} f(t)\,dt = \sum_1^m c_k r_k(t_0)$$

i, per tant, per (2.23)

$$\sum_1^\infty c_k r_k(t_0)$$

convergeix.

L'argument que acabem de veure s'estén immediatament per a demostrar que la sèrie

$$\sum_{k=1}^{\infty} c_k \sin 2\pi 2^k t \tag{2.30}$$

convergeix gairebé pertot si

$$\sum c_k^2 < \infty. \tag{2.31}$$

Aquest teorema sorgeix de manera natural si hom s'adona que

$$r_k(t) = \operatorname{sgn} \sin 2\pi 2^{k-1} t.$$

De fet, la nostra demostració gira al voltant de tres propietats de les funcions de Rademacher:

1a l'ortonormalitat,

2a (2.28),

3a (2.29).

D'aquestes, la 1a i la 2a es satisfan quan $r_k(t)$ es substitueix per $\sin 2\pi 2^k t$. La propietat 3a no es satisfà estrictament, però per $k \leq m$ tenim

$$\int_{\alpha_m}^{\beta_m} \sin 2\pi 2^k t \, dt = \\ (\beta_m - \alpha_m) \sin 2\pi 2^k t_0 + \int_{\alpha_m}^{\beta_m} (\sin 2\pi 2^k t - \sin 2\pi 2^k t_0) \, dt \tag{2.32}$$

i

$$\begin{aligned} &\left| \int_{\alpha_m}^{\beta_m} (\sin 2\pi 2^k t - \sin 2\pi 2^k t_0) \, dt \right| \\ &\leq \int_{\alpha_m}^{\beta_m} \left| \sin 2\pi 2^k t - \sin 2\pi 2^k t_0) \right| dt \leq 2\pi 2^k \int_{\alpha_m}^{\beta_m} |\, t - t_0 \,| \, dt \\ &< 2\pi 2^k (\beta_m - \alpha_m)^2 = 2\pi \frac{2^k}{2^m} (\beta_m - \alpha_m). \end{aligned}$$

Ara, en lloc de

$$\frac{1}{\beta_m - \alpha_m} \int_{\alpha_m}^{\beta_m} f(t) \, dt = \sum_{1}^{m} c_k r_k(t_0),$$

obtenim

$$\left| \frac{1}{\beta_m - \alpha_m} \int_{\alpha_m}^{\beta_m} f(t) \, dt - \sum_{1}^{m} c_k \sin 2\pi 2^k t_0 \right| \leq \sum_{1}^{m} |\, c_k \,| \frac{2\pi}{2^{m-k}}$$

i, atès que $c_n \to 0$ quan $n \to \infty$ (recordem que $\sum c_k^2 < \infty\,!$), hom té que

$$\lim_{m\to\infty} \sum_1^m |\, c_k \,| \frac{2\pi}{2^{m-k}} = 0,$$

i això és suficient per a acabar la demostració.

El teorema que acabem de demostrar respecte a la convergència de

$$\sum_1^\infty c_k \sin 2\pi 2^k t$$

és, de fet, un cas particular d'un teorema famós de Kolmogoroff que diu que

$$\sum_1^\infty c_k^2 < \infty$$

implica la convergència gairebé pertot de

$$\sum_{k=1}^\infty c_k \sin 2\pi n_k t$$

sempre que existeixi un nombre q tal que

$$\frac{n_{k+1}}{n_k} > q > 1.$$

La demostració de Kolmogoroff feia servir, de manera essencial, el fet que les sèries en qüestió eren sèries trigonomètriques, però mitjançant una extensió de l'argument de Paley-Zygmund, hom pot demostrar el següent teorema molt més general.

Si $g(t)$ és periòdica amb període 1, i si

a) $$\int_0^1 g(t)\, dt = 0$$

b) $$|\, g(t') - g(t'') \,| < M \,|\, t' - t'' \,|^\alpha, \quad 0 < \alpha < 1,$$

llavors la convergència de $\sum c_k^2$ implica la convergència gairebé pertot de

$$\sum_1^\infty c_k g(n_k t)$$

sempre que els enters n_k siguin tals que

$$\frac{n_{k+1}}{n_k} > q > 1.$$

La demostració d'aquesta afirmació és una mica massa tècnica per a ser reproduïda aquí, encara que no hi apareix cap nova idea més enllà de les que calien per a Paley-Zygmund.

Quina lliçó hem de treure de tot això? El fet, aparentment accidental, que

$$r_k(t) = \operatorname{sgn} \sin 2\pi 2^{k-1} t$$

suggereix que existeixen analogies entre $r_k(t)$ i $\sin 2\pi 2^{k-1} t$. Atès que les $r_k(t)$ tenen una interpretació clarament probabilística, s'obre un camí per a connectar «cares i creus» amb un reialme matemàtic sense cap relació amb el món de la sort, la probabilitat, les monedes, etc. Podíem haver-ho aconseguit si haguéssim insistit a tractar les «cares i creus» de manera abstracta? Potser sí, però, francament ho dubto.

2,6. Exemple 2. Divergència d'una sèrie amb signes aleatoris

¿Què passa amb la sèrie

$$\sum_{k=1}^{\infty} \pm c_k \tag{2.33}$$

quan

$$\sum_{1}^{\infty} c_k^2 = \infty? \tag{2.34}$$

La resposta és que (2.33) divergeix amb probabilitat 1. La demostració és força senzilla. Primer de tot, observem que el nostre problema és simplement determinar la mesura del conjunt de convergència de

$$\sum_{1}^{\infty} c_k r_k(t) \tag{2.35}$$

sota la condició (2.34). Després, observem que el conjunt de convergència de (2.35) ha de tenir, o bé mesura 0 o bé mesura 1 (un cas particular de l'anomenada llei zero-u). Recordem que

$$r_k(t) = r_1(2^{k-1}t)\ ^*,$$

i, per tant, si t està en el conjunt de convergència, també ho està

$$t + \frac{1}{2^l},$$

*S'hauria d'entendre que $r_k(t)$ es defineix de manera que sigui periòdica de període 1. En altres paraules: $r_k(t+1) = r_k(t)$.

per a $l = 0, 1, 2, \cdots$. De fet, si t es substitueix per $t + 2^{-l}$ només canvien un nombre *finit* de termes de (2.35), cosa que no pot afectar la convergència. D'aquesta manera, la funció característica del conjunt de convergència té períodes arbitràriament petits i, per un teorema ben conegut, ha de ser constant gairebé pertot —la constant ha de ser 0 o 1.[†]

Podem suposar que $c_n \to 0$, perquè d'altra manera l'enunciat del nostre teorema seria trivial.

Suposem ara que (2.34) és veritat, $c_n \to 0$ i que la sèrie (2.35) convergeix en un conjunt de mesura positiva. Per l'observació de més amunt, ha de convergir gairebé pertot. Per tant, existeix una funció mesurable $g(t)$ tal que

$$\lim_{n\to\infty} \sum_1^n c_k r_k(t) = g(t) \tag{2.36}$$

gairebé pertot. De (2.36) es desprèn que, per a tot real $\xi \neq 0$,

$$\lim_{n\to\infty} \exp\left[i\xi \sum_1^n c_k r_k(t)\right] = e^{i\xi g(t)}$$

gairebé pertot. Pel teorema de Lebesgue de la convergència dominada, concloem que

$$\lim_{n\to\infty} \int_0^1 \exp\left[i\xi \sum_1^n c_k r_k(t)\right] dt = \int_0^1 e^{i\xi g(t)}\, dt. \tag{2.37}$$

Però sabem que

$$\int_0^1 \exp\left[i\xi \sum_1^n c_k r_k(t)\right] dt = \prod_{k=1}^n \cos \xi c_k, \tag{2.38}$$

[†]Per a una funció $\phi(t)$ fitada i mesurable (i, per tant, integrable Lebesgue!) la demostració seria la següent. Tenim

$$I = \int_0^1 \phi(t)\, dt = \sum_{k=0}^{2^l-1} \int_{k/2^l}^{(k+1)/2^l} \phi(t)\, dt = 2^l \int_{k/2^l}^{(k+1)/2^l} \phi(t)\, dt.$$

Sigui t_0 tal que

$$\lim_{l\to\infty} 2^l \int_{k_l/2^l}^{(k_l+1)/2^l} \phi(t)\, dt = \phi(t_0)$$

per $k_l/2^l < t_0 < (k_l+1)/2^l$. Pel teorema fonamental del càlcul (vegeu § 2.5) gairebé tot t_0 gaudeix d'aquesta propietat. Així, $\phi(t_0) = I$ per a gairebé tot t_0. Si no suposem $\phi(t)$ fitada, n'hi ha prou amb aplicar l'argument a $e^{i\phi(t)}$. Aquesta demostració és deguda a Hartman i Kershner; el teorema va ser demostrat per primera vegada per Burstin d'una manera més complicada. El fet que la funció característica del conjunt de convergència és mesurable és clar, atès que el conjunt de convergència d'una sèrie de funcions mesurables és mesurable.

i deixem al lector la demostració que (2.34) i $c_n \to 0$ impliquen que

$$\lim_{n\to\infty} \prod_{k=1}^{n} \cos \xi c_k = 0.$$

Així,

$$\int_0^1 e^{i\xi g(t)}\, dt = 0 \tag{2.39}$$

per a tot real $\xi \neq 0$.

Considerem ara una successió $\xi_n \to 0$, però assegurem-nos que cada $\xi_n \neq 0$ (per exemple, $\xi_n = n^{-1}$); tenim que

$$\lim_{n\to\infty} \xi_n g(t) = 0$$

per a gairebé tot t i, en conseqüència

$$\lim_{n\to\infty} e^{i\xi_n g(t)} = 1$$

per a gairebé tot t.

Novament pel teorema de Lebesgue de la convergència dominada

$$\lim_{n\to\infty} \int_0^1 e^{i\xi_n g(t)}\, dt = 1$$

cosa que implica que $0 = 1$, una contradicció. Conseqüentment, (2.35) no podia convergir en un conjunt de mesura positiva i, per tant, ha de divergir gairebé pertot.

Aquest mètode de demostració fa servir la independència de les $r_k(t)$ de manera essencial [vegeu (2.38)], i no sembla immediatament aplicable a l'estudi de la sèrie

$$\sum_{k=1}^{\infty} c_k \sin 2\pi n_k t, \qquad \frac{n_{k+1}}{n_k} > q > 1$$

amb la condició

$$\sum_{1}^{\infty} c_k^2 = \infty.$$

De fet, el mètode encara es pot adaptar, però deixarem la discussió d'aquest punt per a més endavant.

PROBLEMES

1. Sigui $\sum_1^\infty c_k^2 = \infty$, $c_k \to 0$ i considerem la sèrie

$$\sum_{k=1}^{\infty} c_k \sin 2\pi 2^{k-1} t.$$

a) Demostreu que el límit

$$\lim_{n\to\infty}\int_0^1\left(\frac{\sum_1^n c_k\sin 2\pi 2^{k-1}t}{\sqrt{\sum_1^n c_k^2}}\right)^4 dt$$

existeix i trobeu el seu valor.

b) Demostreu que si $F_n(t)$, $0\le t\le 1$, és una successió de funcions tal que

$$\lim_{n\to\infty}\int_0^1 F_n^2(t)\,dt=\alpha,\qquad \lim_{n\to\infty}\int_0^1 F_n^4(t)\,dt=\beta$$

llavors la mesura del conjunt E en el qual $F_n(t)$ tendeix a 0 no pot excedir

$$1-\frac{\alpha^2}{\beta}.$$

c) Fent servir els apartats a) i b), demostreu que, en les condicions del problema, la sèrie

$$\sum_{k=1}^{\infty} c_k\sin 2\pi 2^{k-1}t$$

divergeix gairebé pertot.

2. El proper exemple mostra que el sinus del teorema del problema 1 no pot ser subtituït per una funció periòdica «arbitrària» $f(t)$ de període 1 (sotmesa, això sí, a la condició $\int_0^1 f(t)\,dt=0$).

Sigui

$$f(t)=\sin 2\pi t-\sin 4\pi t\,;$$

demostreu que

$$\sum_1^{\infty}\frac{1}{\sqrt{k}}f(2^{k-1}t)$$

convergeix gairebé pertot.

BIBLIOGRAFIA

E. Borel, «Les probabilités dénombrables et leurs applications arithmétiques», *Rend. Circ. Mat. Palermo*, **27** (1909), 247–271.

D. G. Champernowne, «The construction of decimals normal in the scale of ten», *Jour. London Math. Soc.*, **8** (1933), 254–260.

H. Steinhaus, «Les probabilités dénombrables et leur rapport à la théorie de la mesure», *Fund. Math.*, **4** (1922), 286–310.

H. Rademacher, «Einige Sätze über Reihen von Allgemeinen Orthogonalfunktionen», *Math. Ann.*, **87** (1922), 112–138.

M. Kac, «Convergence of certain gap series», *Ann. Math.*, **44** (1943), 411–415, (aquí es donen les referències als articles originals de Paley i Zygmund).

Ph. Hartman i R. Kershner, «The Structure of Monotone Functions», *Amer. Jour. Math.*, **59** (1937), 809–822.

Capítol 3

La llei normal

3.1. De Moivre.

En el § 2.1 hem vist la *llei feble dels grans nombres.* De Moivre va obtenir la següent formulació més precisa:

$$\lim_{n\to\infty} \mu \left\{ \omega_1\sqrt{n} < r_1(t) + \cdots + r_n(t) < \omega_2\sqrt{n} \right\} = \frac{1}{\sqrt{2\pi}} \int_{\omega_1}^{\omega_2} e^{-y^2/2}\, dy\,. \tag{3.1}$$

El lector no ha de tenir cap dificultat per a interpretar aquest resultat en termes de probabilitat. Una demostració elemental es basa en la fórmula (1.17) del capítol 1, i (3.1) esdevé equivalent a la fórmula, purament combinatòria,

$$\lim_{n\to\infty} \sum_{\frac{n}{2}+\frac{\omega_1}{2}\sqrt{n}<l<\frac{n}{2}+\frac{\omega_2}{2}\sqrt{n}} \frac{1}{2^n}\binom{n}{l} = \frac{1}{\sqrt{2\pi}} \int_{\omega_1}^{\omega_2} e^{-y^2/2}\, dy\,. \tag{3.2}$$

El resultat (3.2) es pot aconseguir amb un ús adequat de la fórmula de Stirling, però aquest camí de demostració té l'inconvenient d'enfosquir la naturalesa del teorema. Els intents per a generalitzar (3.1) varen proporcionar una de les motivacions més importants per a desenvolupar les eines analítiques de la teoria de la probabilitat. Markoff va proposar un mètode molt potent, però no va aconseguir fer-lo prou rigorós. Uns trenta anys més tard, Lévy el va justificar. Les dues properes seccions es dediquen al mètode de Markoff.

