

Tests de generadors de nombres a l'atzar

Robert Tarragona Llovera *

Motivació

La simulació usant variables aleatòries o simulació de *Montecarlo* és una eina que avui en dia té un gran nombre d'aplicacions en la física d'altres energies, la matèria condensada, etc. S'utilitza amb gran efectivitat en el tractament de molts problemes per als quals és molt difícil o impossible trobar una solució analítica. Les tècniques de Montecarlo incorporen en els seus algorismes una part d'atzar destinada a compensar la manca (o l'excés) d'informació sobre el sistema en estudi. Dins d'aquesta part, l'algorisme ha de prendre decisions sobre si accepta o no el resultat d'un càlcul basant-se en unes determinades lleis de probabilitat. En aquest procés prenen gran importància les sèries de nombres aleatoris generades pels ordinadors.

Tots els llenguatges de programació disposen de rutines de generació de nombres a l'atzar i sovint ens pot interessar avaluar-ne l'eficàcia per comprovar si poden ser la font d'algun resultat erroni en executar un programa que les cridi.

Així doncs, amb l'objectiu que qualsevol usuari pugui fer tests d'una determinada rutina, es presenten un grup de proves senzilles, fàcilment implementables, que permeten comprovar si una sèrie de nombres que suposem independents i distribuïts uniformement en $(0, 1)$ són aptes per a la utilització en programes de simulació. Abans, però, es descriuen breument els procediments que es fan servir en els ordinadors per retornar nombres aleatoris.¹

Algorismes de generació

Els nombres generats en cridar una rutina en un ordinador provenen d'un algorisme que utilitza com a entrada un cert nombre de llavors.² A partir d'aquestes llavors l'ordinador fa uns determinats càlculs i retorna un grup de nombres pretesament aleatoris.

Durant anys s'han anat cercant algorismes que permetessin obtenir bons resultats en aquest sentit. El més utilitzat actualment és l'anomenat *mètode lineal congru-*

encial. Aquest algorisme té una gran importància per diverses raons:

- És un procediment molt senzill d'entendre i programar.
- S'ha estat utilitzant durant pràcticament vint anys i se n'ha estudiat el comportament i les propietats amb tot detall.
- És un algorisme que treballa a una alta velocitat.
- L'última raó i la més important és que la seva sortida, si s'elegeixen de forma adequada els paràmetres que hi apareixen, és d'una gran qualitat estadística, sovint millor que la d'algorismes més complicats.

El mètode parteix de les quantitats senceres següents: el mòdul m ($m > 0$), el multiplicador a ($0 \leq a < m$), l'increment c ($0 \leq c < m$) i la llavor o valor inicial X_0 ($0 \leq X_0 < m$). Es genera a continuació la seqüència de nombres sencers X_n donada per:

$$X_{n+1} = (aX_n + c) \bmod m \quad n \geq 0, \quad (1)$$

on mod denota la resta de la divisió per m . Per tal d'obtenir una sèrie de nombres pseudoaleatoris entre $[0,1)$ només cal fer:

$$U_n = \frac{X_n}{m}.$$

A partir de l'equació (1) veiem que, si en un determinat moment es repeteix un nombre dels obtinguts prèviament, tots els següents també es repetiran. La seqüència de nombres obtinguts U_n és, per tant, cíclica. A la quantitat de nombres generats quan es tanca un cicle complet se l'anomena *període* del generador i com a molt pot valer m . Així doncs, una condició necessària per obtenir una major quantitat de nombres diferents és prendre m gran.

L'elecció òptima dels altres paràmetres es fa utilitzant enunciats teòrics demostrats (Knuth, 1968 i 1982). Concretament se sap que el màxim període possible es pot assolir quan

$$m = 2^e, \quad a = 2^k + 1, \quad 2 \leq k < e,$$

*Robert Tarragona Llovera és becari de col·laboració del Departament d'Estructura i Constituents de la Matèria, de la Facultat de Física de la UB.

