Upptäckter med datorn — 5

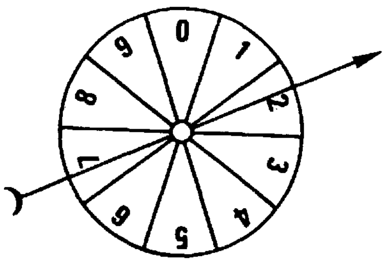
Att pröva och att samla slumptal

LENNART RÅDE

Denna gång tar *Lennart Råde* upp hur man simulerar slumptal. Han visar därefter på metoder att pröva hur väl den använda slumptalsgeneratorn fungerar.

Att bilda slumptal

På många av marknadens miniräknare finns det en "slumpknapp", som vid nedtryckning ger ett decimaltal mellan 0 till 1 till synes helt slumpmässigt. Det är egentligen ganska märkligt att man på detta sätt kan få en miniräknare att härma eller simulera slumpen. I vårt vanligaste programmeringsspråk Basic finns också en "slumpinstruktion" nämligen "RND" (RANDOM). Den får datorn att helt på måfå bilda ett slumptal mellan 0 och 1.

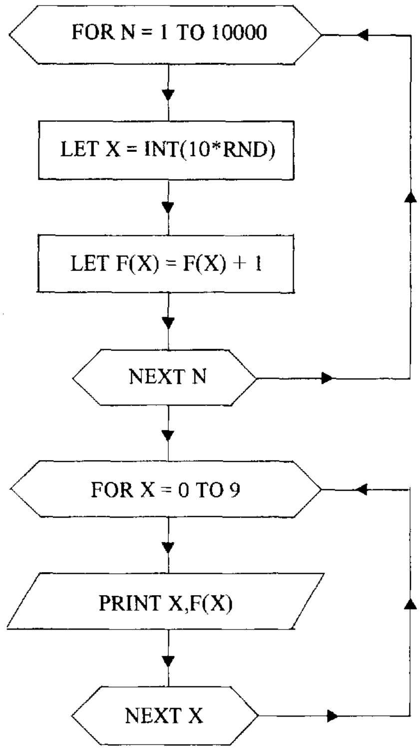


Med slumptal menar man i allmänhet sådana utfall, som man får då man snurrar på lyckohjulet i figuren ovan, dvs då man gör oberoende upprepningar av ett slumpförsök, vid vilket det föreligger tio möjliga utfall 0, 1, 2, ..., 9 med samma sannolikhet 1/10 för vartdera utfallet. En dator eller miniräknare kan ju inte vara slumpmässig på samma sätt som en tärning eller ett lyckohjul. Men det går att utforma program, som får datorn eller miniräknaren att simulera slumpen. Sådana program kallas *slumptalsgeneratorer.* Slumptal används vid alla simuleringar. Det är därför viktigt att man prövar att en använd slumptalsgenerator fungerar väl.

Frekvenstest och partest

Det verkar rimligt att kräva att en slumptalsgenerator i det långa loppet ska ge ungefär samma antal nollor, ettor osv. Detta kan undersökas med ett frekvenstest. Härvid bildar man t ex 10 000 slumptal med den slumptalsgenerator som ska undersökas och bestämmer frekvenserna för de olika utfallen 0, 1, 2, ...

Nedan beskrivs ett program i Basic, som genomför ett sådant frekvenstest för den slumptalsgenerator, som initieras av instruktionen "RND". Lägg märke till att denna slumptalsgenerator bildar slumptal mellan 0 och 1, som sedan med instruktionen "INT (10*RND)" förvandlas till slumptal.

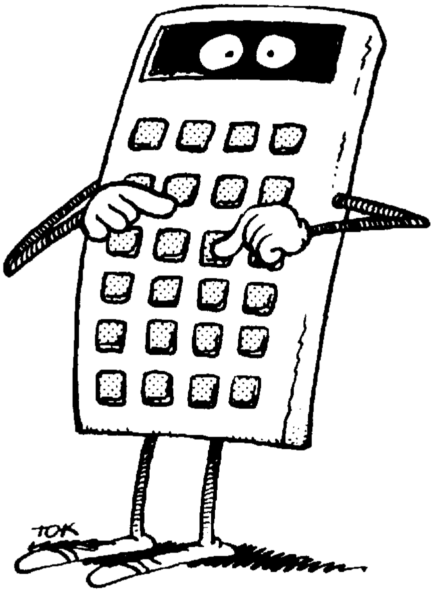


Det är i allmänhet inte tillräckligt att endast genomföra ett frekvenstest av en slumptalsgenerator. Anta t ex att en slumptalsgenerator alltid bildar samma tal två gånger så att den t ex kan tänkas bilda följande "slumptal"

7766220044331l ...

Det är ju då knappast fråga om slumptal men ett frekvenstest kan ändå ge den godkänt. Ett enkelt test vid vilket sådana avvikelser lätt upptäcks är *partestet*. Härvid bildar man slumptalen två i taget och bestämmer frekvenserna för de 100 möjliga utfallen. Jag ger ett exempel i nästa avsnitt.

När man genomför frekvens- eller partest har man behov av att jämföra observerade frekvenser med förväntade frekvenser. Hur stora eller små får avvikelserna vid ett frekvenstest vara för att den prövade slumptalsgeneratorn ska underkännas? En sådan jämförelse kan genomföras med hjälp av statistisk hypotesprövningsteori men vi går inte in på det här.