3.2. La idea

Sigui

$$g(x) = \begin{cases} 1, & \omega_1 < x < \omega_2, \\ 0, & \text{en cas contrari.} \end{cases} \tag{3.3}$$

De la teoria elemental de les integrals de Fourier hom sap que

$$g(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{i\omega_2\xi} - e^{i\omega_1\xi}}{i\xi} e^{-ix\xi}\, d\xi \tag{3.4}$$

amb la convenció habitual que, per $x = \omega_1$ i $x = \omega_2$, hom obté 1/2. Ara, llevat que ω_1 i ω_2 siguin múltiples enters de $\sqrt{n}$, tenim que

$$\begin{aligned} &\mu \left\{ \omega_1 < \frac{r_1(t) + \cdots + r_n(t)}{\sqrt{n}} < \omega_2 \right\} \\ &= \int_0^1 g\left(\frac{r_1(t) + \cdots + r_n(t)}{\sqrt{n}} \right) dt \\ &= \int_0^1 \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{i\omega_2\xi} - e^{i\omega_1\xi}}{i\,\xi} \\ &\qquad \times \exp\left(-i\,\xi \frac{r_1(t) + \cdots + r_n(t)}{\sqrt{n}} \right) d\xi\, dt\,. \end{aligned} \tag{3.5}$$

Intercanviant l'ordre d'integració [que en aquest cas es justifica fàcilment atès que $r_1(t) + \cdots + r_n(t)$ només pren un nombre finit de valors] obtenim

$$\begin{aligned} &\mu \left\{ \omega_1 < \frac{r_1(t) + \cdots + r_n(t)}{\sqrt{n}} < \omega_2 \right\} \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{i\omega_2\xi} - e^{i\omega_1\xi}}{i\,\xi} \\ &\qquad \times \left[\int_0^1 \exp\left(-i\,\xi \frac{r_1(t) + \cdots + r_n(t)}{\sqrt{n}} \right) dt \right] d\xi \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{i\omega_2\xi} - e^{i\omega_1\xi}}{i\,\xi} \left(\cos \frac{\xi}{\sqrt{n}} \right)^n d\xi\,. \end{aligned} \tag{3.6}$$

Ara, per a tot real ξ,

$$\lim_{n\to\infty} \left(\cos \frac{\xi}{\sqrt{n}} \right)^n = e^{-\xi^2/2}\,, \tag{3.7}$$

i és molt temptador arribar a la conclusió que

$$\begin{aligned} &\lim_{n\to\infty} \mu \left\{ \omega_1 < \frac{r_1(t) + \cdots + r_n(t)}{\sqrt{n}} < \omega_2 \right\} \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{i\omega_2\xi} - e^{i\omega_1\xi}}{i\,\xi} e^{-\xi^2/2}\, d\xi = \frac{1}{\sqrt{2\pi}} \int_{\omega_1}^{\omega_2} e^{-y^2/2}\, dy\,. \end{aligned} \tag{3.8}$$

Quin és el problema d'aquest mètode? L'únic pas que requereix justificació és l'intercanvi de les operacions d'integració i de pas al límit quan $n \to \infty$. Malauradament, els límits d'integració són $-\infty$ i $+\infty$, i la funció

$$\frac{e^{i\omega_2\xi} - e^{i\omega_1\xi}}{i\,\xi}$$

no és *absolutament* integrable.

Markoff, tot i que era un matemàtic de primera, va ser incapaç de superar aquesta dificultat i va abandonar el mètode!

Els físics, que tenen un concepte del rigor menys estricte que el nostre, encara anomenen el mètode *mètode de Markoff*, mentre que els matemàtics amb prou feines reconeixen el seu origen.

3.3. El mètode de Markoff fet rigorós

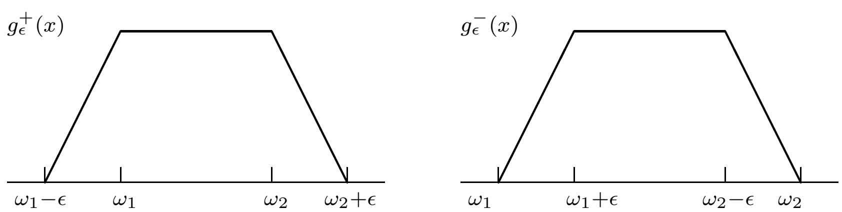
De fet, la justificació del mètode de Markoff és prou senzilla. Es basa en una simple idea d'abast força ampli.

Primer de tot, examinem la fórmula (3.4). És simplement la fórmula de Fourier

$$g(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(y) e^{i\xi(y-x)}\, dy\, d\xi \tag{3.9}$$

aplicada a la funció concreta (3.3).

Introduïm ara dues funcions auxiliars, $g_\epsilon^+(x)$ i $g_\epsilon^-(x)$ amb les gràfiques que es mostren a la figura ($\epsilon > 0$, $2\epsilon < \omega_2 - \omega_1$).*



Tenim que

$$g_\epsilon^-(x) \le g(x) \le g_\epsilon^+(x) \tag{3.10}$$

i, conseqüentment,

$$\begin{aligned}\int_0^1 g_\epsilon^- \left(\frac{r_1(t) + \cdots + r_n(t)}{\sqrt{n}} \right) dt \qquad\qquad (3.11)\\ \le \mu \left\{ \omega_1 < \frac{r_1(t) + \cdots + r_n(t)}{\sqrt{n}} < \omega_2 \right\} \\ \le \int_0^1 g_\epsilon^+ \left(\frac{r_1(t) + \cdots + r_n(t)}{\sqrt{n}} \right) dt\,.\end{aligned}$$

Ara, les funcions

$$G_\epsilon^-(\xi) = \int_{-\infty}^{\infty} g_\epsilon^-(y) e^{iy\xi}\, dy \quad \text{i} \quad G_\epsilon^+(\xi) = \int_{-\infty}^{\infty} g_\epsilon^+(y) e^{iy\xi}\, dy$$

*L'alçària de cada gràfica és igual a 1.

són funcions de ξ *absolutament integrables* a $(-\infty, \infty)$, i, a causa d'això, l'argument del § 3.2 demostra de manera *rigorosa* que

$$\begin{aligned}\lim_{n\to\infty} \int_0^1 g_\epsilon^- \left(\frac{r_1(t)+\cdots+r_n(t)}{\sqrt{n}}\right) dt \\ = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-\xi^2/2}\int_{-\infty}^{\infty} g_\epsilon^-(y)e^{i\xi\, y}\, dy\, d\xi \\ = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} g_\epsilon^-(y)e^{-y^2/2}\, dy\end{aligned} \tag{3.12}$$

i, també, que

$$\begin{aligned}\lim_{n\to\infty} \int_0^1 g_\epsilon^+ \left(\frac{r_1(t)+\cdots+r_n(t)}{\sqrt{n}}\right) dt \\ = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-\xi^2/2}\int_{-\infty}^{\infty} g_\epsilon^+(y)e^{i\xi\, y}\, dy\, d\xi \\ = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} g_\epsilon^+(y)e^{-y^2/2}\, dy\,.\end{aligned} \tag{3.13}$$

Combinant (3.12) i (3.13) amb (3.11) obtenim

$$\begin{aligned}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} g_\epsilon^-(y)e^{-y^2/2}\, dy \\ \le \liminf_{n\to\infty} \mu\left\{\omega_1 < \frac{r_1(t)+\cdots+r_n(t)}{\sqrt{n}} < \omega_2\right\} \\ \le \limsup_{n\to\infty} \mu\left\{\omega_1 < \frac{r_1(t)+\cdots+r_n(t)}{\sqrt{n}} < \omega_2\right\} \\ \le \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} g_\epsilon^+(y)e^{-y^2/2}\, dy\,.\end{aligned} \tag{3.14}$$

Atès que (3.14) és vàlid per a tot $\epsilon > 0$, obtenim immediatament que

$$\begin{aligned}\lim_{n\to\infty} \mu\left\{\omega_1 < \frac{r_1(t)+\cdots+r_n(t)}{\sqrt{n}} < \omega_2\right\} \\ = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} g(y)e^{-y^2/2}\, dy = \frac{1}{\sqrt{2\pi}}\int_{\omega_1}^{\omega_2} e^{-y^2/2}\, dy\,.\end{aligned} \tag{3.15}$$

PROBLEMES

1. El 1917, el desaparegut H. Weyl va demostrar que per a tot irracional α la successió $\alpha_n = n\alpha - [n\alpha]$, $n = 1, 2, \cdots$, està equidistribuïda a

$(0, 1)$. En altres paraules, si $0 \leq \omega_1 < \omega_2 \leq 1$ i $k_n(\omega_1, \omega_2)$ denota el nombre de α_j, $1 \leq j \leq n$, que cauen dintre (ω_1, ω_2), llavors

$$\lim_{n\to\infty} \frac{k_n(\omega_1, \omega_2)}{n} = \omega_2 - \omega_1 \,.$$

Considereu la funció $g(x)$, periòdica de període 1, donada per (3.3) a $(0, 1)$ i, fent servir sèries de Fourier en comptes d'integrals de Fourier, demostreu el teorema de Weyl.

2. Feu servir el mètode de Markoff per a demostrar la fórmula de Laplace

$$\lim_{x\to\infty} e^{-x} \sum_{x+\omega_1\sqrt{x}<k<x+\omega_2\sqrt{x}} \frac{x^k}{k!} = \frac{1}{\sqrt{2\pi}} \int_{\omega_1}^{\omega_2} e^{-y^2/2}\, dy \,.$$

3.4. El mètode vist més de prop

Una anàlisi de l'argumentació del § 3.3 revela que, en realitat, hem demostrat el teorema següent:

Sigui $f_n(t)$, $0 \leq t \leq 1$, una successió de funcions mesurables tals que per a tot real ξ

$$\lim_{n\to\infty} \int_0^1 e^{i\xi f_n(t)}\, dt = e^{-\xi^2/2} \,. \tag{3.16}$$

Llavors,

$$\lim_{n\to\infty} \mu \{ \omega_1 \leq f_n(t) < \omega_2 \} = \frac{1}{\sqrt{2\pi}} \int_{\omega_1}^{\omega_2} e^{-y^2/2}\, dy \,. \tag{3.17}$$

Sigui

$$\sigma_n(\omega) = \mu \{ f_n(t) < \omega \} \,, \tag{3.18}$$

llavors $\sigma_n(\omega)$ té les propietats següents:

1a $\sigma_n(-\infty) = 0$, $\sigma_n(+\infty) = 1$.

2a $\sigma_n(\omega)$ és no decreixent.

3a $\sigma_n(\omega)$ és contínua per l'esquerra.

(Observem que la propietat 3a és una conseqüència de l'additivitat completa de la mesura de Lebesgue.) Una funció que posseeix les propietats 1a, 2a i 3a s'anomena una *funció de distribució*. Ara,

$$\int_0^1 e^{i\xi f_n(t)}\, dt = \int_{-\infty}^{\infty} e^{i\xi\omega}\, d\sigma_n(\omega) \,. \tag{3.19}$$

i el nostre teorema també es pot enunciar així.

Si una successió de funcions de distribució $\sigma_n(\omega)$ és tal que, per a tot real ξ

$$\lim_{n\to\infty} \int_{-\infty}^{\infty} e^{i\xi\omega}\, d\sigma_n(\omega) = e^{-\xi^2/2}\,, \tag{3.20}$$

llavors,

$$\sigma_n(\omega_2) - \sigma_n(\omega_1) \longrightarrow G(\omega_2) - G(\omega_1), \tag{3.21}$$

on

$$G(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\omega} e^{-y^2/2}\, dy\,. \tag{3.22}$$

El lector atent haurà notat un petit forat lògic. Si només ens donen una successió de funcions de distribució $\sigma_n(\omega)$, l'última formulació només es dedueix de la precedent en el cas que pUguem exhibir una successió de funcions $f_n(t)$, $0 \le t \le 1$, tals que

$$\mu\{\, f_n(t) < \omega \,\} = \sigma_n(\omega). \tag{3.23}$$

Podríem arreglar-ho repetint, essencialment, l'argument del § 3.3, però de fet, la construcció de les funcions $f_n(t)$ és extremadament simple. Podem prendre com a $f_n(t)$ la funció *inversa* de $\sigma_n(\omega)$, amb el benentès que els intervals de constància de $\sigma_n(\omega)$ es reflecteixen en discontinuïtats de $f_n(t)$ i les discontinuïtats de $\sigma_n(\omega)$ es reflecteixen en intervals de constància de $f_n(t)$. Deixem els detalls al lector. La conclusió que (3.20) implica (3.21) és un cas especial d'un teorema general molt important conegut com el teorema de continuïtat de la transformada de Fourier-Stieltjes. Aquest teorema es pot enunciar així.

Si $\sigma_n(\omega)$ és una successió de funcions de distribució tals que, per a tot real ξ

$$\lim_{n\to\infty} \int_{-\infty}^{\infty} e^{i\xi\omega}\, d\sigma_n(\omega) = c(\xi) \tag{3.24}$$

i $c(\xi)$ és una funció contínua per $\xi = 0$, llavors existeix una única funció de distribució, $\sigma(\omega)$ tal que

$$\int_{-\infty}^{\infty} e^{i\xi\omega}\, d\sigma(\omega) = c(\xi) \tag{3.25}$$

i

$$\lim_{n\to\infty} \sigma_n(\omega) = \sigma(\omega) \tag{3.26}$$

per a tot ω pel qual $\sigma(\omega)$ sigui contínua.

La demostració, a més a més de les idees que ja s'han explicat, utilitza l'anomenat *principi de selecció de Helly* i és una mica massa tècnica per a ser presentada aquí. En conseqüència l'ometem, encara que farem servir amb tota llibertat el teorema en allò que segueix.

PROBLEMES

1. Siguin $f_n(t),\ 0 \le t \le 1$, tals que per $k = 0, 1, 2, \cdots$. Tenim que

$$\lim_{n\to\infty} \int_0^1 f_n^k(t)\, dt$$
$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} y^k e^{-y^2/2}\, dy = \begin{cases} 0, & k \text{ imparell,} \\ \dfrac{k!}{2^{k/2}\left(\frac{k}{2}\right)!}\,, & k \text{ parell.} \end{cases}$$

Demostreu que, per a tot real ξ

$$\lim_{n\to\infty} \int_0^1 e^{i\xi f_n(t)}\, dt = e^{-\xi^2/2}$$

i que, en conseqüència, (3.17) és vàlid.

2. Sigui $\{n_m\}$ una successió d'enters tals que

$$\lim_{m\to\infty} \frac{n_{m+1}}{n_m} = \infty.$$

Demostreu que per a $k = 0, 1, 2, \cdots$

$$\lim_{m\to\infty} \int_0^1 \left(\sqrt{2}\, \frac{\cos 2\pi n_1 t + \cos 2\pi n_2 t + \cdots + \cos 2\pi n_m t}{\sqrt{m}}\right)^k dt$$
$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} y^k e^{-y^2/2}\, dy$$

i, per tant,

$$\lim_{m\to\infty} \mu \left\{ \omega_1 < \sqrt{2}\, \frac{\cos 2\pi n_1 t + \cos 2\pi n_2 t + \cdots + \cos 2\pi n_m t}{\sqrt{m}} < \omega_2 \right\}$$
$$= \frac{1}{\sqrt{2\pi}} \int_{\omega_1}^{\omega_2} e^{-y^2/2}\, dy\,.$$

Nota. Pel mateix mètode, però fent servir arguments combinatoris més enginyosos, hom pot demostrar que si

$$\sum_{k=1}^{\infty} c_k^2 = \infty \qquad \text{i} \qquad |\, c_k \,| < M\,,$$

i si

$$\frac{n_{k+1}}{n_k} > q > 1,$$

llavors,

$$\lim_{n\to\infty} \mu \left\{ \omega_1 < \sqrt{2}\, \frac{\sum_1^n c_k \cos 2\pi n_k t}{\sqrt{\sum_1^n c_k^2}} < \omega_2 \right\} = \frac{1}{\sqrt{2\pi}} \int_{\omega_1}^{\omega_2} e^{-y^2/2}\, dy\,.$$

En particular, es dedueix que $\sum_1^\infty c_k^2 = \infty$ implica la divergència gairebé pertot de $\sum_1^\infty c_k \cos 2\pi n_k t$ (l'argument, òbviament, també es pot aplicar si hom canvia el cosinus pel sinus).