¹En el nostre cas seria més propi parlar de nombres *pseudoaleatoris* ja que tractem amb nombres construïts mitjançant un algorisme. En endavant utilitzarem ambdues nomenclatures indistintament.

²Valors numèrics que inicialitzen la rutina i en determinen l'evolució.

i que en aquest cas tenim llibertat per escollir $c = 0$. La llavor ens pot desplaçar l'inici de la sèrie, però no canviarà el període de la sèrie resultant.

Una possible millora a l'algorisme s'obté si es genera una taula de nombres a l'atzar i després es desordenen mitjançant un altre generador. Aquesta millora és fàcilment implementable en els ordinadors vectorials o paral·lels, la qual cosa n'ha popularitzat l'ús.

Tests empírics de nombres a l'atzar

Sota aquest nom s'engloben un conjunt de proves estadístiques que es realitzen sobre una gran quantitat de nombres provinents d'un generador per avaluar-ne la qualitat estadística. En distingim dos casos: els tests d'uniformitat i els tests de correlacions.

Els tests d'uniformitat

Aquests tests comproven si els nombres entrats estan uniformement distribuïts en l'interval $(0, 1)$. Això vol dir que si partim l'interval en parts iguals i fem un histograma per veure quants nombres dels generats cauen en cada part, les barres tendiran a tenir la mateixa alçada quan tinguem una quantitat prou gran de nombres.

Aquesta informació, tot i ser una contribució necessària per a qualsevol generador, no és en cap cas suficient. Un exemple simple pot posar de manifest que sèries que es comporten bé en aquest sentit no poden en cap cas ser catalogades d'aleatòries:

$$\frac{1}{n}, \frac{2}{n}, \frac{3}{n}, \dots, \frac{n-1}{n}, 1.$$

En aquesta sèrie els nombres es distribueixen uniformement, però van apareixent consecutivament fent completament previsible el següent element del grup.

Així doncs, no n'hi ha prou a mirar les dades en conjunt, sinó que cal tenir en compte les possibles correlacions que hi pugui haver entre els elements que componen la llista.

Els tests de correlacions

Els tests de correlacions són importants per intentar trobar connexions entre els nombres que componen una sèrie U_1, U_2, \dots, U_n . Hi ha dos procediments estàndard per fer-ho:

En primer lloc es pot estendre el concepte de test d'uniformitat de manera que en lloc d'estudiar quants nombres U_i cauen en un determinat segment $(a_1, b_1) \subset (0, 1)$ es comptin quantes k -tuples $(U_{i+1}, U_{i+2}, \dots, U_{i+k})$ cauen en $(a_1, b_1) \times (a_2, b_2) \times \dots \times (a_k, b_k) \subset (0, 1)^k$. Diem que estem avaluant l'aleatorietat k -dimensional d'un generador.

Un mètode alternatiu de fer tests consisteix a estudiar el caràcter local de la sèrie mesurant dades com ara la llargada de subsuccions monòtones, la quantitat de nombres que cauen consecutivament en un determinat

segment, la longitud de subsèries necessària per omplir un recobriment, etc.

Tests estadístics

Per tal de realitzar els tests empírics explicats en l'apartat anterior és necessari estudiar alguns tests que s'usen en l'estadística per analitzar dades que suposem que s'ajusten a una determinada distribució. D'entre aquests en destaquen dos: el test de χ^2 i el de Kolmogorov-Smirnov.

El test de χ^2

El test de χ^2 és, possiblement, el test més utilitzat en les ciències experimentals i matemàtiques, amb múltiples aplicacions en l'acceptació o refús d'hipòtesis.

Per realitzar el test de χ^2 , suposarem que s'han realitzat una sèrie d'experiments *independents* i s'han obtingut n resultats $U_0, U_1, U_2, \dots, U_n$ que es poden classificar en k parts diferents. Considerarem, a més, que se'ns ha proporcionat un determinat model d'acord amb el qual atribuïm a cada part una probabilitat p_j , $j = 1, \dots, k$. Es pretén determinar si les dades estudiades s'ajusten o no a aquesta distribució de probabilitat *discreta*.