En miljon slumptal

År 1955 publicerade RAND Cooperation i USA en stor volym med titeln "A Million Random Digits". På sammanlagt 400 sidor ges i denna bok tabeller med tillsammans 1 miljon slumptal. Dessa slumptal bildades med en fysikalisk metod med hjälp av en "elektronisk roulette". Man hade stora svårigheter att få den att fungera tillfredsställande och man fick bl a "omrandomisera" de bildade slumptalen. Man genomförde ett flertal frekvens- och partest. Ett frekvenstest med samtliga slumptal gav följande resultat.

Utfall	Frekvens
0	99802
1	100050
2	100641
3	100311
4	100094
5	100214
6	99942
7	99559
8	100107
9	99280

Följande tabell ger resultaten av ett partest av 50 000 av slumptalen. I tabellen anger X första siffran och Y andra siffran i de bildade paren.

X \ Y	0	1	2	3	4	5	6	7	8	9
0	508	456	509	507	502	489	471	504	488	489
1	510	514	474	514	504	481	496	486	507	527
2	451	523	493	484	502	466	514	506	493	484
3	500	472	476	466	513	478	540	513	530	463
4	513	561	481	485	526	513	485	510	524	511
5	475	490	527	507	493	481	489	512	465	554
6	494	486	491	483	525	504	530	539	513	490
7	508	512	454	498	550	533	516	504	485	520
8	463	503	475	514	520	544	514	491	520	442
9	501	496	536	493	474	504	500	515	461	494

Samla slumptal med dator

Som ett exempel på simulering av ett slumpförsök med hjälp av slumptal kan man betrakta samlande av slumptal tills man fått en fullständig uppsättning av de tio möjliga utfallen. Som bekant har ofta samlare av frimärken, tändsticksetiketter, jultallrikar osv ambitionen att samla till en komplett serie. Det kan vara ganska svårt att förverkliga detta mål. Speciellt svårt kan det vara att få tag på den sista biten. Att samla slumptal tills man fått en fullständig uppsättning är givetvis enklare. Här kan vi överlämna arbetet till en dator. Här är ett exempel på resultat vid ett sådant samlande

23385561116648084110773148286060124711025885579

I detta fall krävdes det att man bildade inte mindre än 47 slumptal tills man fick en fullständig serie.

Vi beskriver inte här i detalj hur man kan utforma ett program, t ex i Basic, som bildar slumptal tills alla de tio möjliga utfallen har erhållits och som bestämmer antalet N av bildade slumptal. Jag vill inte beröva läsaren nöjet att utforma ett sådant program på egen hand.

Följande stam- och bladdiagram visar observerade värden på variabeln N för 30 upprepningar.

1	7899
2	122233445567779
3	00569
4	0011
8	3

Som minst krävdes här att man fick en fullständig uppsättning redan efter det att man bildat 17 slumptal. Som mest krävdes det 83 slumptal.

Om Du har lyckats utforma ett program, som bestämmer antalet N av erforderliga slumptal tills man fått en fullständig serie är det lätt att modifiera detta program så att det bildar ett antal sådana samlingar och bestämmer medelvärdet av de observerade värdena på variabeln N, dvs den genomsnittliga storleken av en sådan samling. Författaren lät en fickdator bilda inte mindre än 1 000 sådana samlingar, varvid den genomsnittliga storleken blev 29,804.

Man kan härleda en teoretisk formel för det förväntade antalet. Man kan nämligen visa att det är

$$10\left(1+\frac{1}{2}+\frac{1}{3}+\ldots+\frac{1}{10}\right) \approx 29{,}2897$$

Datoraktiviteter

1. Utforma program för frekvens- och partest av en slumptalsgenerator. Använd programmen för att pröva några olika slumptalsgeneratorer. Pröva också om man som slumptal X mellan 0 och 1 kan använda $X=Z^2$ och $Z=$RND.
2. Utforma ett program som bildar slumptal tills alla de tio möjliga utfallen har erhållits. Programmet ska bestämma antalet N av totala antalet slumptal. Utvidga programmet så att det bestämmer medelvärdet av N i t ex 1 000 upprepningar.
3. Upprepa föregående övning men anta nu att man ska bilda slumptal tills man fått alla de möjliga utfallen två gånger.
4. När man bildat slumptal tills alla de olika möjligheterna har erhållits kommer ett visst antal utfall att ha erhållits exakt en gång. Låt E vara antalet utfall, som erhållits exakt en gång. Uppskatta ett genomsnittsvärde på E genom simulering.
5. I stället för slumptal kan man allmännare betrakta utfall vid ett försök med utfallen 0, 1, ..., N-1 och med samma sannolikhet $1/N$ för vardera utfallet. Upprepa övningarna ovan för några olika värden på N.

Hör gärna av Dig med kommentarer, resultat och program till datoraktiviteterna ovan. Är Du lärare så visa gärna denna uppsats för datorintresserade elever. Adressen till författaren är

Lennart Råde
Matematiska Institutionen,
Chalmers Tekniska Högskola,
412 96 Göteborg.

Du kan läsa mer om slumptal och simulering i L Råde, *Simulering*, Studentlitteratur 1987. I denna bok ges också mer utförliga litteraturhänvisningar.