Com el lector deu haver observat, tot això està íntimament relacionat amb el mètode emprat en l'exemple 2 del § 2.5.

3. Sigui $\sigma_n(\omega)$ una funció de distribució, i sigui

$$c(\xi) = \int_{-\infty}^{\infty} e^{i\xi\omega}\, d\sigma(\omega)\,.$$

Demostreu que

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T |\, c(\xi)\,|^2\, d\xi = \text{la suma dels quadrats dels salts de } \sigma(\omega).$$

(Aquest teorema tan simple, però alhora tan elegant, és degut a N. Wiener. Una demostració es pot basar en l'observació que

$$c(\xi) = \int_0^1 e^{i\xi f(t)}\, dt\,,$$

on $f(t)$ és la inversa de $\sigma(\omega)$, tal com s'ha descrit més amunt. Així

$$\frac{1}{T} \int_0^T |\, c(\xi)\,|^2\, d\xi = \int_0^1 \int_0^1 \frac{1}{T} \int_0^T e^{i\xi(f(s)-f(t))}\, d\xi\, ds\, dt$$

i

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T e^{i\xi(f(t)-f(s))}\, d\xi = \begin{cases} 0, & f(t) \neq f(s), \\ 1, & f(t) = f(s). \end{cases}$$

Pel teorema de la convergència fitada, es dedueix que el límit

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T |\, c(\xi)\,|^2\, d\xi$$

existeix i és igual a la mesura plana dels punts (t, s), $(0 \le t, s \le 1)$, pels quals $f(t) = f(s)$. Això és equivalent al nostre teorema.)

4. Demostreu que

$$f(t) = \sum_{k=1}^{\infty} c_k r_k(t), \quad \sum_1^{\infty} c_k^2 < \infty,$$

no pot ser constant en un conjunt de mesura positiva, si no és que totes les c, llevat d'un nombre finit, són 0.

3.5. ¿Una llei de la natura o un teorema matemàtic?

Per a concloure aquest capítol, considerarem un exemple que és molt instructiu, tant des del punt de vista conceptual com tècnic.

En primer lloc, necessitarem tres definicions.

1a. *La mesura relativa.* Sigui A un conjunt de nombres reals, i considerem el subconjunt d'A que cau a $(-T,T)$, és a dir, $A \cap (-T,T)$. La mesura relativa, $\mu_R\{A\}$, d'A es defineix com el límit

$$\mu_R\{A\} = \lim_{T\to\infty} \frac{1}{2T}\, \mu\, \{\, A \cap (-T,T)\,\}\,, \tag{3.27}$$

quan aquest límit existeix. La mesura relativa no és *completament additiva*, perquè si $A_i = (i, i+1)$, $i = 0, \pm 1, \pm 2, \cdots$, llavors

$$\mu_R \left\{ \bigcup_{i=-\infty}^{\infty} A_i \right\} = 1,$$

mentre que

$$\sum_{i=-\infty}^{\infty} \mu_R\{A_i\} = 0.$$

2a. *El valor mitjà d'una funció.* El valor mitjà $M\{f(t)\}$ de la funció $f(t)$, $-\infty < t < \infty$, es defineix com el límit

$$M\{f(t)\} = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} f(t)\, dt, \tag{3.28}$$

quan el límit existeix.

3a. *La independència lineal d'un conjunt de nombres reals.* Els nombres reals $\lambda_1, \lambda_2,\ \lambda_3, \cdots$ s'anomenen *linealment independents* (o *independents*) *sobre el cos dels racionals* quan l'única solució $(k_1, k_2, \cdots)$ en *enters* de l'equació

$$k_1\lambda_1 + k_2\lambda_2 + \cdots = 0 \tag{3.29}$$

és

$$k_1 = k_2 = \cdots = 0.$$

L'exemple més conegut de nombres linealment independents és la successió

$$\log p_1, \log p_2, \log p_3, \cdots \tag{3.30}$$

dels logaritmes dels nombres primers $(p_1 = 2, p_2 = 3, \cdots)$. Com el lector no deu haver deixat de notar, la independència lineal de (3.30) és equivalent al teorema de factorització única. Aquesta senzilla i bonica observació va ser feta el 1919 per H. Bohr, que en va fer el punt de partida d'un nou atac sobre molts problemes relacionats amb la cèlebre funció ζ de Riemann.

Siguin ara $\lambda_1, \lambda_2, \cdots$ linealment independents i considerem la funció

$$\sqrt{2}\,\frac{\cos\lambda_1 t + \cdots + \cos\lambda_n t}{\sqrt{n}}\,. \tag{3.31}$$

Sigui $A_n(\omega_1, \omega_2)$ el conjunt en el qual

$$\omega_1 < \sqrt{2}\,\frac{\cos\lambda_1 t + \cdots + \cos\lambda_n t}{\sqrt{n}} < \omega_2\,. \tag{3.32}$$

Demostrarem ara que la mesura relativa, $\mu_R\{\,A_n(\omega_1, \omega_2)\,\}$, està definida i, a més a més, veurem que

$$\lim_{n\to\infty} \mu_R\{\,A_n(\omega_1, \omega_2)\,\} = \frac{1}{\sqrt{2\pi}} \int_{\omega_1}^{\omega_2} e^{-y^2/2}\,dy\,. \tag{3.33}$$

Fent servir la notació del § 3.3, tenim

$$\begin{aligned}
&\frac{1}{2T}\int_{-T}^{T} g_\epsilon^- \left(\sqrt{2}\,\frac{\cos\lambda_1 t + \cdots + \cos\lambda_n t}{\sqrt{n}}\right) dt \\
&\qquad \le \frac{1}{2T}\int_{-T}^{T} g \left(\sqrt{2}\,\frac{\cos\lambda_1 t + \cdots + \cos\lambda_n t}{\sqrt{n}}\right) dt \\
&\qquad \le \frac{1}{2T}\int_{-T}^{T} g_\epsilon^+ \left(\sqrt{2}\,\frac{\cos\lambda_1 t + \cdots + \cos\lambda_n t}{\sqrt{n}}\right) dt
\end{aligned} \tag{3.34}$$

i

$$\begin{aligned}
&\frac{1}{2T}\int_{-T}^{T} g_\epsilon^\pm \left(\sqrt{2}\,\frac{\cos\lambda_1 t + \cdots + \cos\lambda_n t}{\sqrt{n}}\right) dt \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty} G_\epsilon^\pm(\xi) \left[\frac{1}{2T}\int_{-T}^{T} \exp\left(i\xi\sqrt{2}\,\frac{\cos\lambda_1 t + \cdots + \cos\lambda_n t}{\sqrt{n}}\right) dt\right] d\xi\ ^*,
\end{aligned} \tag{3.35}$$

on $G_\epsilon^+(\xi)$ i $G_\epsilon^-(\xi)$ són *absolutament integrables* a $(-\infty, \infty)$. (D'aquesta manera, l'intercanvi d'ordre d'integració es justifica fàcilment.)

Demostrem ara que

$$\lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} \exp\left(i\xi\sqrt{2}\,\frac{\cos\lambda_1 t + \cdots + \cos\lambda_n t}{\sqrt{n}}\right) dt = J_0^n\left(\sqrt{2}\frac{\xi}{\sqrt{n}}\right), \tag{3.36}$$

on J_0 és la coneguda funció de Bessel.

Ens limitarem a veure el cas $n = 2$, atès que la demostració per un valor arbitrari de n és la mateixa.

---

* Recordeu que fem servir l'abreviació $G_\epsilon^\pm(\xi) = \int_{-\infty}^{\infty} g_\epsilon^\pm(x) e^{i\xi x}\,dx\,.$

Tenim (posant $\eta = \xi\sqrt{2}/\sqrt{n}$),

$$\frac{1}{2T}\int_{-T}^{T} e^{i\eta(\cos\lambda_1 t+\cos\lambda_2 t)}$$
$$= \sum_{k,\,l=0}^{\infty} \frac{(i\eta)^k (i\eta)^l}{k!\,l!}\,\frac{1}{2T}\int_{-T}^{T} \cos^k\lambda_1 t\,\cos^l\lambda_2 t\,dt\,, \quad (3.37)$$

i hem d'arribar a

$$\lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}\cos^k\lambda_1 t\,\cos^l\lambda_2 t\,dt = M\{\cos^k\lambda_1 t\,\cos^l\lambda_2 t\}\,.$$

Ara,

$$\cos^k\lambda_1 t\,\cos^l\lambda_2 t = \frac{1}{2^k}\,\frac{1}{2^l}\left(e^{i\lambda_1 t}+e^{-i\lambda_1 t}\right)^k\left(e^{i\lambda_2 t}+e^{-i\lambda_2 t}\right)^l$$
$$= \frac{1}{2^k}\,\frac{1}{2^l}\sum_{r=0}^{k}\sum_{s=0}^{l}\binom{k}{r}\binom{l}{s}e^{i[(2r-k)\lambda_1+(2s-l)\lambda_2]t}\,,$$

i

$$M\{e^{i\alpha t}\} = \lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}e^{i\alpha t}\,dt = \begin{cases} 1, & \alpha = 0,\\ 0, & \alpha\neq 0.\end{cases}$$

A causa de la independència lineal tenim que l'expressió

$$(2r-k)\lambda_1 + (2s-l)\lambda_2$$

només pot ser 0 si $2r = k$ i $2s = l$. En conseqüència, es dedueix immediatament que

$$M\{\cos^k\lambda_1 t\,\cos^l\lambda_2 t\} = \frac{1}{2^k}\binom{k}{k/2}\frac{1}{2^l}\binom{l}{l/2} \quad (3.38)$$

en cas que tant k com l siguin parells, i el valor és 0 en d'altres casos. Podem escriure (3.38) de la manera següent:

$$M\{\cos^k\lambda_1 t\,\cos^l\lambda_2 t\} = M\{\cos^k\lambda_1 t\}M\{\cos^l\lambda_2 t\}\,, \quad (3.39)$$

que, combinada amb (3.37), ens proporciona

$$M\{e^{i\eta(\cos\lambda_1 t+\cos\lambda_2 t)}\} = M\{e^{i\eta\cos\lambda_1 t}\}M\{e^{i\eta\cos\lambda_2 t}\}\,. \quad (3.40)$$

És clar que

$$M\{e^{i\eta\,\cos\lambda t}\} = \frac{1}{2\pi}\int_0^{2\pi} e^{i\eta\cos\theta}\,d\theta = J_0(\eta) \quad (3.41)$$

i, per tant [a partir de (3.40)],

$$M\{e^{i\eta(\cos\lambda_1 t+\cos\lambda_2 t)}\} = J_0^2(\eta)\,.$$

Podem considerar, doncs, (3.36) com a demostrat. Fent $T \to \infty$ a (3.34) i fent servir (3.35) i (3.36), obtenim

$$
\begin{aligned}
&\frac{1}{2\pi} \int_{-\infty}^{\infty} G_\epsilon^-(\xi)\, J_0^n \left(\sqrt{2}\, \frac{\xi}{\sqrt{n}} \right) d\xi \\
&\le \liminf_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} g\left(\sqrt{2}\, \frac{\cos\lambda_1 t + \cdots + \cos\lambda_n t}{\sqrt{n}} \right) dt \\
&\le \limsup_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} g\left(\sqrt{2}\, \frac{\cos\lambda_1 t + \cdots + \cos\lambda_n t}{\sqrt{n}} \right) dt \\
&\le \frac{1}{2\pi} \int_{-\infty}^{\infty} G_\epsilon^+\, J_0^n \left(\sqrt{2}\, \frac{\xi}{\sqrt{n}} \right) d\xi\,.
\end{aligned}
\tag{3.42}
$$

És ben conegut que, per $\eta \to \pm\infty$,

$$
J_0(\eta) = \mathcal{O}\left(\frac{1}{\sqrt{|\,\eta\,|}} \right),
$$

i, en conseqüència, per $n \ge 3$,

$$
J_0^n \left(\sqrt{2}\, \frac{\xi}{\sqrt{n}} \right)
$$

és absolutament integrable respecte de ξ. Això implica que ($n \ge 3$)

$$
\lim_{\epsilon\to 0} \frac{1}{2\pi} \int_{-\infty}^{\infty} G_\epsilon^-(\xi)\, J_0^n \left(\sqrt{2}\, \frac{\xi}{\sqrt{n}} \right) d\xi = \lim_{\epsilon\to 0} \frac{1}{2\pi} \int_{-\infty}^{\infty} G_\epsilon^+(\xi)\, J_0 \left(\sqrt{2}\, \frac{\xi}{\sqrt{n}} \right) d\xi
$$

i, per tant, que el límit

$$
\lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} g\left(\sqrt{2}\, \frac{\cos\lambda_1 t + \cdots + \cos\lambda_n t}{\sqrt{n}} \right) dt = \mu_R\,\{\, A_n(\omega_1, \omega_2)\,\}
$$

existeix!* Ara, (3.42) es pot escriure de la manera següent:

$$
\begin{aligned}
\frac{1}{2\pi} \int_{-\infty}^{\infty} G_\epsilon^-(\xi)\, J_0^n \left(\sqrt{2}\, \frac{\xi}{\sqrt{n}} \right) d\xi &\le \mu_R\,\{\, A_n(\omega_1, \omega_2)\,\} \\
&\le \frac{1}{2\pi} \int_{-\infty}^{\infty} G_\epsilon^+(\xi)\, J_0^n \left(\sqrt{2}\, \frac{\xi}{\sqrt{n}} \right) d\xi
\end{aligned}
$$

i hom comprova fàcilment que

$$
\lim_{n\to\infty} J_0^n \left(\sqrt{2}\, \frac{\xi}{\sqrt{n}} \right) = e^{-\xi^2/2}\,.
$$

*Per a $n = 1$ i $n = 2$ això també és veritat, però la demostració s'ha de modificar.

La demostració de (3.33) es pot ara completar exactament igual que en el § 3.3. Si considerem

$$q_n(t) = \sqrt{2}\,\frac{\cos\lambda_1 t + \cdots + \cos\lambda_n t}{\sqrt{n}}$$

com el resultat de la superposició de vibracions amb freqüències incommensurables, el teorema de l'equació (3.33) aporta informació precisa sobre el temps relatiu que $q_n(t)$ passa entre ω_1 i ω_2. El fet que això ens porti a la llei normal

$$\frac{1}{2\pi}\int_{\omega_1}^{\omega_2} e^{-y^2/2}\,dy\,,$$

generalment associada amb fenòmens aleatoris, és potser una indicació que els punts de vista determinista i probabilista no són tan irreconciliables com pot semblar a primer cop d'ull. Penetrar més en aquesta qüestió ens portaria massa lluny, però podem recordar unes paraules de Poincaré que va dir (segur que mig de broma) que alguna cosa de misteriosa hi devia haver en la llei normal, perquè els matemàtics opinaven que era una llei de la naturalesa, mentre que els físics estaven convençuts que es tractava d'un teorema matemàtic.

