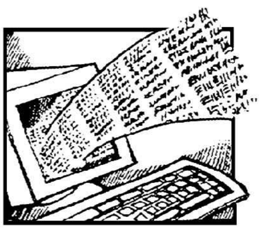

Pietarin paradoksi ja tietokonesimulaatio

Jani Isohanni
Ohjelmistotekniikan laitos, Tampereen teknillinen yliopisto
jani.isohanni@tut.fi

Pietarissa 1700-luvulla vaikuttanut sveitsiläinen matemaatikko Nicolaus Bernoulli esitti kuvitteellisen uhkapelin, jossa kolikkoa heitetään, kunnes saadaan kruuna, ja pelaaja voittaa rahaa sen mukaan, monennellako heitolla kruuna tulee. Jos kruuna tulee heti ensimmäisellä heitolla, voittaa pelaaja kaksi euroa. Jos toisella, niin hän voittaa neljä euroa jne. Yleisesti, jos kruuna tulee n:nnellä heitolla, pelaaja voittaa 2^n euroa, ja klaavalla peli jatkuu mahdollisesti rajattoman kauan.

Tämä peliasetelma osoittautuu oivaksi esimerkiksi siitä, miten tietokonesimulaatioilla voidaan saada näennäisesti ristiriitaisen oloisia tuloksia.

Voiton odotusarvo

Olemme kiinnostuneita, miten paljon pelaajan kannattaa maksaa pelatakseen kuvattua peliä. Lähestytään asiaa laskemalla ensin voiton odotusarvo. Jotta voisimme puhua asioista täsmällisesti, niin sovitaan muutamasta tarvittavasta käsitteestä.

Pelikierros tarkoittaa kolikon heittämistä, kunnes saadaan kruuna, *peli* on jokin valittu määrä pelikierroksia, ja *heitto* on yksi kolikon heitto. Nyt peli koostuu valitusta määrästä pelikierroksia, jotka puolestaan kestävät korkeintaan valitun määrän heittoja.

Panos on yhdellä pelikierroksella asetettu panos, *voitto* yhdeltä kierrokselta voitettu rahamäärä, ja *voittosumma* on yhden pelin kaikilta pelikierroksilta yhteensä voitettu summa.

Tarkastellaan ensin heittojen maksimimäärän suhteen rajoitettua tapausta.

Jos pelikierros päättyy ensimmäiseen heittoon, niin vain kruunalla voittaa ja voiton todennäköisyys on 0,5 voiton ollessa 2 euroa. Tästä voidaan laskea, että voiton odotusarvo yhdelle kierrokselle on $0{,}5 \cdot 2 = 1$, ja päätellään, että yhdeltä pelikierrokselta voittaa keskimäärin yhden euron eli pelaajan kannattaa osallistua peliin kierrokseksi alle eurolla.

Jos pelikerros venytetään maksimissaan kahteen peräkkäiseen heittoon, on voiton odotusarvo $0{,}5 \cdot 2 + 0{,}25 \cdot 4 = 2$ euroa eli pelaajan kannattaa ostaa yksi pelikierros alle kahdella eurolla. Tämä yleistyy suoraviivaisesti tapaukseen, jossa sallitaan korkeintaan n perättäistä heittoa, ja voiton odotusarvoksi saadaan

$$\sum_{i=1}^{n} \frac{1}{2^n} 2^n = \sum_{i=1}^{n} 1 = n.$$

Siis mitä enemmän perättäisiä heittoja kierroksella sallitaan, sitä enemmän pelaaja voi odottaa voittavansa yhdellä pelikierroksella. Jos heittoja sallitaan rajattoman paljon, niin periaatteessa pelaajan kannattaa pelata yksi kierros millä tahansa äärellisellä summalla, koska voiton odotusarvo kasvaa äärettömäksi.

Simulaatio

Edellisen pohjalta tuntuu loogiselta odottaa, että jos simuloimme peliä tietokoneella, niin peli kääntyy pelaajalle voitolliseksi, kunhan sisällytämme simulaatioon tarpeeksi monta pelikierrosta.

Luodaan ohjelma, joka simuloi halutun määrän pelikierroksia, joissa kussakin voi olla rajaton määrä heittoja[1]. Ohjelma tulostaa pelatuista pelikierroksista keskimääräisen voiton kierrokselta eli paljonko peli keskimäärin palauttaa pelatulta kierrokselta, sekä pelatun pelin voittosumman. Yhden pelikierroksen osalta voiton odotusarvo on ääretön, joten keskimääräisen toteutuneen voiton voisi olettaa kipuavan kohti ääretöntä, kunhan lisäämme pelikierroksia tarpeeksi. Koska nyt kierroksia voi tulla rajattoman paljon, kannattaa äskeisten laskujen pohjalta pelaajan lähteä peliin mukaan millä tahansa äärellisellä panoksella kierrosta kohti.

Ohessa on taulukoituna lyhyellä C-ohjelmalla tehdyn simulaation tuloksia. Viimeisessä sarakkeessa on pelaajan nettovoitto koko pelistä, kun oletetaan, että hän on maksanut yhdestä pelikierroksesta 20 euroa.

