Normale tal

Matematiklærerdag 2017

Simon Kristensen

Institut for Matematik
Aarhus Universitet

Aarhus Universitet, 24/03/2017

Outline

Hvad er tilfældighed?

- I statistik, sandsynlighedsteori og ikke mindst i programmering er det vigtigt at kunne finde tilfældige tal.
- I min tid (80'erne) genererede min Commodore64 tilfældige tal ud fra clock-frekvensen. Det var noget hø!
- Moderne metoder er bedre, men render alle ind i problemet med faktisk at definere tilfældighed.
- Idag ser vi på en mulig måde ved hjælp af talteori og sandsynlighedsteori.
- Forhåbentlig er der gymnasie-egnet materiale herinde et sted.

Uafhængighed

- Med en vis rimelighed kan man kræve, at elementerne i en tilfældig følge opfører sig som udfald af en *uniformt fordelt* og *uafhængig* følge af stokastiske variable.
- Altså, der skal være lige stor sandsynlighed for, at ethvert tilladt ciffer forekommer på plads n.
- Og cifferfordelingen på plads n må ikke på nogen måde afhænge af cifferfordelingen på de andre pladser.
- Så kan vi gange sandsynligheder sammen.
- Læg mærke til, at jeg på dette tidspunkt ikke siger noget som helst om udfaldsrummet. Det vender vi tilbage til.

Infinite Monkey Theorem

- "En abe, der taster tilfældigt på en skrivemaskine i uendelig lang til, vil på et tidspunkt skrive Shakespears samlede værker (med 100% sandsynlighed)!"
- Antag at der er 65 taster på skrivemaskinen, og at aben skal skrive "årsopgørelse".
- Hvis aben opfører sig uniformt fordelt, rammer den en given tast med sandsynlighed $\frac{1}{65}$.
- Hvis aben opfører sig uafhængigt, kan jeg gange sammen, så sandsynligheden for at ramme "årsopgørelse" er 65^{-12}.
- Sandsynligheden for ikke at ramme "årsopgørelse" er dermed $1 - 65^{-12}$.
- Og dermed er sandsynligheden for ikke at ramme "årsopgørelse" i n forsøg $(1 - 65^{-12})^n \to 0$.

Rigtige aber

- I 2003 forsøgte en gruppe studerende og undervisere fra University of Plymouth med 6 aber i en måned.
- De producerede 5 sider, mest bestående af bogstavet "s".
- Alfa-hannen forsøgte at smadre tastaturet med en sten.
- Alle som en forrettede de deres nødtørft i tastaturet.
- Aber er altså på samme tid nogle svin, hærværks-væsner og rigtig dårlige tilfældighedsgeneratorer.

De store tals stærke lov

- De store tals stærke lov siger, at for uafhængige og identisk fordelte stokastiske variable (X_n), vil gennemsnittet konvergere mod middelværdien næsten sikkert, dvs.

$$\frac{1}{N}\sum_{n=1}^{N} X_n \to \mathbb{E}(X_1) \quad \text{n.s.}$$

- Hvis den uendelige abe var rigtig tilfældig, skulle den altså ikke blot skrive Shakespears samlede en enkelt gang, men uendeligt mange gange.
- Det er denne type tilfældighed, vi leder efter. Men nu ved hjælp af talsystemer.

Hvad er normale tal?

- Et reelt tal er simpelt normalt til base b, hvis ethvert ciffer optræder med frekvens b^{-1} i base b-opskrivningen af tallet.
- Et reelt tal er normalt til base b, hvis enhver blok af cifre af længde n optræder med frekvens b^{-n} i base b-opskrivningen af tallet.
- Et reelt tal er absolut normalt, hvis det er normalt til enhver heltallig base $b \geq 2$.
- Som vi skal se, er normalitet til en enkelt basis et dårligt mål for tilfældighed.

Borels Sætning

- Émile Borel viste i 1909, at Lebesgue næsten alle tal er absolut normale. Vi tager den på tavlen.
- Det er umådeligt svært faktisk at finde sådan et tal. Det kommer vi til.
- Ingen af de klassiske konstanter (π, e, $\log 2$, $\sqrt{2}$ osv.) er vist at være absolut normale, selvom de formodes at være det.

Borels Formodning

- En måde at lave formodninger på i talteorien er, at hvis der ikke er nogen oplagte obstruktioner til et givet udsagn, så er det sikkert sandt.
- Det er klart, at et rationalt tal p/q ikke er normalt til base q. Dermed er rationale tal ikke absolut normale.
- Imidlertid er der ikke nogen oplagt obstruktion for algebraiske tal af højere grad. Borels Formodning siger, at disse er absolut normale.
- Vi er skræmmende langt fra formodningen. Det bedste resultat (af Bugeaud og Adamczewski, med en lille forbedring af Bugeaud og Evertse) siger, at blok-kompleksiteten af et algebraisk irrationalt tal vokser lige lidt hurtigere end lineært (på tavlen).