PROBLEMES

1. Demostreu que si $\lambda_1, \cdots, \lambda_n$ són linealment independents, llavors les funcions $\cos\lambda_1 t, \cdots, \cos\lambda_n t$ són estadísticament independents, és a dir, per a tots els reals $\alpha_1, \cdots, \alpha_n$

$$\mu_R\{\,\cos\lambda_1 t < \alpha_1, \cdots, \cos\lambda_n t < \alpha_n\,\} = \prod_{k=1}^{n} \mu_R\{\,\cos\lambda_k t < \alpha_k\,\}\,.$$

[Aquesta és, evidentment, la propietat que és l'essència de la demostració de (3.33)].

2. Sigui $s = \sigma + it$, $\sigma > 1$, i considerem la funció ζ de Riemann

$$\zeta(s) = \sum_{n=1}^{\infty}\frac{1}{n^s} = \prod_p \frac{1}{1-\dfrac{1}{p^s}}\,.$$

Demostreu que per $l > 0$

$$M\{|\,\zeta(\sigma+it)\,|^l\} = M\left\{\frac{1}{|\,\zeta(\sigma+it)\,|^{l-2}}\right\}\zeta^{l-1}(2\sigma)\,.$$

BIBLIOGRAFIA

A. Markoff, *Wahrscheinlichkeitsrechnung*, Teubner, Leipzig, 1912.

M. Loève, *Probability Theory*, Van Nostrand and Co., Princeton, 1955. Aquest llibre conté una exposició detallada de la teoria de les funcions de distribució i, en particular, de l'obra de Paul Lévy.

M. Kac i H. Steinhaus, ≪Sur les fonctions indépendantes IV≫, *Studia Math.*, **7** (1938), 1–15.

Capítol 4

Els nombres primers juguen un joc d'atzar

4.1. Funcions de la teoria de nombres, densitat, independència

Una funció de la teoria de nombres* $f(n)$ és una funció definida en els enters positius $1, 2, 3, \cdots$. El valor mitjà de f, $M\{f(n)\}$, es defineix com el límit (quan existeix),

$$M\{f(n)\} = \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} f(n). \tag{4.1}$$

Si A és un conjunt d'enters positius, denotem com a $A(N)$ el nombre d'elements d'A que es troben entre els primers N enters. La *densitat* d'A és el límit

$$\lim_{N\to\infty} \frac{A(N)}{N} = D\{A\} \tag{4.2}$$

quan existeix. El concepte de densitat és anàleg al concepte de mesura relativa (vegeu § 3.5) i, com en el cas de la mesura relativa, no és completament additiva. Considerem els enters divisibles per un primer p. La densitat del conjunt d'aquests enters és clarament $1/p$. Prenem ara el conjunt dels enters divisibles alhora per p i q (on q és un altre primer). Ser divisible per p i q és equivalent a ser divisible per pq i, en conseqüència, la densitat del nou conjunt és $1/pq$. Ara

$$\frac{1}{pq} = \frac{1}{p} \cdot \frac{1}{q}\,, \tag{4.3}$$

i podem interpretar aquest fet dient que els «esdeveniments» ser divisible per p i ser divisible per q són independents. Això es manté, evidentment, per a qualsevol nombre de primers i podem dir, d'una manera una mica

*N. del tr. Avui en diríem una *funció entera*.

pintoresca, que els primers juguen un joc d'atzar! Aquesta observació tan senzilla, gairebé trivial, és el començament d'un nou desenvolupament que relaciona d'una manera significativa, d'una banda la teoria de nombres i d'una altra, la teoria de la probabilitat.

Il·lustrarem en detall alguns dels aspectes elementals d'aquest desenvolupament i n'esbossarem breument els aspectes més avançats.

4.2. L'estadística de la funció ϕ d'Euler

Anomenem $\phi(n)$ el nombre d'enters, menors que n, que són relativament primers amb n. Aquesta funció de la teoria de nombres va ser introduïda per Euler i té un gran nombre d'aplicacions, a banda d'un considerable interès per si mateixa.

Hom comprova immediatament que si

$$(m, n) = 1$$

(és a dir, m i n són relativament primers entre si), llavors

$$\phi(mn) = \phi(m)\phi(n) \tag{4.4}$$

i

$$\phi(p^\alpha) = p^\alpha - p^{\alpha-1}. \tag{4.5}$$

Així,

$$\phi(n) = \prod_{\substack{p^\alpha \mid n \\ p^{\alpha+1} \nmid n}} \left(p^\alpha - p^{\alpha-1}\right), \tag{4.6}$$

o bé, atès que

$$n = \prod_{\substack{p^\alpha \mid n \\ p^{\alpha+1} \nmid n}} p^\alpha \tag{4.7}$$

(la factorització única!),

$$\frac{\phi(n)}{n} = \prod_{p \mid n} \left(1 - \frac{1}{p}\right). \tag{4.8}$$

Introduïm ara les funcions $\rho_p(n)$ definides de la manera següent:

$$\rho_p(n) = \begin{cases} 1, & p \mid n, \\ \\ 0, & p \nmid n. \end{cases} \tag{4.9}$$

En termes d'aquestes funcions $\rho_p(n)$, podem escriure

$$\frac{\phi(n)}{n} = \prod_{p} \left(1 - \frac{\rho_p(n)}{p}\right). \tag{4.10}$$

Observem ara que, si ϵ_j és o bé 0 o bé 1, llavors

$$\begin{aligned} &D\{\rho_{p_1}(n) = \epsilon_1,\ \rho_{p_2}(n) = \epsilon_2, \cdots, \rho_{p_k}(n) = \epsilon_k\} \\ &= D\{\rho_{p_1}(n) = \epsilon_1\} \cdot D\{\rho_{p_2}(n) = \epsilon_2\} \cdots D\{\rho_{p_k}(n) = \epsilon_k\}\,. \end{aligned} \tag{4.11}$$

Això és simplement una altra manera de dir que els «esdeveniments» ser divisible per $p_1, p_2, \cdots, p_k$ són independents (o que les funcions $\rho_p(n)$ són independents).

La propietat (4.11) implica que

$$\begin{aligned} &M\left\{\prod_{p\le p_k}\left(1-\frac{\rho_p(n)}{p}\right)\right\} \\ &\qquad = \prod_{p\le p_k} M\left\{\left(1-\frac{\rho_p(n)}{p}\right)\right\} = \prod_{p\le p_k}\left(1-\frac{1}{p^2}\right), \end{aligned} \tag{4.12}$$

i suggereix que

$$\begin{aligned} M\left\{\frac{\phi(n)}{n}\right\} &= M\left\{\prod_p\left(1-\frac{\rho_p(n)}{p}\right)\right\} \\ &= \prod_p M\left\{\left(1-\frac{\rho_p(n)}{p}\right)\right\} \\ &= \prod_p\left(1-\frac{1}{p^2}\right) = \frac{1}{\zeta(2)} = \frac{6}{\pi^2}\,. \end{aligned} \tag{4.13}$$

Malauradament, (4.13) no es pot deduir directament de (4.12), perquè la densitat D no és completament additiva.

Nogensmenys, (4.13) es pot deduir fàcilment de la manera següent. A partir de (4.8) es dedueix que

$$\frac{\phi(n)}{n} = \sum_{d\,|\,n}\frac{\mu(d)}{d}\,, \tag{4.14}$$

on $\mu(d)$ és la funció de Möbius definida com a

1. $\mu(1) = 1$.

2. $\mu(m) = 0$ si m és divisible pel quadrat d'un primer.

3. $\mu(m) = (-1)^\nu$ si m és un producte de ν primers diferents.

Es dedueix ara que

$$\frac{1}{N}\sum_{n=1}^{N}\frac{\phi(n)}{n} = \frac{1}{N}\sum_{d=1}^{N}\frac{\mu(d)}{d}\left[\frac{N}{d}\right] \tag{4.15}$$

i, per tant, que

$$M\left\{\frac{\phi(n)}{n}\right\} = \sum_{d=1}^{\infty}\frac{\mu(d)}{d^2} = \prod_p\left(1-\frac{1}{p^2}\right) = \frac{1}{\zeta(2)} = \frac{6}{\pi^2}\,. \tag{4.16}$$

Posem ara

$$f_k(n) = \prod_{p\le p_k}\left(1-\frac{\rho_p(n)}{p}\right), \tag{4.17}$$

i considerem

$$f_k(n) - \frac{\phi(n)}{n}\,.$$

Tenim, clarament, que

$$0 \le f_k(n) - \frac{\phi(n)}{n} \le 1, \tag{4.18}$$

i, encara més, per (4.16) i (4.12)

$$M\left\{f_k(n) - \frac{\phi(n)}{n}\right\} = \prod_{p\le p_k}\left(1-\frac{1}{p^2}\right) - \prod_p\left(1-\frac{1}{p^2}\right). \tag{4.19}$$

Ara, per $l > 1$,

$$0 \le f_k^l(n) - \left(\frac{\phi(n)}{n}\right)^l \le l\left(f_k(n) - \frac{\phi(n)}{n}\right), \tag{4.20}$$

i, per tant,

$$\frac{1}{N}\sum_{n=1}^{N} f_k^l(n) \ge \frac{1}{N}\sum_{n=1}^{N}\left(\frac{\phi(n)}{n}\right)^l \ge \frac{1}{N}\sum_{n=1}^{N} f_k^l(n) - \frac{l}{N}\sum_{n=1}^{N}\left(f_k(n) - \frac{\phi(n)}{n}\right).$$

Per $N\to\infty$ obtenim

$$\begin{aligned} & M\{f_k^l(n)\} \\ \ge\ & \limsup_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\left(\frac{\phi(n)}{n}\right)^l \ge \liminf_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\left(\frac{\phi(n)}{n}\right)^l \\ \ge\ & M\{f_k^l(n)\} - lM\left\{f_k(n) - \frac{\phi(n)}{n}\right\}. \end{aligned} \tag{4.21}$$

Però,

$$\begin{aligned} & M\{f_k^l(n)\} \\ =\ & M\left\{\prod_{p\le p_k}\left(1-\frac{\rho_p(n)}{p}\right)^l\right\} = \prod_{p\le p_k} M\left\{\left(1-\frac{\rho_p(n)}{p}\right)^l\right\} \\ =\ & \prod_{p\le p_k}\left[1-\frac{1}{p}+\frac{1}{p}\left(1-\frac{1}{p}\right)^l\right], \end{aligned}$$

i, combinant això amb (4.19) i (4.21) obtenim, tot fent $k \to \infty$,

$$M\left\{\left(\frac{\phi(n)}{n}\right)^l\right\} = \prod_p \left[1 - \frac{1}{p} + \frac{1}{p}\left(1 - \frac{1}{p}\right)^l\right], \tag{4.22}$$

fórmula deguda a I. Schur.

Formalment, (4.22) es dedueix en una línia:

$$\begin{aligned} M\left\{\left(\frac{\phi(n)}{n}\right)^l\right\} &= M\left\{\prod_p \left(1 - \frac{\rho_p(n)}{p}\right)^l\right\} \\ &= \prod_p M\left\{\left(1 - \frac{\rho_p(n)}{p}\right)^l\right\} \\ &= \prod_p \left[1 - \frac{1}{p} + \frac{1}{p}\left(1 - \frac{1}{p}\right)^l\right], \end{aligned}$$

però com que D no és completament additiva, la justificació donada més amunt és necessària.

A partir de (4.10), tenim

$$\log \frac{\phi(n)}{n} = \sum_p \log\left(1 - \frac{\rho_p(n)}{p}\right) = \sum_p \rho_p(n) \log\left(1 - \frac{1}{p}\right) \tag{4.23}$$

i, formalment, per a tot real ξ,

$$\begin{aligned} & M\left\{\exp\left(i\xi \log \frac{\phi(n)}{n}\right)\right\} \qquad (4.24) \\ &= \prod_p M\left\{\exp\left(i\xi\rho_p(n)\log\left(1 - \frac{1}{p}\right)\right)\right\} \\ &= \prod_p \left(1 - \frac{1}{p} + \frac{1}{p}\exp\left(i\xi \log\left(1 - \frac{1}{p}\right)\right)\right) \\ &= c(\xi). \end{aligned}$$

La justificació rigorosa de (4.24) és gairebé idèntica a la que ja s'ha vist de (4.22) i la deixem al lector.

Sigui ara $K_N(\omega)$ el nombre d'enters n menors o iguals que N pels quals

$$\log \frac{\phi(n)}{n} < \omega.$$

Posem

$$\sigma_N(\omega) = \frac{K_N(\omega)}{N}, \tag{4.25}$$

i observem que $\sigma_N(\omega)$ és una *funció de distribució* i que

$$\int_{-\infty}^{\infty} e^{i\xi\omega}\, d\sigma_N(\omega) \tag{4.26}$$
$$= \frac{\exp\left(i\xi \log \dfrac{\phi(1)}{1}\right) + \cdots + \exp\left(i\xi \log \dfrac{\phi(N)}{N}\right)}{N}\,.$$

De (4.24) es dedueix que

$$\lim_{N\to\infty} \int_{-\infty}^{\infty} e^{i\xi\omega}\, d\sigma_N(\omega) = M\left\{\exp\left(i\xi \log \frac{\phi(n)}{n}\right)\right\} = c(\xi)\,, \tag{4.27}$$

i es veu fàcilment que $c(\xi)$ és contínua a $\xi = 0$. Així, pel teorema enunciat al final del § 3.4, existeix una funció de distribució $\sigma(\omega)$ tal que

$$\int_{-\infty}^{\infty} e^{i\xi\omega}\, d\sigma(\omega) = c(\xi) = \prod_p \left(1 - \frac{1}{p} + \frac{1}{p}\exp\left(i\xi \log\left(1 - \frac{1}{p}\right)\right)\right) \tag{4.28}$$

i que compleix que

$$\lim_{N\to\infty} \sigma_N(\omega) = \sigma(\omega) \tag{4.29}$$

en cada punt de continuïtat de $\sigma(\omega)$. Ara és fàcil demostrar que $\sigma(\omega)$ és contínua per a tot ω. Per a fer-ho, farem servir el resultat del problema 3 (pàgina 38, capítol 3).

Tenim

$$|\, c(\xi)\,|^2 \tag{4.30}$$
$$= \prod_p \left[\left(1 - \frac{1}{p}\right)^2 + \frac{2}{p}\left(1 - \frac{1}{p}\right)\cos\left(\xi \log\left(1 - \frac{1}{p}\right)\right) + \frac{1}{p^2}\right]$$
$$= \prod_{p \le p_k} \left[\left(1 - \frac{1}{p}\right)^2 + \frac{2}{p}\left(1 - \frac{1}{p}\right)\cos\left(\xi \log\left(1 - \frac{1}{p}\right)\right) + \frac{1}{p^2}\right],$$

i hom pot veure (vegeu el problema 1 darrere d'aquesta secció) que els nombres

$$\log\left(1 - \frac{1}{p}\right)$$

són linealment independents.