Per fer-ho es designa per Y_j el nombre de cops que una de les U ha caigut en la part j -èssima ($j = 1, \dots, k$) i es compara aquest valor amb el resultat que s'esperaria obtenir partint del model, és a dir np_j . Concretament, es calcula l'estadístic de χ^2 que ve donat per:

$$C = \sum_{j=1}^k \frac{(Y_j - np_j)^2}{np_j}. \quad (2)$$

C és un valor positiu obtingut com a suma de les desviacions $Y_j - np_j$ al quadrat i a les quals s'ha afegit un divisor per donar més importància a les parts amb p_j petita. En general esperarem doncs que els valors de C que s'obtinguin siguin 'petits'.

Per tal d'interpretar quan un valor de C és correcte o no s'utilitza una dada important de l'estadístic C : la seva distribució asimptòtica, és a dir si n és prou gran,³ és una χ^2 amb $k - 1$ graus de llibertat (Marsaglia, 1985). La funció de distribució de la llei χ^2 està tabulada en detall en diversos llibres, per a diferents valors de k . A més, poden generar-se fàcilment nombres aleatoris ajustats a aquesta distribució tenint en compte que es construeix sumant els quadrats de k lleis normals $N(0, 1)$.

Per entendre millor aquesta descripció s'explica l'exemple següent: suposem que es vol estudiar si una parella de daus estan trucats. Per fer-ho es llença la parella $n = 200$ cops; en cada tirada se sumen els punts i se n'noten els resultats. En cada prova hi ha $k=11$ resultats possibles i el que s'obtidria seria una taula com la següent:

³Cal que es verifiqui que $np_j \geq 5$ per a tot j . El test pot donar estimacions altament errònies per a valors de n menors!

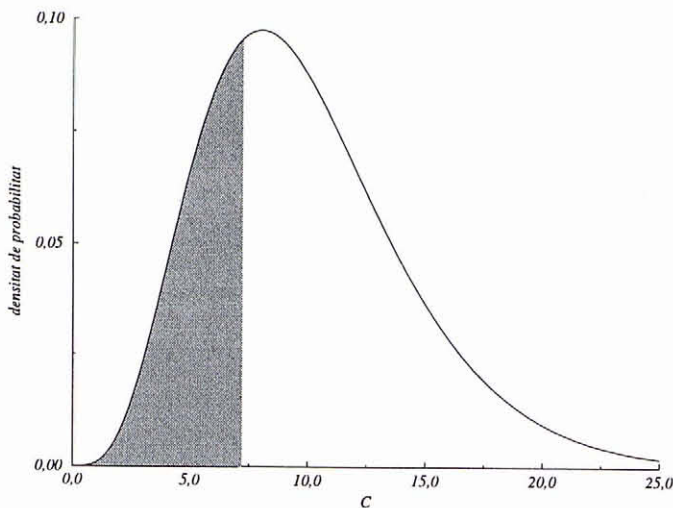


Figura 1: Llei de χ^2 amb 10 graus de llibertat

j	1	2	3	4	5	6	7	8	9	10	11
S_j	2	3	4	5	6	7	8	9	10	11	12
Y_j	9	13	19	20	24	30	27	23	19	14	2

on S_j és el resultat de la suma i Y_j el nombre de cops que ha sortit. Clarament les probabilitats que s'ajusten al nostre model ideal són $p_1 = p_{11} = 1/36$, $p_2 = p_{10} = 2/36, \dots, p_j = (6 - |j - 6|)/36; j = 1, \dots, 11$. A continuació calculem el valor de l'estadístic corresponent:

$$C = \frac{(9 - 200 \cdot \frac{1}{36})^2}{200 \cdot \frac{1}{36}} + \frac{(13 - 200 \cdot \frac{1}{18})^2}{200 \cdot \frac{1}{18}} + \dots + \frac{(2 - 200 \cdot \frac{1}{36})^2}{200 \cdot \frac{1}{36}} = 7,26.$$

En la figura 1 s'ha representat la funció densitat de probabilitat de la llei de χ^2 amb 10 graus de llibertat i s'ha ombrejat l'àrea sota la corba compresa entre les abscisses 0 i 7,26; aquesta àrea val 0,364.⁴ Això vol dir que el 36,4 % dels cops que es realitzi l'experiment, per a una parella de daus no trucada, s'obindrà un valor de C menor.