Uniform fordeling

- En følge $\{x_n\}$ i $[0,1]$ siges at være uniformt fordelt hvis

$$\lim_{N\to\infty} \left| \frac{\#\{n \le N : x_n \in [a,b]\}}{N} - (b-a) \right| = 0$$

  for ethvert interval $[a,b] \subseteq [0,1]$.
- Med andre ord kan følgen bruges til at lave numerisk integration med.
- En sætning af Wall siger, at x er et normalt tal til base b hvis og kun hvis følgen af brøkdele $\{b^n x\}$ er uniformt fordelt.

Champernownes tal

- Som bachelorstuderende i 1933 viste Champernowne at tallet

$$0,1234567891011121314\ldots$$

  er normalt til base 10.
- Jeg er sikker på, at vi alle er enige om, at det er en ret skidt tilfældighedsgenerator!
- Det vides ikke, om tallet er normalt til andre baser.
- Vi tager lige en (ikke historisk korrekt) skitse.

Et mere generelt resultat

Sætning

Lad b en base, lad (c_j) være en voksende følge af heltal, så at for ethvert $\theta > 1$ er $c_N < N^\theta$ for N stor nok. Så er tallet

$$0, (c_1)_b (c_2)_b \dots$$

normalt til base b.

Det følger umiddelbart at konkatenering af primtallene også virker, da $p_n \leq Cn \log(3n)$.

The Hot Spot Lemma

Skriv $x = 0, x_1 x_2 x_3 \dots$ i base b. For en blok D_k af længde k i en base b, lad

$$A_b(D_k, N, x) = \#\{j \leq N - k : x_{j+1} = d_1, \dots x_{j+k} = d_k\}$$

Lemma

Tallet x er normalt til base b hvis og kun hvis der findes et $C > 0$ så at for ethvert $k \geq 1$ og enhver k-blok D_k:

$$\limsup_{N \to \infty} \frac{A_b(D_k, N, x)}{N} \leq \frac{C}{b^k}.$$

Endelige strenge

Definition

Lad $b \geq 2, k \geq 1$ og $\ell \geq 1$ være heltal. Lad $\epsilon > 0$. En streng W på alfabetet $\{0, \dots, b-1\}$ af længde ℓ er (ϵ, k)-normal til base b hvis antallet af forekomster af enhver længde k-streng i W ligger mellem $(b^{-k} - \epsilon)\ell$ og $(b^{-k} + \epsilon)\ell$.
Et positivt heltal c er (ϵ, k)-normalt til base b, hvis strengen $(c)_b$ er det.

De fleste tal er endeligt normale

Lemma

Lad $b, k \in \mathbb{N}$, $b \geq 2$ og lad $\epsilon \in (0, 1/2)$. Så findes $\delta = \delta(b, k)$, så at

$$\#\{n \leq N : n \text{ ikke } (\epsilon, k)\text{-normalt til base } b\} \leq N^{1-\delta\epsilon^2}$$

når N er stor nok.

Coup de grâce

- Da c_n vokser sub-eksponentielt, er der stadig mange (ϵ, k)-normale tal derinde.
- Dette giver god kontrol over optræden af længde k-strenge i starten af det ord, vi kigger på.
- Faktisk så god, at vi kan balancere startblok, ϵ og k ud mod hinanden og anvende the Hot Spot Lemma.
- Dermed bliver Chapernownes tal normalt.

Algoritmer

- I skal da lige have lidt frontforskning!
- Det er svært at afgøre, om et givent tal er absolut normalt, men man kan konstruere et.
- Alan Turing havde en dobbelt-eksponentiel algoritme (upubliceret)
- Sierpiński (1917) havde ligeledes en dobbelt-eksponentiel algoritme.
- I nyere tid findes algoritmer i polynomiel tid, der er ret tæt på den optimale grænse. Becher, Heiber og Slaman (2013) har en algoritme, der spytter de første i cifre (i base b) ud i $O(f(i)i^2)$ elementære operationer, hvor f er en vilkårlig, ikke-aftagende, ubegrænset funktion.

Håbløst a-normale tal

- Et rimeligt spørgsmål er, hvorfor det er så pokkers svært at finde et tal med en egenskab, som næsten alle tal har.
- En del af svaret er, at en nulmængde faktisk kan være ganske stor.
- En mængde er topologisk stor (residual), hvis den indeholder en tæt G_δ-mængde.
- Olsen (2004) viste, at hvis vi ser på følgen af simultane cifferfordelinger i base b op til længde N og kræver at alle mulige sandsynlighedsvektorer skal være akkumulationspunkter for denne følge, så er mængden af sådanne tal residual.

Tak fordi I kom!
Og god weekend!