D'acord amb les consideracions del § 3.5, tenim que

$$\begin{aligned}
\lim_{T\to\infty} \frac{1}{T} \int_0^T \prod_{p\le p_k} &\left[\left(1-\frac{1}{p}\right)^2 + \frac{2}{p}\left(1-\frac{1}{p}\right)\right. \\
&\qquad \times \cos\left(\xi \log\left(1-\frac{1}{p}\right)\right) + \left.\frac{1}{p^2}\right] d\xi \\
= \prod_{p\le p_k} \lim_{T\to\infty} \frac{1}{T} \int_0^T &\left[\left(1-\frac{1}{p}\right)^2 + \frac{2}{p}\left(1-\frac{1}{p}\right)\right. \\
&\qquad \times \cos\left(\xi \log\left(1-\frac{1}{p}\right)\right) + \left.\frac{1}{p^2}\right] d\xi \\
= \prod_{p\le p_k} &\left[\left(1-\frac{1}{p}\right)^2 + \frac{1}{p^2}\right],
\end{aligned}$$

i per resultats elementals sobre nombres primers, sabem que

$$\lim_{k\to\infty} \prod_{p\le p_k} \left[\left(1-\frac{1}{p}\right)^2 + \frac{1}{p^2}\right] = \prod_p \left[\left(1-\frac{1}{p}\right)^2 + \frac{1}{p^2}\right] = 0.$$

En conseqüència, es dedueix que

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T |\, c(\xi)\,|^2 \, d\xi = 0, \tag{4.31}$$

i, per tant, $\sigma(\omega)$ és contínua per a tot ω. Resumint, la densitat

$$D\left\{\log \frac{\phi(n)}{n} < \omega\right\} = \sigma(\omega)$$

existeix per a tot ω, $\sigma(\omega)$ és contínua i

$$\int_{-\infty}^{\infty} e^{i\xi\omega}\, d\sigma(\omega) = \prod_p \left[\left(1-\frac{1}{p}\right) + \frac{1}{p}\exp\left(i\xi \log\left(1-\frac{1}{p}\right)\right)\right].$$

Aquest resultat (obtingut per primera vegada per I. Schoenberg), es pot deduir d'una manera més elemental i ha estat a bastament generalitzat per P. Erdös.* Hem escollit el camí més tortuós per tal d'extraure el peculiar aroma probabilístic del resultat i exhibir, al mateix temps, la interrelació d'una varietat d'idees i tècniques.

La fórmula (4.24) és un anàleg molt clar de la fórmula

$$\frac{\sin\xi}{\xi} = \prod_{k=1}^{\infty} \cos\frac{\xi}{2^k}$$

*Erdös també ha demostrat el notable teorema que la nostra $\sigma(\omega)$ és singular, és a dir, que $\sigma'(\omega) = 0$ gairebé pertot.

amb la qual hem començat. És, per a dir-ho d'alguna manera, una variació sobre el mateix tema, i el fet que el tema permeti tantes variacions diferents és un clar homenatge a la «melodia» que conté.

PROBLEMES

1. Demostreu que tant els nombres de la forma $\log(1-(1/p))$ com els de la forma $\log(1+(1/p))$ són linealment independents.
2. *Estadística de $\sigma(n)$ (suma dels divisors de n).*

   a) Sigui $\alpha_p(n)$ la potència amb la qual el primer p apareix a la representació (única) de n com a producte de potències del seus divisors primers, és a dir,
   $$n=\prod_p p^{\alpha_p(n)}\,.$$
   Demostreu que les funcions $\alpha_p(n)$ són estadísticament independents.

   b) Demostreu que si $\sigma(n)$ denota la suma de tots els divisors de n llavors,
   $$\frac{\sigma(n)}{n}=\prod_p\left(1+\frac{1}{p}+\cdots+\frac{1}{p^{\alpha_p(n)}}\right).$$

   c) Fent servir el fet que
   $$\frac{\sigma(n)}{n}=\sum_{k\,|\,n}\frac{1}{k}$$
   demostreu que
   $$M\left\{\frac{\sigma(n)}{n}\right\}=\frac{\pi^2}{6}\,.$$

   d) Demostreu que
   $$M\left\{\prod_{p\le p_k}\left(1+\frac{1}{p}+\cdots+\frac{1}{p^{\alpha_p(n)}}\right)\right\}=\prod_{p\le p_k}\frac{1}{1-\dfrac{1}{p^2}}\,.$$

   e) Poseu
   $$f_k(n)=\prod_{p\le p_k}\left(1+\frac{1}{p}+\cdots+\frac{1}{p^{\alpha_p(n)}}\right);$$
   observeu que
   $$\frac{f_k(n)}{\dfrac{\sigma(n)}{n}}=\prod_{p>p_k}\frac{1}{1+\dfrac{1}{p}+\cdots+\dfrac{1}{p^{\alpha_p(n)}}}\,,$$

i deduïu d'aquí la desigualtat

$$\prod_{p>p_k}\left(1-\frac{\rho_p(n)}{p}\right) \le \frac{f_k(n)}{\dfrac{\sigma(n)}{n}} \le \prod_{p>p_k}\frac{1}{1+\dfrac{\rho_p(n)}{p}}\,.$$

f) Demostreu que

$$\begin{aligned}
&M\left\{e^{i\xi\log\frac{\sigma(n)}{n}}\right\}\\
&\quad=\prod_p M\left\{\exp\left[i\xi\log\left(1+\frac{1}{p}+\cdots+\frac{1}{p^{\alpha_p(n)}}\right)\right]\right\}\\
&\quad=\prod_p\left[1-\frac{1}{p}+\sum_{\alpha=1}^{\infty}\left(\frac{1}{p^\alpha}-\frac{1}{p^{\alpha+1}}\right)\right.\\
&\qquad\qquad\left.\times\exp\left[i\xi\log\left(1+\frac{1}{p}+\cdots+\frac{1}{p^{\alpha_p(n)}}\right)\right]\right]\\
&\quad=c(\xi)\,.
\end{aligned}$$

g) Fent servir el fet que

$$\begin{aligned}
&\left|1-\frac{1}{p}+\sum_{\alpha=1}^{\infty}\left(\frac{1}{p^\alpha}-\frac{1}{p^{\alpha+1}}\right)\exp\left[i\xi\log\left(1+\frac{1}{p}+\cdots+\frac{1}{p^{\alpha_p(n)}}\right)\right]\right|\\
&\qquad\le\left|1-\frac{1}{p}+\frac{1}{p}\left(1-\frac{1}{p}\right)\exp\left[i\xi\log\left(1+\frac{1}{p}\right)\right]\right|+\frac{1}{p^2}\\
&\qquad=\left(1-\frac{1}{p}\right)\sqrt{1+\frac{2}{p}\cos\left[\xi\log\left(1+\frac{1}{p}\right)\right]+\frac{1}{p^2}}+\frac{1}{p^2}
\end{aligned}$$

així com el fet que els nombres $\log(1+(1/p))$ són linealment independents, demostreu que la densitat

$$D\left\{\frac{\sigma(n)}{n}<\omega\right\}=\tau(\omega)$$

existeix i és una funció contínua de ω.

Aquest resultat va ser demostrat per primera vegada per H. Davenport i queda englobat en un teorema molt més general d'Erdös.

El cas $\omega = 2$ és particularment interessant perquè mostra que els *nombres abundants* (és a dir, nombres pels quals $\sigma(n) > 2n$), així com els *nombres deficients* (és a dir, nombres pels quals $\sigma(n) < 2n$), tenen densitat. També es dedueix que els *nombres perfectes* (pels quals $\sigma(n) = 2n$) tenen densitat 0. Hom conjectura que només existeixen un nombre finit de nombres perfectes.

3. *La fórmula d'inversió.* Demostreu que si

$$\int_{-\infty}^{\infty} e^{i\xi\omega}\, d\sigma(\omega) = c(\xi)\,,$$

on $\sigma(\omega)$ és una funció de distribució, llavors

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{e^{i\omega_2\xi} - e^{i\omega_1\xi}}{i\xi}\, c(\xi)\, d\xi = \sigma(\omega_2) - \sigma(\omega_1)$$

sempre que ω_1 i ω_2 siguin punts de continuïtat de σ. Si ω_1 o bé ω_2 (o tots dos alhora) són punts de discontinuïtat, llavors $\sigma(\omega_1)$ o $\sigma(\omega_2)$ (o tots dos alhora) s'han de substituir per

$$\frac{\sigma(\omega_1 - 0) + \sigma(\omega_1 + 0)}{2} \quad \text{o bé} \quad \frac{\sigma(\omega_2 - 0) + \sigma(\omega_2 + 0)}{2}\,.$$

En particular, demostreu que

$$D\left\{\omega_1 < \log\frac{\phi(n)}{n} < \omega_2\right\}$$
$$= \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{e^{i\omega_2\xi} - e^{i\omega_1\xi}}{i\xi} \prod_p \left(1 - \frac{1}{p} + \frac{1}{p}\exp\left[i\xi\log\left(1 - \frac{1}{p}\right)\right]\right) d\xi,$$

una fórmula explícita però gairebé inútil!

4.3. Una altra aplicació

Sigui $\omega(n)$ el nombre de divisors primers de n comptant la seva multiplicitat, és a dir,

$$\omega(n) = \sum_p \alpha_p(n), \tag{4.32}$$

on les α s'han definit en el problema 2, § 4.2.

Sigui $\nu(n)$ el nombre de divisors primers de n sense comptar la seva multiplicitat, és a dir,

$$\nu(n) = \sum_p \rho_p(n). \tag{4.33}$$

La diferència $\omega(n)-\nu(n)$ s'anomena *excés*. Volem ara determinar la densitat del conjunt d'enters per als quals l'excés és igual a k ($k \geq 0$ un enter), és a dir,

$$d_k = D\{\omega(n) - \nu(n) = k\}. \tag{4.34}$$

No cal dir que l'existència d'aquesta densitat no és evident i cal trobar-la.

Comencem amb la fórmula (1.14) del capítol 1,

$$\frac{1}{2\pi}\int_0^{2\pi} e^{imx}\, dx = \begin{cases} 1, & m = 0 \\ 0, & m \neq 0, \end{cases} \tag{4.35}$$

on m és un enter. Considerem també

$$\frac{1}{N}\sum_{n=1}^{N}\frac{1}{2\pi}\int_0^{2\pi} e^{i(\omega(n)-\nu(n)-k)x}\,dx = \frac{1}{2\pi}\int_0^{2\pi} e^{-ikx}\frac{1}{N}\sum_{n=1}^{N} e^{i(\omega(n)-\nu(n))x}\,dx. \tag{4.36}$$

El membre esquerre de (4.36) representa [a la vista de (4.35)] el nombre d'enters $n \leq N$ que tenen excés exactament igual a k. Així,

$$d_k = \lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\frac{1}{2\pi}\int_0^{2\pi} e^{i(\omega(n)-\nu(n)-k)x}\,dx \tag{4.37}$$

quan el límit existeix. Invocant novament el principi de la convergència dominada, veiem a partir de (4.36) que n'hi ha prou amb demostrar que, per a tot real x, el límit

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N} e^{i(\omega(n)-\nu(n))x} = M\{e^{i(\omega(n)-\nu(n))x}\} \tag{4.38}$$

existeix. Ara bé,

$$\omega(n)-\nu(n) = \sum_p(\alpha_p(n)-\rho_p(n)),$$

i es veu fàcilment que les funcions $\alpha_p(n)-\rho_p(n)$ són independents. Això suggereix que el límit (4.38) no només existeix, sinó que també

$$\begin{aligned} &M\{e^{i(\omega(n)-\nu(n))x}\} \qquad (4.39)\\ &= M\left\{\exp\left[ix\sum_p(\alpha_p(n)-\rho_p(n))\right]\right\}\\ &= \prod_p M\{e^{ix(\alpha_p(n)-\rho_p(n))}\}\\ &= \prod_p\left[1-\frac{1}{p}+\sum_{\alpha=1}^{\infty}\left(\frac{1}{p^\alpha}-\frac{1}{p^{\alpha+1}}\right)e^{ix(\alpha-1)}\right]\\ &= \prod_p\left(1-\frac{1}{p}\right)\left(1+\frac{1}{p-e^{ix}}\right). \end{aligned}$$

És fàcil donar una justificació rigorosa d'aquest resultat, tot seguint línies similars a les § 4.2. Preneu primer

$$\sum_{n=1}^{N}(\alpha_p(n)-\rho_p(n)),$$

i considereu els enters n, $1 \leq n \leq N$, pels quals

$$\alpha_p(n) = \beta.$$

Aquests són exactament els enters divisibles per p^β, però no divisibles per $p^{\beta+1}$ i, per tant, el seu nombre és igual a

$$\left[\frac{N}{p^\beta}\right] - \left[\frac{N}{p^{\beta+1}}\right].$$

Es dedueix, doncs, que

$$\sum_{n=1}^{N}(\alpha_p(n) - \rho_p(n)) = \sum_{\beta\geq 2}(\beta-1)\left\{\left[\frac{N}{p^\beta}\right] - \left[\frac{N}{p^{\beta+1}}\right]\right\}. \tag{4.40}$$

Sigui ara

$$g_k(n) = \sum_{p>p_k}(\alpha_p(n) - \rho_p(n)), \tag{4.41}$$

i observem que (4.40) implica que

$$\begin{aligned} M\{g_k(n)\} &= \sum_{p>p_k}\sum_{\beta\geq 2}(\beta-1)\left(\frac{1}{p^\beta} - \frac{1}{p^{\beta+1}}\right) \\ &< \sum_{p>p_k}\sum_{\beta\geq 2}(\beta-1)\frac{1}{p^\beta} = \sum_{p>p_k}\frac{1}{(p-1)^2}. \end{aligned} \tag{4.42}$$

Ara,

$$\begin{aligned} &\frac{1}{N}\sum_{n=1}^{N} e^{ix(\omega(n)-\nu(n))} \\ &\quad = \frac{1}{N}\sum_{n=1}^{N}\exp\left[ix\sum_{p\leq p_k}(\alpha_p(n) - \rho_p(n))\right] e^{ixg_k(n)}, \end{aligned} \tag{4.43}$$

i, per tant,

$$\begin{aligned} &\left|\frac{1}{N}\sum_{n=1}^{N} e^{ix(\omega(n)-\nu(n))} - \frac{1}{N}\sum_{n=1}^{N}\exp\left[ix\sum_{p\leq p_k}(\alpha_p(n) - \rho_p(n))\right]\right| \\ &\quad = \left|\frac{1}{N}\sum_{n=1}^{N}\exp\left[ix\sum_{p\leq p_k}(\alpha_p(n) - \rho_p(n))\right](e^{ixg_k(n)} - 1)\right| \\ &\quad \leq \frac{1}{N}\sum_{n=1}^{N}\left|e^{ixg_k(n)} - 1\right| \leq \frac{|x|}{N}\sum_{n=1}^{N} g_k(n). \end{aligned}$$

Atès que

$$\begin{aligned}
&\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\exp\left[ix\sum_{p\le p_k}(\alpha_p(n)-\rho_p(n))\right]\\
&\quad = M\left\{\exp\left[ix\sum_{p\le p_k}(\alpha_p(n)-\rho_p(n))\right]\right\}\\
&\quad = \prod_{p\le p_k} M\{e^{ix(\alpha_p(n)-\rho_p(n))}\} = \prod_{p\le p_k}\left(1-\frac{1}{p}\right)\left(1+\frac{1}{p-e^{ix}}\right),
\end{aligned}$$

veiem que (4.42) implica que la distància desde qualsevol *punt d'acumulació* de la successsió

$$\frac{1}{N}\sum_{n=1}^{N} e^{ix(\omega(n)-\nu(n))}$$

a

$$\prod_{p\le p_k}\left(1-\frac{1}{p}\right)\left(1+\frac{1}{p-e^{ix}}\right)$$

és menor que

$$|\,x\,|\sum_{p>p_k}\frac{1}{(p-1)^2}\,.$$

Atès que k és arbitrari, es dedueix immediatament que

$$\begin{aligned}
&\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N} e^{ix(\omega(n)-\nu(n))}\\
&\quad = M\{e^{ix(\omega(n)-\nu(n))}\}\\
&\quad = \prod_{p}\left(1-\frac{1}{p}\right)\left(1+\frac{1}{p-e^{ix}}\right),
\end{aligned}\tag{4.44}$$

i (4.39) queda, doncs, justificat.