En aquest cas sembla lògic acceptar la parella de daus que ens han ofert. En general acceptem el model proposat quan el valor de la funció de distribució en l'abscissa que hem obtingut està comprès entre 0,05 i 0,95. Cal remarcar, però, que fins i tot per a uns daus trucats hi ha una certa probabilitat d'obtenir uns valors de C en els límits especificats. I, recíprocament, per a uns daus no trucats, es poden tenir, eventualment, valors de C deficientes. Seria convenient, doncs, repetir el test diversos cops (de 3 a 5) i mirar-ne els resultats globalment. I, encara millor, podríem fer uns quants experiments canviant el nombre de llançaments per tal de millorar la

⁴El valor s'ha extret buscant el valor de la funció de distribució en el punt 7,26 mitjançant interpolació dels valors tabulats (Abramowitz i Stegun, 1965).

fiabilitat de la prova. En fer-ho, obtenim els següents valors de C :

	Prova 1	Prova 2	Prova 3	Prova 4
$n = 200$	7,26	10,11	9,28	5,11
$n = 1.000$	12,11	6,85	4,92	19,12
$n = 2.000$	11,72	17,00	12,06	14,61
$n = 10.000$	17,92	5,11	8,91	12,42

El motiu pel qual hem usat diferents nombres de llançaments és que els daus podrien proporcionar valors correctes per a un cert rang de valors i incorrectes per a uns altres (per exemple, podria passar que els daus sofrissin un desgast per a un n gran). Aquesta situació pot passar quan es fan tests de nombres aleatoris. Hi ha generadors que es comporten correctament quan generen una quantitat relativament petita de nombres però que fallen per a n alts. Caldrà preveure una varietat suficient de proves per mirar de compensar aquesta possibilitat. Dels setze valors de la taula anterior només el 19,12 no passaria individualment el test de χ^2 . Tenint en compte que una parella de daus no trucada no el passaria un de cada deu cops, obtenir un únic valor discordant resulta ser una bona proporció, la qual cosa ens permet, ara sí, acceptar el test amb major seguretat.

El test de Kolmogorov-Smirnov

El test de Kolmogorov-Smirnov (KS) és una prova que es realitza sobre una seqüència de nombres X_1, X_2, \dots, X_n per tal de determinar si és factible que provinquin d'una determinada distribució *contínua*.⁵ Per tal d'entendre com ho fa és necessari introduir el concepte de funció de distribució empírica associada a la sèrie $\{X_i\}$:

$$F_n(x) = \frac{\text{Nombre de } X \leq x}{n}.$$

Per tal d'imaginar més clarament el seu aspecte, suposarem que la sèrie està ordenada de forma creixent i de manera que no hi ha nombres repetits (la generalització és senzilla). Aleshores F_n és una funció esglaonada definida en tota la recta real de la manera següent:

- En $(-\infty, X_1)$ val 0;
- En $[X_i, X_{i+1})$ val $i/n; (i = 1, \dots, n - 1)$;
- En $[X_n, +\infty)$ val 1.

Per exemple, si prenem la sèrie formada pels següents valors ja ordenats:

-1, 11 - 0,52 - 0,31 - 0,17 0,03 0,27 0,62 1,22

i representem gràficament $F_n(x)$, obtindrem la funció en forma d'escala de la figura 2. Aquesta funció de distribució s'ajusta perfectament a la sèrie de partida. Ara bé, com es pot determinar si els X s'ajusten a una determinada llei de probabilitat amb funció de distribució

⁵S'entén per distribució contínua tota aquella que té una funció de distribució absolutament contínua, per exemple una gaussiana.