Kierroksia	Keskim. voitto/kierros	Nettovoitto
10	2,8	-172
10^2	4,62	n. -1540
10^3	12,91	n. -7000
10^4	8,73	n. -130000
10^5	19,29	n. -71000
10^6	12,83	n. $-7{,}2$M
10^7	11,40	n. -86M
10^8	18,97	n. -100M

Tulokset vaikuttavat hieman epäintuitiivisilta. Laskujen mukaan pelistä kannattaa maksaa mikä tahansa äärellinen summa rahaa per kierros, mutta jo kahdenkymmenen euron panoksella jää tappiolle sitä enemmän mitä pidempään peliä pelaa. Voisi odottaa, että tällaisilla pelimäärillä voittosumma kasvaisi nähtävästi, mutta esimerkiksi sadan miljoonan pelatun kierroksen osalta jäädään tappiolle peräti sata miljoonaa. Tämän kaiken päälle keskimääräinen voitto vaihtelee aika paljon. Peli ei vaikuta oikein kannattavalta. Onko ohjelmassamme vikaa? Tai alun laskuissa?

Lähdetään oletuksesta, että ohjelmassa tai sen ajossa ei ole mitään outoa eli simulaation tulokset sinällään ovat kirjautuneet oikein. Tällöin voidaan päätellä, että viimeisessä pelissä ei ollut yhtään yli kolmenkymmenen kruunun sarjaa, koska jos olisi ollut, niin siitä saatu voitto yksinään olisi isompi kuin nyt saatu koko voittosumma (koska $2^{31} \approx 2{,}15 \cdot 10^9 > 100000000 \cdot 18{,}97 \approx 1{,}9 \cdot 10^9$).

Tämä antaa syytä pohtia, kuinka todennäköistä ylipäätään on saada tietyn pituisia heittosarjoja, kun pelissä pelataan jokin määrä kierroksia. Todennäköisyys saada yli k:n heiton kierros on

$$\lim_{N\to\infty} \sum_{i=k+1}^{N} (1/2)^i = ((1/2)^{k+1})/(1-1/2) = (1/2)^k.$$

Suoraan komplementin todennäköisyydestä seuraa, että todennäköisyys, ettei pelikierroksella saada yli k:ta heittoa, on $1-(1/2)^k$, joten jos pelataan n pelikierrosta, niin todennäköisyys p, ettei yhdelläkään kierroksella tule yli k:ta heittoa, on

$$p = (1-(1/2)^k)^n. \qquad (1)$$

Lasketaan nyt, mikä on pisin esiintyvä heittosarja todennäköisyydellä p, kun pelikierrosten lukumäärä n tunnetaan. Tämä edellyttää, että ratkaisemme edellisen yhtälön k:n suhteen, ja pienellä algebrallisella pyörittelyllä saammekin

$$k = \frac{\ln(1-\sqrt[n]{p})}{\ln 0{,}5}.$$

Asetetaan nyt $p = 0{,}95$ ja lasketaan k edellisessä simulaatiossa ajetuille kierroslukumäärille, jolloin saadaan oheinen taulukko.

Pelikierroksia n	Pisin heittosarja k tn:llä $p = 0{,}95$
10	7,6
10^2	10,9
10^3	14,3
10^4	17,6
10^5	20,9
10^6	24,2
10^7	27,5
10^8	30,9

Huomaamme, että vaikka pelaisimme sata miljoonaa kierrosta, niin todennäköisyydellä 0,95 jokaisella kierroksella jäädään alle 31 peräkkäisen heiton. Tästä seuraa, että todennäköisyydellä 0,95 voittosumma sadan miljoonan kierroksen pelissä on alle 2^{31}, joka tekee noin $2^{31}/10^8 \approx 21{,}5$ euroa per kierros. Siis todennäköisyydellä 0,95 jää pelissä tappiolle, jos sijoittaa yli 21,5 euroa kierrosta kohti. Tätä vasten simulaation tulos rupeaakin tuntumaan järkevältä.

Simulaation aluksi paradoksaaliselta vaikuttaneet tulokset osoittautuvat näin lähemmässä tarkastelussa loogisiksi. Peli ei vaikutakaan enää (ainakaan ajankäytöllisesti) kovinkaan houkuttelevalta. Vaikka siinä on

[1] Periaatteessa rajaton määrä. Käytännössä tietokoneella laskettaessa on käytössä aina rajallinen määrä resursseja (esim. aika). Tässä tapauksessa nämä rajat eivät kuitenkaan tule lähellekään ja on mielekästä puhua rajattomasta määrästä kierroksia.

mahdollista voittaa huimia summia, niiden todennäköisyys on hyvin pieni. Odotusarvoa laskettaessa tämä jää helposti huomaamatta, koska näille epätodennäköisille pelikierroksille kasattu voitto on vastaavasti hyvin suuri, jolloin ne odotusarvoa laskiessa tietyssä mielessä kumoavat toisensa.

Lopuksi

Alussa esitetyt laskut antoivat kuvan, että Bernoullin esittämää peliä kannattaa pelata millä tahansa äärellisellä panoksella. Kuitenkin tietokonesimulaatiossa kävi ilmi, että jo pienellä kierroskohtaisella panoksella tappiota kertyi huimasti. Tämä on hyvä esimerkki siitä, miten simulointi ja puhtaasti teoreettinen tarkastelu tukevat toisiaan; simulaatiossa saatu hieman yllättävä tulos pakotti miettimään tilannetta tarkemmin. Tämän tarkastelun jälkeen kävikin ilmi, että pelkän odotusarvon laskeminen antoi pelistä ja sen luonteesta harhaanjohtavan kuvan.