Tornant a (4.36) i (4.37) obtenim

$$\begin{aligned}
d_k &= D\{\omega(n)-\nu(n)=k\}\\
&= \frac{1}{2\pi}\int_0^{2\pi} e^{ikx}\prod_p\left(1-\frac{1}{p}\right)\left(1+\frac{1}{p-e^{ix}}\right)dx\,.
\end{aligned}\tag{4.45}$$

Considerem ara la funció

$$F(z)=\prod_p\left(1-\frac{1}{p}\right)\left(1+\frac{1}{p-z}\right),\tag{4.46}$$

i observem que és analítica en el pla sencer, excepte en els pols simples $z = 2, 3, 4, \cdots$. En particular, $F(z)$ és analítica en el cercle $|\, z\,| < 2$ i la podem desenvolupar en sèrie de potències

$$F(z) = \sum_{k=0}^{\infty} a_k z^k \,,$$

amb radi de convergència 2.

Quins són els coeficients a_k? Si fem servir la coneguda fórmula

$$a_k = \frac{1}{2\pi i} \int \frac{F(z)}{z^{k+1}}\, dz \,,$$

on la integral es pren al llarg del cercle $|\, z\,| = 1$, obtenim, tot substituint $z = e^{ix}$, que

$$a_k = d_k.$$

En altres paraules,

$$\sum_{k=0}^{\infty} d_k z^k = \prod_p \left(1 - \frac{1}{p}\right)\left(1 + \frac{1}{p-z}\right) . \tag{4.47}$$

Aquesta fórmula tan bonica va ser descoberta per A. Rényi d'una manera diferent.

Malgrat la pesadesa d'obtenir fórmules explícites per d_k, és força senzill determinar-ne el comportament asimptòtic per a grans valors de k.

De fet, $F(z)$ es pot escriure de la manera següent:

$$F(z) = \frac{A}{z-2} + G(z) \,,$$

on $G(z)$ és analítica en el cercle $|\, z\,| < 3$ i A (el residu en el pol 2) és donat per la fórmula

$$A = -\frac{1}{2} \prod_{p>2} \left(1 - \frac{1}{p}\right)\left(1 + \frac{1}{p-2}\right) .$$

Així

$$F(z) = \frac{1}{4} \prod_{p>2} \left(1 - \frac{1}{p}\right)\left(1 + \frac{1}{p-2}\right) \sum_{k=0}^{\infty} \frac{z^k}{2^k} + \sum_{k=0}^{\infty} b_k z^k \,,$$

on el radi de convergència de $\sum b_k z^k$ és 3. Atès que

$$d_k = \frac{1}{4} \prod_{p>2} \left(1 - \frac{1}{p}\right)\left(1 + \frac{1}{p-2}\right) \frac{1}{2^k} + b_k$$

i que

$$\limsup_{k\to\infty} |\, b_k\,|^{1/k} = \frac{1}{3} \,,$$

tenim, per $k \to \infty$,

$$d_k \sim \frac{1}{2^{k+2}} \prod_{p>2} \left(1 - \frac{1}{p}\right) \left(1 + \frac{1}{p-2}\right) \tag{4.48}$$

o bé

$$\lim_{k\to\infty} 2^{k+2} d_k = \prod_{p>2} \left(1 - \frac{1}{p}\right) \left(1 + \frac{1}{p-2}\right) . \tag{4.49}$$

Hi ha dos casos concrets de (4.47) que mereixen la nostra atenció. Posant-hi $z = 0$ obtenim

$$d_0 = \prod_p \left(1 - \frac{1}{p^2}\right) = \frac{1}{\zeta(2)} = \frac{6}{\pi^2} .$$

Aquest és el resultat, prou ben conegut, que diu que la densitat dels nombres *lliures de quadrats* (és a dir, que no són divisibles per un quadrat perfecte) és $6/\pi^2$.

Posant-hi $z = 1$, obtenim

$$\sum_{k=0}^{\infty} d_k = \prod_p \left(1 - \frac{1}{p}\right) \left(1 + \frac{1}{p-1}\right) = 1 .$$

Atès que els conjunts d'enters pels quals $\omega(n) - \nu(n) = k$ són disjunts i la reunió de tots constitueix el conjunt de tots els enters, el resultat seria totalment trivial si la densitat fos completament additiva. Com que no ho és, el fet que de totes les maneres obtenim

$$\sum_{k=0}^{\infty} d_k = 1$$

és, com a mínim, una mica divertit.

4.4. Gairebé tots els enters m tenen, aproximadament, $\log \log m$ divisors primers

Considerem els enters $1 \le m \le n$ pels quals, o bé

$$\nu(m) < \log \log n - g_n \sqrt{\log \log n} \tag{4.50}$$

o bé

$$\nu(m) > \log \log n + g_n \sqrt{\log \log n} ,$$

on g_n és una successió que tendeix a infinit:

$$\lim_{n\to\infty} g_n = \infty . \tag{4.51}$$

Sigui K_n el seu nombre i intentem fer-ne una estimació. Farem servir l'esquema de Tchebyscheff que hem explicat al § 2.1.

Tenim

$$\sum_{m=1}^{n}(\nu(m)-\log\log n)^2 \geq {\sum}'(\nu(m)-\log\log n)^2\,, \tag{4.52}$$

on la prima del sumatori indica que la suma s'estén només als enters m que satisfan (4.50). Clarament

$${\sum}'(\nu(m)-\log\log n)^2 \geq K_n\, g_n^2 \log\log n\,, \tag{4.53}$$

i, d'aquí, per (4.52)

$$\frac{K_n}{n} \leq \frac{1}{n\, g_n^2 \log\log n}\sum_{m=1}^{n}(\nu(m)-\log\log n)^2\,. \tag{4.54}$$

Ens cal, doncs, trobar una estimació de

$$\begin{aligned}&\sum_{m=1}^{n}(\nu(m)-\log\log n)^2 \qquad\qquad (4.55)\\ &\quad=\sum_{m=1}^{n}\nu^2(m)-2\log\log n\sum_{m=1}^{n}\nu(m)+n(\log\log n)^2\,.\end{aligned}$$

Ara bé,

$$\nu(m)=\sum_p \rho_p(m)\,,$$

i

$$\nu^2(m)=\sum_p \rho_p(m)+2\sum_{p<q}\rho_p(m)\rho_q(m)$$

$(\rho_p^2=\rho_p)$; en conseqüència,

$$\sum_{m=1}^{n}\nu(m)=\sum_p\left[\frac{n}{p}\right]\,, \tag{4.56}$$

i

$$\sum_{m=1}^{n}\nu^2(m)=\sum_p\left[\frac{n}{p}\right]+2\sum_{p<q}\left[\frac{n}{pq}\right]\,. \tag{4.57}$$

A (4.56) i (4.57) el sumatori s'estén només sobre els primers p i q menors o iguals que n. Per tant,

$$\sum_{m=1}^{n}\nu(m)\geq n\sum_{p\leq n}\frac{1}{p}-\pi(n)\,, \tag{4.58}$$

on $\pi(n)$ denota el nombre de primers menors o iguals a n. De manera semblant

$$\sum_{m=1}^{n} \nu^2(m) \leq n \sum_{p\leq n} \frac{1}{p} + 2n \sum_{p<q\leq n} \frac{1}{pq} < n \sum_{p\leq n} \frac{1}{p} + n \left(\sum_{p\leq n} \frac{1}{p} \right)^2 . \tag{4.59}$$

És conegut que

$$\sum_{p\leq n} \frac{1}{p} = \log\, \log\, n + e_n\,, \tag{4.60}$$

on e_n és fitada i, per tant,

$$\sum_{m=1}^{n} \nu^2(m) \leq n(\log\, \log\, n)^2 + 2n \log\, \log\, ne_n + ne_n^2 + n \log\, \log\, n + ne_n$$

i

$$\sum_{m=1}^{n} \nu(m) \geq n \log\, \log\, n + ne_n - \pi(n)\,.$$

Finalment, (4.55) proporciona

$$\sum_{m=1}^{n} (\nu(m) - \log\, \log\, n)^2 \leq ne_n^2 + n \log\, \log\, n + ne_n + 2 \log\, \log\, n\pi(n)\,,$$

i, en conseqüència,

$$\frac{K_n}{n} \leq \frac{1}{g_n^2} + \frac{e_n^2}{g_n^2 \log\, \log\, n} + \frac{e_n}{g_n^2 \log\, \log\, n} + 2\, \frac{\pi(n)}{n} \frac{1}{g_n^2}\,.$$

Atès que e_n és fitada, $\pi(n) < n$ i $g_n \to \infty$, es dedueix que

$$\lim_{n\to\infty} \frac{K_n}{n} = 0. \tag{4.61}$$

A causa de la lentitud amb la qual $\log \log m$ canvia, (4.61) implica el següent.

Si l_n denota el nombre d'enters, $1 \leq m \leq n$, pels quals es compleix o bé

$$\nu(m) < \log\, \log\, m - g_m \sqrt{\log\, \log\, m} \tag{4.62}$$

o bé

$$\nu(m) > \log\, \log\, m + g_m \sqrt{\log\, \log\, m}\,,$$

llavors

$$\lim_{n\to\infty} \frac{l_n}{n} = 0. \tag{4.63}$$

La demostració es deixa al lector (vegeu el problema 1 al final d'aquesta secció). El teorema de l'equació (4.63) va ser demostrat per primera vegada per Hardy i Ramanujan el 1917. A ells es deu l'enunciat pintoresc que gairebé tots els enters tenen, aproximadament, log log m divisors primers. La demostració que hem donat és de P. Turán, i és molt més senzilla que la demostració original de Hardy i Ramanujan. El lector ha vist que la demostració de Turán és un anàleg directe de la demostració de la llei feble dels grans nombres que hem vist en el § 2.1. Aquesta és una altra mostra d'una idea prestada d'un camp que dóna fruit quan és aplicada en un altre.

PROBLEMES

1. Demostreu (4.63). (*Indicació*: sigui $0 < \alpha < 1$; considereu només els enters en el rang $n^\alpha \leq m \leq n$, i demostreu que tot enter m d'aquest rang que satisfà
$$|\,\nu(m) - \log\log m\,| > g_m\sqrt{\log\log m}$$
també satisfà
$$|\,\nu(m) - \log\log n\,| > h_n\sqrt{\log\log n}\,,$$
amb $h_n \to \infty$ adequadament escollit.)
2. Demostreu (4.61) per a $\omega(m)$.

4.5. La llei normal en teoria de nombres.

El fet que $\nu(m)$, el nombre de divisors primers de m, sigui igual a la suma
$$\sum_p \rho_p(m) \tag{4.64}$$
de funcions independents suggereix que, d'alguna manera, la distribució dels valors de $\nu(m)$ pot venir donada per la llei normal. Aquest és realment el cas, i el 1939 Erdös i Kac van demostrar el teorema següent.

Sigui $K_n(\omega_1, \omega_2)$ el nombre d'enters m, $1 \leq m \leq n$, pels quals
$$\log\log n + \omega_1\sqrt{\log\log n} < \nu(m) < \log\log n + \omega_2\sqrt{\log\log n}\,; \tag{4.65}$$
llavors,
$$\lim_{n\to\infty} \frac{K_n(\omega_1,\omega_2)}{n} = \frac{1}{\sqrt{2}} \int_{\omega_1}^{\omega_2} e^{-y^2/2}\,dy\,. \tag{4.66}$$
A causa de la lentitud amb la qual log log n canvia (vegeu el problema 1 al final del § 4.4) el resultat (4.66) és equivalent a:
$$D\left\{\log\log n + \omega_1\sqrt{\log\log n} < \nu(n) < \log\log n + \omega_2\sqrt{\log\log n}\right\} = \frac{1}{\sqrt{2\pi}} \int_{\omega_1}^{\omega_2} e^{-y^2/2}\,dy\,. \tag{4.67}$$

Existeixen actualment diverses demostracions d'aquest resultat (la millor, en la meva opinió, és una demostració recent de Rényi i Turán), però malauradament, cap és suficientment curta o elemental per a ser reproduïda aquí. En conseqüència, ens hem de conformar amb un argument heurístic basat en el resultat següent, ja clàssic, de Landau.

Si $\pi_k(n)$ denota el nombre d'enters menors o iguals que n que tenen exactament k divisors primers, llavors

$$\pi_k(n) \sim \frac{1}{(k-1)!}\,\frac{n}{\log n}\,(\log\,\log\,n)^{k-1}\,. \tag{4.68}$$

Per $k = 1$, obtenim el conegut teorema dels nombres primers; per $k > 1$, (4.68) es pot deduir del teorema dels nombres primers mitjançant consideracions totalment elementals.

Ara,

$$K_n(\omega_1, \omega_2) = \sum_{\log\,\log\,n+\omega_1\sqrt{\log\,\log\,n}<k<\log\,\log\,n+\omega_2\sqrt{\log\,\log\,n}} \pi_k(n)\,, \tag{4.69}$$

i, per tant, hom pot esperar que

$$\begin{aligned} &\frac{K_n(\omega_1, \omega_2)}{n} \\ &\sim \frac{1}{\log\,n} \sum_{\log\,\log\,n+\omega_1\sqrt{\log\,\log\,n}<k<\log\,\log\,n+\omega_2\sqrt{\log\,\log\,n}} \frac{(\log\,\log\,n)^{k-1}}{(k-1)!}\,. \end{aligned} \tag{4.70}$$

Si hom recorda el problema 2, § 3.3 i escriu

$$x = \log\,\log\,n \qquad \left(\text{o sigui } e^{-x} = \frac{1}{\log\,n}\right)\,, \tag{4.71}$$

hom obté,

$$\frac{K_n(\omega_1, \omega_2)}{n} \sim \frac{1}{\sqrt{2\pi}} \int_{\omega_1}^{\omega_2} e^{-y^2/2}\,dy$$

que és precisament (4.66).