F ? Doncs estudiant si F i F_n són semblants o no. Això és el que fa el test de KS. Per a tal fi el test calcula els dos estadístics següents:

$$K_n^+ = \sqrt{n} \max_{-\infty < x < +\infty} \{F_n(x) - F(x)\}; \quad (3)$$

$$K_n^- = \sqrt{n} \max_{-\infty < x < +\infty} \{F(x) - F_n(x)\}.$$

Fixem-nos que aquests K_n mesuren una mena de 'distància' entre ambdues funcions. Més concretament, el màxim de les distàncies verticals en tots els punts de la recta real.⁶ Per tant, desitjaríem que els valors que se'n desprenguin no fossin gaire alts.

Com passava en el test de χ^2 , la funció de distribució de K_n^+ i K_n^- és coneguda (és la mateixa per a ambdós) i, per tant, podem determinar quan un valor obtingut és correcte o no. Aquests valors estan ben tabulats pels diferents valors de n .

Cal dir que la fórmula (3), que s'ha utilitzat per definir els dos estadístics de KS, no és de gaire utilitat pràctica per calcular-los. Però si tenim en compte que tant la distribució empírica com la teòrica són funcions monòtones, es pot verificar que el següent algorisme permet calcular-los ràpidament:

- S'ordenen les observacions X_1, X_2, \dots, X_n en ordre creixent, de manera que $X_1 \leq X_2 \leq \dots \leq X_n$.
- Es calculen els estadístics utilitzant les fórmules:

$$K_n^+ = \sqrt{n} \max_{1 \leq j \leq n} \left(\frac{j}{n} - F(X_j) \right); \quad (4)$$

$$K_n^- = \sqrt{n} \max_{1 \leq j \leq n} \left(F(X_j) - \frac{j-1}{n} \right).$$

D'aquesta manera, només cal que s'evaluin les funcions en els n punts que componen la sèrie.⁷

Tornant a l'exemple anterior, en la mateixa figura 2 s'ha dibuixat amb traç discontinu la funció de distribució a la qual suposem que s'ajusten les nostres dades, en aquest cas una llei $N(0, 1)$. Per realitzar el test cal construir una taula com la següent:

X_j	-1,11	-0,52	-0,31	-0,17	0,03	0,27	0,62	1,22
$F(X_j)$	0,134	0,302	0,378	0,433	0,512	0,606	0,732	0,889
$j/n - F(X_j)$	-0,009	-0,052	-0,003	0,067	0,113	0,144	0,143	0,111
$F(X_j) - (j-1)/n$	0,134	0,177	0,128	0,058	0,012	-0,019	0,018	0,014

D'aquí es poden llegir directament $K_g^+ = 0,144 \times \sqrt{8} = 0,407$ i $K_g^- = 0,177 \times \sqrt{8} = 0,501$. Mirant les taules es comprova que la funció de distribució de K_g^+ i K_g^- pren valors menors el 52% i el 68% dels cops

⁶Per tal d'interpretar el factor \sqrt{n} que hi apareix es pot consultar Morgan, (1984).

⁷Mirant la figura 2 es pot entendre per què només és necessari avaluar els estadístics en els punts de salt de F_n .