Malauradament, fer rigorós aquest argument tan atractiu no és fàcil perquè fa falta una estimació uniforme de l'error en el teorema de Landau (4.68) que no és fàcil d'obtenir. Podria ser interessant esmentar que la demostració original de Hardy i Ramanujan del teorema del § 4.4 es basava essencialment en (4.68), malgrat que només buscaven certes estimacions més aviat que un resultat asimptòtic precís. La teoria que hem desenvolupat en el capítol 3 suggereix un mètode per a demostrar (4.66). Sigui $K_n(\omega)$ el nombre d'enters m, $1 \le m \le n$, pels quals

$$\nu(m) < \log\,\log\,n + \omega\sqrt{\log\,\log\,n}\,,$$

i posem

$$\sigma_n(\omega) = \frac{K_n(\omega)}{n}\,. \tag{4.72}$$

És clar que $\sigma_n(\omega)$ és una funció de distribució i que

$$\frac{1}{n\log\log n}\sum_{m=1}^{n}(\nu(m)-\log\log n)^2 = \int_{-\infty}^{\infty}\omega^2\,d\sigma_n(\omega)\,. \tag{4.73}$$

Si fem servir l'estimació precisa

$$\sum_{p\le n}\frac{1}{p} = \log\log n + C + \epsilon_n, \qquad \epsilon_n \to 0\,, \tag{4.74}$$

llavors l'argument del § 4.4 ens proporciona

$$\lim_{n\to\infty}\int_{-\infty}^{\infty}\omega^2\,d\sigma_n(\omega) = 1 = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}y^2e^{-y^2/2}\,dy\,. \tag{4.75}$$

També tenim (gairebé de manera trivial!) que

$$\lim_{n\to\infty}\frac{1}{n\sqrt{\log\log n}}\sum_{m=1}^{n}(\nu(m)-\log\log n) = 0\,,$$

i, per tant,

$$\lim_{n\to\infty}\int_{-\infty}^{\infty}\omega\,d\sigma_n(\omega) = 0 = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}ye^{-y^2/2}\,dy\,. \tag{4.76}$$

Si poguéssim demostrar que per a tot enter $k > 2$

$$\lim_{n\to\infty}\int_{-\infty}^{\infty}\omega^k\,d\sigma_n(\omega) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}y^ke^{-y^2/2}\,dy\,, \tag{4.77}$$

en deduiríem que

$$\lim_{n\to\infty}\int_{-\infty}^{\infty}e^{i\xi\omega}\,d\sigma_n(\omega) = e^{-\xi^2/2}$$

per a tot real ξ i, per tant,

$$\lim_{n\to\infty}\sigma_n(\omega) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\omega}e^{-y^2/2}\,dy\,. \tag{4.78}$$

Això, a la vista de (4.72), no és altra cosa que el nostre teorema (4.66). Demostrar (4.77) és, òbviament, equivalent a

$$\lim_{n\to\infty}\frac{1}{n(\log\log n)^{k/2}}\sum_{m=1}^{n}(\nu(m)-\log\log n)^k = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}y^ke^{-y^2/2}\,dy\,, \tag{4.79}$$

i això, al seu torn, depèn de l'avaluació asimptòtica de les sumes

$$\sum_{p_{l_1}\cdots p_{l_k}<n} \frac{1}{p_{l_1}\cdots p_{l_k}}\,.$$

(Recordem que en el § 4.4 la demostració de Turán depenia de l'estimació de

$$\sum_{pq\leq n} \frac{1}{pq}\,.)$$

Això, per estrany que sembli, no és gens fàcil. Halberstam, però, recentment ha aconseguit fer la demostració seguint aquestes línies. Aquest apropament, sens dubte, és el més directe i el que s'ajusta més, en essència, a les línies tradicionals de la teoria de la probabilitat. El triomf definitiu del mètode probabilista en la teoria de nombres va arribar amb la demostració de Rényi i Turán del resultat que deia que el terme d'error a

$$\frac{K_n(\omega)}{n} - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-y^2/2}\, dy$$

és de l'ordre de

$$\frac{1}{\sqrt{\log\, \log\, n}}\,.$$

Le Vecque, per analogia amb estimacions similars en teoria de probabilitats, va conjecturar que l'error és de l'ordre de $(\log\,\log\, n)^{-1/2}$.

Els nombres primers, en veritat, juguen a un joc d'atzar!

PROBLEMES

1. Demostreu que (4.67) val si $\nu(n)$ es substitueix per $\omega(n)$ (és a dir, el nombre de divisors primers comptant la multiplicitat). (*Indicació*: del fet que $M\{\omega(n) - \nu(n)\} < \infty$, deduïu, en primer lloc, que el conjunt d'enters pels quals $\omega(n) - \nu(n) > g_n, g_n \to \infty$, té densitat 0.)

2. Sigui $d(n)$ el nombre de divisors de n.

   (a) Demostreu que
   $$d(n) = \prod_p (\alpha_p(n) + 1)\,.$$
   (Per a la definició de $\alpha_p(n)$ vegeu el problema 2, § 4.3.)

   (b) Demostreu que
   $$M\left\{\frac{d(n)}{2^{\nu(n)}}\right\} = \prod_p \left(1 + \frac{1}{2p(p-1)}\right) < \infty.$$

(c) Fent servir (4.67) i la indicació del problema 1, demostreu que

$$D\{2^{\log\,\log\,n+\omega_1\sqrt{\log\,\log\,n}} < d(n) < 2^{\log\,\log\,n+\omega_2\sqrt{\log\,\log\,n}}\} \\ = \frac{1}{2\pi}\int_{\omega_1}^{\omega_2} e^{-y^2/2}\,dy\,.$$

BIBLIOGRAFIA

Per a les referències al treball de Davenport, Erdös, Erdös i Kac, Halsberstam i Schoenberg i Turán vegeu els articles: M. Kac, «Probability methods in some problems of analysis and number theory», *Bull. Amer. Math. Soc.*, **55** (1949), 641–665 i I. P. Kubilus, «Probability methods in number theory,» (en rus), *Usp. Mat. Nauk*, **68** (1956), 31–66.

A. Rényi, «On the density of certain sequences of integers», *Publ. Inst. Math. Belgrade*, **8** (1955), 157–162.

A. Rényi and P. Turán, «On a theorem of Erdös-Kac», *Acta Arith.*, **4** (1958), 71–84.

Capítol 5

De la teoria cinètica a les fraccions continuades

5.1. Paradoxes de la teoria cinètica

A mitjan segle XIX es van començar a fer intents per tal d'unificar les disciplines de la mecànica i la termodinàmica.

El principal problema consistia a deduir la segona llei de la termodinàmica a partir del model que la matèria consistia en partícules (àtoms o molècules) subjectes a forces i que obeïen les lleis de la mecànica.

En les mans de Maxwell i Boltzmann (i més tard de J. W. Gibbs) aquest apropament cinètic va donar com a fruit un dels resultats més bonics i penetrants de la ciència.

Però l'apropament es veia enterbolit, des del començament, per dues paradoxes. La primera, esbombada el 1876 per Loschmidt, consistia en l'observació que les lleis de la mecànica són reversibles en el temps (és a dir, invariants pel canvi de t per $-t$).

D'altra banda, la segona llei de la termodinàmica postula un comportament típicament irreversible.

Semblava, doncs, impossible, deduir la segona llei a partir merament de consideracions mecàniques.

La segona paradoxa, associada amb el nom de Zermelo, va ser encara més impactant.

Zermelo invocava un teorema senzill, però fonamental, de Poincaré que diu que un sistema dinàmic conservatiu, en unes certes condicions no gaire exigents, té la propietat que «gairebé tot» (en un sentit tècnic que aclarirem més endavant) estat del sistema es veu obligat a repetir-se fins al grau de precisió que es desitgi.

Això també està en contradicció amb un comportament irreversible.

Per tal d'entendre aquestes paradoxes, considerem dos contenidors, un ple de gas i l'altre completament buit.

En un moment donat, connectem els dos contenidors. Llavors, la segona llei prediu que el gas fluirà del primer al segon i que la quantitat de gas en el primer contenidor *disminuirà monòtonament amb el temps*. Aquest comportament del gas mostra una clara *fletxa del temps*.

Des del punt de vista cinètic (mecànic), estem davant d'un sistema dinàmic que no pot de cap manera real mostrar la fletxa del temps i que, a més a més, es comportarà d'una manera quasi-periòdica, tal com prediu el teorema de Poincaré.

Estem, clarament, davant d'una situació paradoxal.

5.2. Preliminars

Per tal d'entendre la resposta de Boltzmann cal que recordem alguns conceptes de dinàmica clàssica.

Un sistema amb n graus de llibertat queda descrit en termes de n coordenades generalitzades $q_1, q_2, \cdots, q_n$ i n moments conjugats $p_1, p_2, \cdots, p_n$. En un sistema dinàmic conservatiu existeix una funció del sistema,

$$H(q_1, \cdots, q_n; p_1, \cdots, p_n)$$

coneguda com la funció hamiltoniana, que representa la seva energia total.

Les equacions del moviment són de la forma

$$\frac{dq_i}{dt} = \frac{\partial H}{\partial p_i}, \quad i = 1, 2, \cdots, n, \tag{5.1}$$

$$\frac{dp_i}{dt} = -\frac{\partial H}{\partial q_i}, \quad i = 1, 2, \cdots, n, \tag{5.2}$$

i, si coneixem les posicions inicials $q_i(0)$ i els moments inicials $p_i(0)$, el moviment [és a dir, les funcions $q_i(t)$ i $p_i(t)$] queda determinat de manera única.

És habitual representar el sistema com un punt en l'espai euclidià $2n$-dimensional (espai de fases o Γ-espai) amb coordenades

$$q_1, \cdots, q_n, p_1, \cdots, p_n.$$

Així, en el moment t del temps, el sistema dinàmic està representat pel punt

$$P_t = (q_1(t), \cdots, q_n(t), p_1(t), \cdots, p_n(t)).$$

El moviment del nostre sistema defineix una família uniparamètrica de transformacions T_t per mitjà de la relació

$$T_t(P_0) = P_t. \tag{5.3}$$

Suposem ara que tenim un conjunt A de punts P_0 i anomenem $T_t(A)$ el conjunt corresponent de punts P_t.

Liouville va fer notar que les equacions hamiltonianes del moviment, (5.1) i (5.2), impliquen el fet remarcable que les mesures $2n$-dimensionals de A i $T_t(A)$ són iguals! (la demostració és molt senzilla i es pot basar en la generalització del conegut i familiar teorema de la divergència en l'espai $2n$-dimensional).

En altres paraules, les transformacions T_t preserven la mesura ordinària de Lebesgue en el Γ-espai.

Les equacions (5.1) i (5.2) tenen una altra conseqüència important, a saber, que

$$H(q_1(t), \cdots, q_n(t), p_1(t), \cdots, p_n(t)) \\ = H(q_1(0), \cdots, q_n(0), p_1(0), \cdots, p_n(0))$$

(conservació de l'energia), i, per tant, el punt que representa el nostre sistema dinàmic està obligat a romandre sobre una *superfície d'energia*, Ω, d'equació

$$H(q_1, \cdots, q_n, p_1, \cdots, p_n) = \text{const.} \tag{5.4}$$

Suposarem que la superfície d'energia Ω és un compacte i que és suficientment «regular» com per què li sigui aplicable la teoria elemental de les integrals de superfície i suposarem també que

$$\|\nabla H\|^2 = \sum_{i=1}^{n} \left(\frac{\partial H}{\partial p_i}\right)^2 + \left(\frac{\partial H}{\partial q_i}\right)^2 > c > 0. \tag{5.5}$$

Sigui $B \subset \Omega$ un conjunt sobre la superfície d'energia tal que estigui definida la integral

$$\int_B \frac{d\sigma}{\|\nabla H\|},$$

on $d\sigma$ és l'element de superfície. Definim la mesura $\mu\{B\}$ de B per mitjà de la fórmula

$$\mu\{B\} = \frac{\displaystyle\int_B \frac{d\sigma}{\|\nabla H\|}}{\displaystyle\int_\Omega \frac{d\sigma}{\|\nabla H\|}} \tag{5.6}$$

de manera que

$$\mu\{\Omega\} = 1. \tag{5.7}$$

Ara, per consideracions geomètriques senzilles i a partir del teorema de Liouville que hem comentat més amunt, es dedueix que

$$\mu\{T_t(B)\} = \mu\{B\}. \tag{5.8}$$

En altres paraules, T_t preserva la mesura μ en Ω.

La fórmula (5.6) només assigna mesura a uns quants conjunts elementals (als quals sigui aplicable la teoria elemental de la integració). No obstant això, la mesura es pot estendre a una col·lecció molt més àmplia de conjunts, de la mateixa manera que, començant amb intervals de la recta real i *definint* la mesura d'un interval com la seva longitud, hom construeix la mesura completament additiva de Lebesgue.

En particular, un conjunt C té μ-mesura 0 si per a tot $\epsilon > 0$ existeix una col·lecció finita o numerable B_i de conjunts elementals tals que

$$C \subset \bigcup_i B_i \qquad \text{i} \qquad \sum_i \mu\{B_i\} < \epsilon.$$

Ara estem en condicions d'enunciar en termes precisos el teorema de Poincaré invocat per Zermelo.

Si B és μ-mesurable llavors gairebé tot $P_0 \in B$ (és a dir, llevat d'un conjunt de μ-mesura 0) gaudeix de la propietat que per algun valor de t (possiblement depenent de P_0) $T_t(P_0) \in B$.

5.3. La resposta de Boltzmann

Per tal d'entendre la resposta de Boltzmann retornem a l'exemple dels dos contenidors. Suposem que coneixem la forma funcional exacta de la hamiltoniana

$$H(q_1, \cdots, q_n, p_1, \cdots, p_n) \tag{5.9}$$

i el seu valor C a $t = 0$. D'aquesta manera, coneixem la superfície d'energia (la seva equació és $H = C$).

A la condició que per $t = 0$ *totes* les partícules estiguin en un dels dos contenidors hi correspon clarament un conjunt B de punts de Ω. El nostre sistema, doncs, parteix del conjunt B.

La primera afirmació de Boltzmann deia que la μ-mesura $\mu\{B\}$ de B és «extremament» petita, cosa que està d'acord amb la nostra intuïció que estem partint d'un estat rar o altament inusual. Considerem, per altra banda, el conjunt R dels punts de Ω que correspon a estats en els quals el nombre de partícules en els dos contenidors és «molt aproximadament» proporcional al volum de cada contenidor. La $\mu\{R\}$ ha de ser «extremament» propera a 1.

Òbviament, aquestes afirmacions depenen en gran mesura del significat d'«extremament» i de «molt aproximadament», però n'hi hauria prou de dir que, a causa de l'enormitat d'àtoms per centímetre cúbic (de l'ordre de 10^{20}) és prou prudent interpretar «extremament» com ser menys de 10^{-10} i «molt aproximadament» com ser dintre d'un marge de 10^{-10} de la proporció correcta.

La segona afirmació de Boltzmann va ser molt agosarada. Deia que la primera afirmació *implica* que el temps relatiu que passa la corba real que descriu el moviment del sistema dintre de B i de R és, respectivament «extremament» petit i «extremament» gran.

En altres paraules, si el sistema es troba en un estat inusual, gairebé immediatament l'abandonarà (malgrat que pel teorema de Poincaré, gairebé amb tota seguretat hi retornarà eventualment) i, un cop dintre d'un conjunt que correspongui als estats «gairebé normals», hi romandrà «essencialment» per sempre.