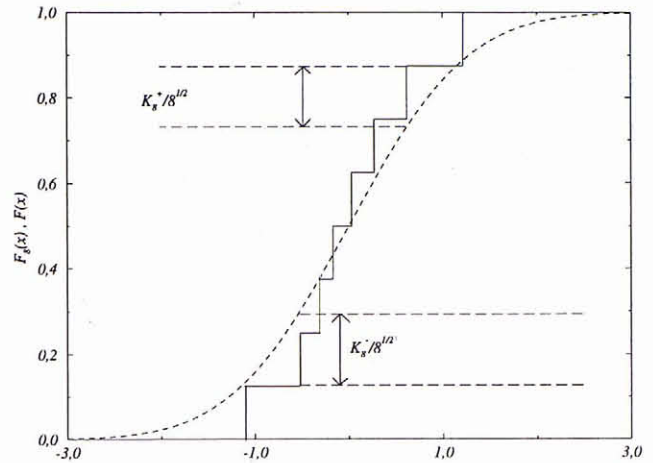


Figura 2: Funcions de distribució empírica i teòrica

respectivament i, per tant, acceptem que els valors de la sèrie s'ajusten a una gaussiana.

Tests de nombres a l'atzar

Test d'uniformitat

El primer test d'uniformitat és un test de χ^2 . S'ha de partir l'interval $(0, 1)$ en k parts diferents i veure en quina de les parts van a parar els n nombres estudiats. Les probabilitats que s'han d'utilitzar per al test són: $p_i = l_i/n$, on l_i és la longitud del segment i -èssim, $i = 1, \dots, k$.

Segon test d'uniformitat

Partim de la hipòtesi que els nombres que ens proporciona el generador en estudi s'ajusten a una distribució uniforme en $(0, 1)$. La funció de distribució corresponent és contínua i ve donada per $F(x) = x, 0 \leq x < 1$. Així doncs podem realitzar un test de KS sobre els nombres provinents del generador en estudi per tal de determinar si la hipòtesi és certa o no.

Com ja hem mencionat anteriorment, a causa de la seva simplicitat, aquests tests acostumen a ser passats per la gran majoria de generadors. S'han inclòs perquè els seus algorismes són molt similars als dos primers tests de correlacions que es presenten a continuació. Així doncs, si obtenim pèssims resultats en aquestes primeres proves no valdrà ja la pena que utilitzem els tests de correlacions, més elaborats i restrictius.

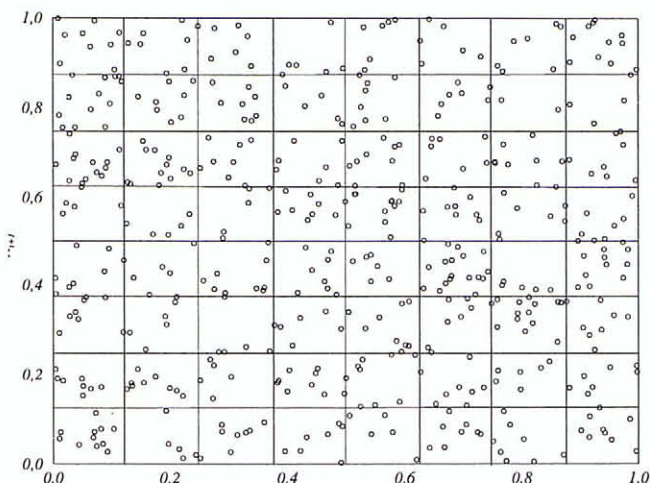


Figura 3: Test de sèries per parelles

Test de sèries

En el test de sèries s'avalua l'aleatorietat de parelles de nombres consecutius (X_j, X_{j+1}) . Per fer-ho podem representar gràficament X_{j+1} en funció de X_j en un quadrat 1×1 . El que s'obté és un núvol de punts com el representat en la figura 3. Si a continuació partim el quadrat en un cert nombre k de regions diferents esperarem que el nombre de punts en cada regió sigui proporcional a la seva àrea.

Per veure si això és així podem realitzar un test de χ^2 amb $k-1$ graus de llibertat amb probabilitats a_j , sent a_j l'àrea de la regió j . De forma anàloga es podrien fer tests de les sèries triples, quàdruples, etc. D'aquesta manera, el fet d'obtenir resultats negatius en fer el test ens indicarà l'existència de regions amb un excés o una manca de punts, incompatible amb la hipòtesi d'uniformitat.