Boltzmann va atacar la demostració de la primera afirmació amb estimacions plausibles però no massa rigoroses. Per a justificar la segona afirmació va introduir la *hipòtesi* que la corba que representa el moviment del sistema passa per *tots* els punts de la superfície d'energia.

Aquesta hipòtesi que Boltzmann va batejar com la *hipòtesi ergòdica* (*Ergodenhypothese*) és falsa (excepte per al cas $n = 1$ en què és trivial).

Boltzmann va intentar rescatar la seva explicació substituint la hipòtesi ergòdica errònia pel que va anomenar la hipòtesi quasiergòdica. Aquesta nova hipòtesi postulava que la corba del moviment passa arbitràriament prop de qualsevol punt de la superfície d'energia. Això, malgrat ser molt plausible, tampoc és suficient per tal d'establir la connexió entre el temps relatiu passat en un conjunt A i la seva μ-mesura, $\mu\{A\}$.

I el nucli de l'assumpte és, clarament, la connexió entre el temps relatiu passat en un conjunt A i la seva μ-mesura, $\mu\{A\}$.

Però, què volem dir exactament amb temps relatiu passat a A? La definició gairebé es suggereix tota sola de manera natural. Sigui $t(\tau, P_0, A)$ el temps que la corba del moviment que comença en P_0 passa dintre d'A fins al temps τ. El temps relatiu és el límit, quan existeix,

$$\lim_{\tau\to\infty} \frac{t(\tau, P_0, A)}{\tau}. \tag{5.10}$$

Resulta que la demostració de l'existència d'aquest límit constitueix la dificultat real. Un cop fet això, hom només necessita una suposició addicional sobre T_t per a concloure que el límit és igual a $\mu\{A\}$.

5.4. La formulació abstracta

Ara que ja hem passat molta estona amb el rerefons de la mecànica estadística, la deixaré de banda en gran part i n'extrauré el contingut purament matemàtic.

En comptes de la superfície d'energia, consideraré un conjunt Ω (de mesura total 1) sobre el qual està definida una mesura, μ, completament additiva.

Suposaré ara que tenim donada una família uniparamètrica de transformacions T_t de Ω en si mateix que preserva la μ-mesura. Aquesta afirmació requereix un comentari. En dinàmica, les transformacions T_t són bijectives (això és una conseqüència immediata de la unicitat de les solucions de les equacions hamiltonianes del moviment). No és, però, necessari suposar que les T_t són bijectives si hom defineix adequadament què s'entén per preservar la mesura.

La definició adequada és la següent. Sigui $T_t^{-1}(A)$ l'antiimatge del conjunt A. És a dir,

$$T_t(T_t^{-1}(A)) = A. \tag{5.11}$$

Direm que la transformació T_t preserva la mesura quan

$$\mu\left\{ T_t^{-1}(A) \right\} = \mu\{ A \}. \tag{5.12}$$

Per a transformacions bijectives, (5.12) és clarament equivalent a la definició habitual de preservació de la mesura, és a dir,

$$\mu\{ T_t(A) \} = \mu\{ A \}. \tag{5.13}$$

Sigui ara $P_0 \in \Omega$ i sigui $g(P)$ la funció característica del conjunt mesurable A. És a dir,

$$g(P) = \begin{cases} 1, & P \in A, \\ 0, & P \notin A. \end{cases} \tag{5.14}$$

Ara és prou clar que $t(\tau, P_0, A)$ vindrà donat per la fórmula

$$t(\tau, P_0, A) = \int_0^\tau g(T_t(P_0))\, dt\,, \tag{5.15}$$

i el problema és l'existència del límit

$$\lim_{\tau\to\infty} \frac{1}{\tau} \int_0^\tau g(T_t(P_0))\, dt\,. \tag{5.16}$$

Al mateix temps que aquesta versió en la qual el temps transcorre contínuament, és convenient considerar una versió discreta del problema.

Sigui T una transformació que preserva la mesura. És a dir,

$$\mu\left\{ T_t^{-1}(A) \right\} = \mu\{ A \}, \tag{5.17}$$

i considerem les seves potències (iteracions) $T^2, T^3, \cdots$. L'anàleg del límit (5.16) seria

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} g(T^k(P_0))\,. \tag{5.18}$$

El 1931, G. D. Birkhoff va aconseguir demostrar que els límits (5.16) i (5.18) existeixen per a gairebé tot P_0 (en el sentit de la μ-mesura). Una mica abans, John von Neumann va demostrar que els límits (5.16) i (5.18) existien en el sentit de la mitjana quadràtica.

Avui dia hi ha diverses demostracions d'aquests teoremes, i la més curta és la de F. Riesz. L'ometem i recomanem el lector l'excel·lent llibret de P. R. Halmos, *Lectures on Ergodic Theory*, publicat per la Societat Matemàtica del Japó.

Què podem dir del límit (5.18) [o (5.16)]?

Denotem aquest límit com a $h(P_0)$. Hom veu immediatament que es tracta d'una funció μ-mesurable, fitada (de fet, $0 \leq h(P_0) \leq 1$) i que, per a gairebé tot P_0,

$$h(T(P_0)) = h(P_0). \tag{5.19}$$

Sigui ara H_α el conjunt de P_0 pels quals

$$h(P_0) < \alpha\,,$$

i sigui $Q \in T^{-1}(H_\alpha)$. Així, $T(Q) \in H_\alpha$ i, per tant,

$$h(T(Q)) < \alpha.$$

Atès que per a gairebé tot Q, $h(T(Q)) = h(Q)$, veiem que $h(Q) < \alpha$ llevat d'un conjunt de Q de μ-mesura 0. En conseqüència, llevat d'un conjunt de μ-mesura 0,

$$T^{-1}(H_\alpha) = H_\alpha$$

per a tot α (el conjunt excepcional pot, òbviament, dependre de α).

En altres paraules, els conjunts H_α són *invariants* (llevat d'un conjunt de mesura 0).

Una transformació s'anomena *mètricament transitiva** si els únics conjunts invariants són de mesura 0 o de mesura 1.

Si suposem que la nostra transformació és mètricament transitiva, veiem que els conjunts H_α són de mesura 0 o de mesura 1 i, en conseqüència, $h(P_0)$ ha de ser constant gairebé pertot.

El valor d'aquesta constant es determina fàcilment observant que

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} g(T^k(P_0)) = h(P_0) \qquad \text{(gairebé pertot)}$$

implica (pel teorema de la convergència fitada) que

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \int_\Omega g(T^k(P_0))\, d\mu = \int_\Omega h(P_0)\, d\mu\,. \tag{5.20}$$

De fet,

$$\int_\Omega g(T^k(P_0))\, d\mu = \int_\Omega g(P_0)\, d\mu = \mu\{\,A\,\}$$

(això és una conseqüència immediata del fet que T preservi la mesura), i, per tant,

$$\int_\Omega h(P_0)\, d\mu = \mu\{\,A\,\}\ .$$

*N. del tr. Avui s'estila més dir-ne *ergòdica*.

La constant, doncs, és igual a $\mu\{A\}$.

Combinant tot això, podem dir que, si T és mètricament transitiva, llavors per a gairebé tot P_0,

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n} g(T^k(P_0)) = \mu\{A\}\,. \tag{5.21}$$

Aquest resultat es generalitza fàcilment de la següent manera. Si $f(P_0)$ és μ-integrable, és a dir,

$$\int_\Omega |\,f(P_0)\,|\;d\mu < \infty,$$

i si T és mètricament transitiva, llavors, per a gairebé tot P_0,

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n} f(T^k(P_0)) = \int_\Omega f(P_0)\,d\mu\,. \tag{5.22}$$

Hom pot pensar que la demostració de (5.22) reivindica completament les idees de Boltzmann, però, malauradament, les transformacions T_t a les quals s'arriba en dinàmica són tan complexes que, llevat d'alguns casos molt simples, no se sap si són mètricament transitives o no. Aquest fet, però, no detrau la bellesa i la importància del teorema ergòdic.

5.5 El teorema ergòdic i les fraccions continuades

Sigui x, $0 < x \leq 1$, un nombre real i desenvolupem-lo en fracció continuada

$$\cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \ddots}}} \tag{5.23}$$

on $a_1, a_2, \cdots$ són enters positius. És fàcil trobar fórmules per a les a.

Tenim

$$a_1(x) = \left[\frac{1}{x}\right], \qquad a_2(x) = \left[\frac{1}{\frac{1}{x} - \left[\frac{1}{x}\right]}\right], \quad \cdots,$$

on, com és habitual, $[y]$ denota la part entera de y.

Les fórmules per a les a es fan progressivament més i més complicades, però una mica de reflexió mostra que es poden incloure totes dintre d'un mateix esquema.

Sigui

$$T(x) = \frac{1}{x} - \left[\frac{1}{x}\right]\;; \tag{5.24}$$

llavors,

$$\begin{aligned} a_2(x) &= a_1(T(x)) \,, &(5.25)\\ a_3(x) &= a_2(T(x)) = a_1(T^2(x)) \,, &(5.26)\end{aligned}$$

etc.

Ara que estem treballant amb iteracions de la transformació $T(x)$ donada per (5.24), és evident la possibilitat d'aplicar el teorema ergòdic.

Quin és l'espai Ω? Simplement l'interval $(0,1)$ amb el 0 exclòs. Quina és la mesura invariant? Això és més difícil de respondre, però ja ho va fer, essencialment, Gauss.

Hom pot procedir de la manera següent. Sigui $\rho(x)$, $0 < x \leq 1$, tal que

$$\text{(a)} \qquad \rho(x) \geq 0, \qquad \text{(b)} \qquad \int_0^1 \rho(x)\, dx = 1, \tag{5.27}$$

i definim $\mu\{\,A\,\}$ per mitjà de la fórmula

$$\mu\{\,A\,\} = \int_A \rho(x)\, dx \,. \tag{5.28}$$

Prenem ara un interval (α, β), $0 < \alpha < \beta < 1$, i considerem la seva antiimatge per la transformació $T(x)$. Tenim

$$T^{-1}(\alpha,\beta) = \bigcup_{k=1}^{\infty} \left(\frac{1}{k+\beta}, \frac{1}{k+\alpha} \right) \tag{5.29}$$

i, per tant,

$$\mu\left\{\, T^{-1}(\alpha,\beta) \,\right\} = \sum_{k=1}^{\infty} \int_{1/(k+\beta)}^{1/(k+\alpha)} \rho(x)\, dx \,. \tag{5.30}$$

Si s'ha de preservar μ, hem de tenir

$$\int_\alpha^\beta \rho(x)\, dx = \sum_{k=1}^{\infty} \int_{1/(k+\beta)}^{1/(k+\alpha)} \rho(x)\, dx \tag{5.31}$$

per a tot α i β. No sabem com procedir sistemàticament per tal de resoldre (5.31), però és fàcil comprovar que

$$\rho(x) = \frac{1}{\log 2} \frac{1}{1+x} \tag{5.32}$$

és una solució i satisfà les condicions (5.27).

Ara tenim tot el que necessitem per tal de comprovar que $T(x)$ és mètricament transitiva, que és completament trivial.*

*Aquí em vaig deixar endur per un impuls d'optimisme inexplicable. Mentre que el fet és intuïtivament obvi, la demostració, malauradament, no ho és. La podeu trobar en el § 2 de l'article de Ryll-Nardzewski citat en el final d'aquest capítol. La demostració és de K. Knopp.

Si $f(x)$ és μ-integrable, és a dir,

$$\frac{1}{\log 2}\int_0^1 |\, f(x)\,|\, \frac{dx}{1+x} < \infty, \tag{5.33}$$

llavors, per (5.22)

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n} f(T(x)) = \frac{1}{\log 2}\int_0^1 f(x)\,\frac{dx}{1+x} \tag{5.34}$$

per a gairebé tot x (observeu que els conjunts de μ-mesura 0 són idèntics als conjunts de mesura 0 en el sentit de la mesura ordinària de Lebesgue).

Sigui ara

$$f(x) = \log a_1(x)\,. \tag{5.35}$$

A partir de (5.34) obtenim que, per a gairebé tot x,

$$\lim_{n\to\infty}(a_1 a_2 \cdots a_n)^{1/n} = C\,, \tag{5.36}$$

on

$$\begin{aligned} C &= \exp\left(\frac{1}{\log 2}\int_0^1 \log a_1(x)\,\frac{dx}{1+x}\right) \\ &= \exp\left(\frac{1}{\log 2}\sum_{k=1}^{\infty}\log k \log\frac{(k+1)^2}{k(k+2)}\right)\,. \end{aligned} \tag{5.37}$$

Aquest teorema tan notable va ser demostrat per primera vegada per Khintchine el 1935 (per un mètode diferent).[†] La demostració que hem donat aquí és de C. Ryll-Nardzewski.

Podria haver estalviat al lector les tres primeres seccions d'aquest capítol. Podria haver començat amb la formulació abstracta del § 5.4 i no fer cap esment a la dinàmica o a la teoria cinètica. Però si hagués procedit d'aquesta manera, hauria suprimit la part més excitant i, al meu entendre, més instructiva de la història. Perquè el camí de la teoria cinètica fins a les fraccions continuades, tal com l'havien concebut Boltzmann i altres és un fantàstic exemple del fet, tantes vegades oblidat, que les matemàtiques no són un ens aïllat sinó que molta de la seva bellesa i del seu poder els deuen a altres disciplines.

PROBLEMES

1. Sigui $B \subset \Omega$ un conjunt μ-mesurable i sigui $\mu\{\,B\,\} \neq 0$. Si T preserva la mesura (però no és necessàriament mètricament transitiva), demostreu que per a gairebé tot $P_0 \in B$ existeix un enter $n \geq 1$ tal que

[†]N.T. La constant $C \approx 2.685452\ldots$ es coneix com *constant de Khintchine*.

$T^n(P_0) \in B$. (Aquesta és la versió discreta del teorema de Poincaré; per a demostrar-ho, considereu el conjunt $C \subset B$ tal que si $P_0 \in C$, llavors $T^n(P_0) \notin B$ per $n = 1, 2, \cdots$. Demostreu, després, que C és μ-mesurable i que $C, T^{-1}(C), T^{-2}(C), \cdots$ són tots disjunts).

2. Sigui $n(P_0)$, $P_0 \in B$, el *primer* enter positiu tal que $T^{n(P_0)}(P_0) \in B$. Si T (a més a més de preservar la mesura) és mètricament transitiva, demostreu que
$$\int_B n(P_0)\, d\mu = 1.$$

3. Sigui x, $0 < x \leq 1$, desenvolupat en fracció continuada
$$x = \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \ddots}}},$$
i sigui B el conjunt en el qual $a_1(x) = k$ (és a dir, $1/(k+1) < x \leq 1/k$). Sigui $n(x, k)$ l'enter més petit *més gran* que 1 i tal que $a_{n(x,k)} = k$. Demostreu que
$$\frac{1}{\log 2} \int_{1/(k+1)}^{1/k} (n(x, k) - 1) \frac{dx}{1 + x} = 1.$$

4. Sigui $0 \leq x \leq 1$ i $T(x) = 2x - [x]$. Deduïu el teorema de Borel del capítol 2 com una aplicació del teorema ergòdic.

BIBLIOGRAFIA

C. Ryll-Nardzewski, «On the ergodic theorems II», *Studia Math.*, **12** (1951), 74–79.