En la figura 3 es mostra la manera habitual de recollir els resultats per a $k = 64$. Les probabilitats són totes iguals a $1/64$. Si projectem tots els punts sobre un dels dos eixos podem llegir les dades per un test d'uniformitat, la qual cosa ens indica que si es passa el test de sèries automàticament es passa el d'uniformitat, però no a la inversa.

Test del màxim de t

Es tracta d'una extensió a t dimensions del test d'uniformitat de KS. Donada una sèrie U_1, U_2, \dots, U_{n-t} es defineix $V_j = \max(U_{tj}, U_{tj+1}, \dots, U_{tj+t-1})$ per $0 \leq j < n$. D'aquesta manera s'obtenen els màxims d'entre t elements per les n subcol·leccions. A continuació es realitza un test de KS a la seqüència V_0, V_1, \dots, V_{n-1} amb funció de distribució $F(x) = x^t, 0 \leq x < 1$.

Com en el cas del test de sèries, convé realitzar proves per a diferents dimensionalitats, ja que es podria

donar el cas que un generador oferís problemes només en una dimensió concreta, d . Si això succeeix podem obtenir resultats erronis en programes que cridin nombres a l'atzar en grups de d , ja que aquests no cobreixen uniformement tot el rang de valors a què poden accedir.

Test dels grups

En un bon generador, amb poques correlacions, els valors que anem obtenint són independents dels anteriors. Sovint, però, pot passar que els valors propers tinguin una certa tendència a estar els uns prop dels altres o a l'inrevés. Per veure si això succeeix en el generador que estem provant podem delimitar una regió de l'interval $(0,1)$ i veure quants nombres seguits hi cauen. Si repetim aquesta operació un nombre n molt gran de cops obtindrem informació sobre possibles influències mútues dins de la sèrie total.

Concretament, donats $a, b \in (0,1)$ tals que $0 \leq a < b \leq 1$ es vol estudiar la longitud de subseqüències consecutives $U_j, U_{j+1}, \dots, U_{j+r}, U_{j+r+1}$ de manera que U_j i U_{j+r+1} cauen fora de (a,b) i les altres U , dins. A una subsèrie com aquesta se l'anomena 'grup' de longitud r . Es tracta, donada una sèrie sencera, de comptabilitzar quants grups de longitud $0,1, \dots$ hi apareixen i avaluar-ne el resultat.

Per exemple, en la següent sèrie aleatòria fictícia arrodonida a dos decimals, on hem pres $a = 0, b = 0,5$, les longituds seran:

0,61 0,72 0,24 0,77 0,32 0,09 0,48 0,50

0 1 3

0,17 0,11 0,72

2

Si escrivim $p = b - a$, la probabilitat de tenir un grup de longitud 0 és $p_0 = 1 - p$, ja que aquesta situació correspon a obtenir un element fora de (a,b) ; la de tenir un grup de longitud 1 és $p_1 = p(1 - p)$, ja que aquesta situació requereix que el primer nombre caigui en (a,b) (probabilitat p) i que el segon caigui fora (probabilitat $1 - p$); de manera semblant obtindriem $p_2 = p^2(1 - p), \dots, p_r = p^r(1 - p)$, etc. Un cop coneixem aquestes probabilitats, podem aplicar un test de χ^2 a les n longituds que hem recollit provinents del nostre generador.

A mesura que augmenta la longitud, la probabilitat de trobar un grup d'aquesta mida disminueix. Convé doncs fixar una longitud màxima d , de manera que tots els grups de mida igual o superior es comptabilitzin en una mateixa classe a la qual associem una probabilitat $p_d = 1 - \sum_{r=0}^{d-1} p_r = 1 - (1 - p) \sum_{r=0}^{d-1} p^r = p^d$.

Test de permutacions

Un altre aspecte important a tenir en compte en fer tests d'un generador és si els nombres que van apareixent tenen una major tendència a créixer que a decreixer. Tot i que a primer cop de vista això sembla que contradigui

el requisit d'uniformitat, podem pensar en un generador que de mitjana tingui el doble de xifres creixents que decreixents, però que el valor del creixement mitjà sigui la meitat del decreixement mitjà. Si es dona una situació similar ni el test d'uniformitat ni el de sèries i, en alguns casos, ni tan sols el de grups podrien detectar aquesta anomalia. El que es fa a continuació és estudiar les possibles ordenacions mútues en què poden aparèixer els valors de la sèrie a estudiar.

Per fer-ho es pren un grup de t nombres de la sèrie $\{U_i, U_{i+1}, \dots, U_{i+t-1}\}$ i es classifica segons l'ordre relatiu dels seus components. Per exemple, si fixem $t = 3$ podem tenir $U_i < U_{i+1} < U_{i+2}$ a la qual se li podria associar el valor 1, $U_i < U_{i+2} < U_{i+1}$ amb valor 2, i així successivament fins al valor 6 de $U_{i+2} < U_{i+1} < U_i$.⁸ Aquesta classificació es pot repetir un nombre n molt gran de cops i realitzar un test de χ^2 amb probabilitats totes iguals a $1/t!$. Com abans, guanyarem representabilitat si variem la llargada de les subsèries, t .

Conclusió

S'han exposat alguns procediments per fer tests de sèries de nombres a l'atzar, però n'hi ha molts d'altres, alguns d'ells molt complexos. La millor manera d'esbrinar si un d'aquests tests ens pot aportar resultats significatius

Bibliografia

- ABRAMOWITZ, M. i STEGUN, I.A., *Handbook of Mathematical Functions*, Dover Publications (Nova York, 1965).
 FISHMAN, G.S., *Principles of discrete event simulation*, Wiley (Nova York, 1976).
 KNUTH, D.E., *The art of computer programming, Vol. 1 i 2*, Addison-Wesley Reading (Massachusetts, 1968 i 1982).
 MARSAGLIA, W.H., *Computer science and statistics: the interface*, Elsevier (Amsterdam, 1985).
 MORGAN, J.P., *Elements of Simulation*, Chapman and Hall (Nova York, 1985).
 PRESS, W.H., FLANNERY, B.P., TEUKOLSKY, S.A. i VETTERLING, W.T., *Numerical recipes, the art of scientific computing*, Cambridge University Press (Cambridge, 1989).

és provant-los en diferents màquines. Es comprovarà que la gran majoria de les sèries passen els tests d'uniformitat, però que fallen en alguna de les altres proves. Com s'ha dit aquestes s'han de repetir un cert nombre de cops i canviar els valors de n en aquells que requereixin l'aplicació d'un de χ^2 . S'escullen típicament valors $n \sim 10^5 - 10^8$. Habitualment cal parar atenció en programes on es cridin una gran quantitat de nombres, de l'ordre del període del generador (vegeu l'anotació (8) al peu de pàgina) o bé programes en què haguem de classificar els nombres obtinguts en una gran quantitat d'apartats (només d'alguns ordres per sota del mateix període), o bé que usin els nombres cridats en petites agrupacions.⁹ És en aquests casos que cal parar una atenció especial a la qualitat dels nombres generats per evitar resultats que s'allunyin progressivament del que desitgem. Per tal d'ampliar aspectes d'algorísmica o computació es poden consultar, respectivament, (Knuth, 1968 i 1982) i (Press et al., 1989).

Agraïments

Voldria agrair al Dr. Eduard Vives la direcció d'aquest treball així com la seva col·laboració a l'hora de redactar-lo. Expresso també el meu agraïment al Dr. Antoni Planes pels suggeriments fets en llegir el text.

⁸Els generadors actuals més senzills tenen un període de l'ordre de 10^{16} . Així, es dona per fet que els U són tots diferents ja que la probabilitat de tenir-ne dos d'iguals en un mateix grup és pràcticament zero.

⁹Per exemple, si volem simular els resultats obtinguts en llençar una parella de daus.