Lottorivin numeroiden summa

Pentti Haukkanen
Matematiikan, tilastotieteen ja filosofian laitos, Tampereen yliopisto
pentti.haukkanen@uta.fi

Järjestellessäni äsken vanhoja Mathematica-kansioitani törmäsin tiedostoon, jossa ratkaistaan pieni lottotehtävä. Muutama vuosi sitten nimittäin keskusteleltiin julkisuudessa (tosin aivan vähäisessä määrin) siitä, voiko lottorivien numeroiden summien avulla ennustaa arvottavaa lottoriviä. Lottorivin numerot tarkoittavat lottorivin seitsemää kokonaislukua väliltä $[1, 39]$, jotka ilmoitetaan suuruusjärjestyksessä. Jotkut olivat huomanneet, että kaikki lottorivin numeroiden summat eivät ole yhtä todennäköisiä. Esimerkiksi, vain yhden lottorivin (rivin 1, 2, 3, 4, 5, 6, 7) numeroiden summa on 28, kun taas summa 30 saadaan kahdesta lottorivistä (1, 2, 3, 4, 5, 6, 9 ja 1, 2, 3, 4, 5, 7, 8). Jotkut maallikot tekivät tästä johtopäätöksen, että kannattaa esimerkiksi ennemmin lotota riviä, jonka numeroiden summa on 30, kuin riviä, jonka numeroiden summa on 28. Todennäköisyyslaskennan alkeet ymmärtävä tietää, että kaikki rivit ovat yhtä todennäköisiä, siis esimerkiksi rivi 1, 2, 3, 4, 5, 6, 7 on yhtä todennäköinen kuin rivi 1, 2, 3, 4, 5, 6, 9. Toki on todennäköisempää, että lottorivin numeroiden summa on 30 kuin, että se on 28, mutta lotossa ei veikatakaan numeroiden summaa vaan itse numeroita. (Huomattakoon, että tarkka lottoaja ei kuitenkaan veikkaa riviä 1, 2, 3, 4, 5, 6, 7, koska sitä veikkaa niin moni, että osuessaan kohdalle se tuottaisi yhdelle voittajalle ilmeisesti tavanomaista pienemmän potin; mutta tämä onkin jo eri asia.)

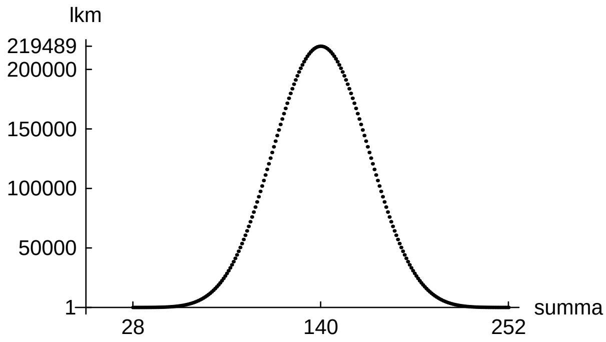
Laskin silloin muutama vuosi sitten jollekulle asiaa pohtineelle, miten lottorivien numeroiden summat jakautuvat. Tein laskutehtävän generoivilla funktioilla seuraavalla tavalla. Tarkastellaan generoivaa funktiota

$$g(u, t, M) = \prod_{i=1}^{M} (1 + u\, t^i),$$

joka laskee joukon $\{1, 2, \ldots, M\}$ osajoukkoja niin, että symbolin u potenssi kertoo, kuinka monta alkiota osajoukossa on, ja symbolin t potenssi kertoo näiden alkioiden summan. Kirjoitetaan $g(u, t, M)$ nyt muotoon

$$g(u, t, M) = \sum_{m=1}^{M} \sum_{n=1}^{s(m,M)} a_{m,n} u^m t^n,$$

jolloin siis kerroin $a_{m,n}$ ilmaisee, kuinka monta kappaletta on sellaisia m-alkioisia joukon $\{1, 2, \ldots, M\}$ osajoukkoja, joiden alkioiden summa on n. (Tässä $s(m, M)$ on suurin mahdollinen summan arvo, ts. $s(m, M) = [M - (m - 1)] + [M - (m - 2)] + \cdots + [M - 1] + M$.) Koska lottorivissä valitaan 7 alkiota joukosta $\{1, 2, \ldots, 39\}$, niin generoivassa funktiossa $g(u, t, 39)$ kerroin $a_{7,n}$ ilmaisee, kuinka monta kappaletta on sellaisia lottorivejä, joiden numeroiden summa on n. Generoivien funktioiden laskenta on helppoa Mathematicalla. Kuvassa 1 on lottorivien numeroiden summan jakauma. Yleisin summan arvo on 140, joka on 219 489 eri rivin numeroiden summa. Kaikkiaan erilaisia lottorivejä on $\binom{39}{7}$ eli 15 380 937 kappaletta. Lottorivin numeroiden summa on 140 todennäköisyydellä 0.0142702; siis todennäköisyys on selvästi yli prosentin. Summan arvo on esimerkiksi välillä $137, \ldots, 143$ noin 10 prosentin todennäköisyydellä. Voisi ehkä sanoa, että keskimmäiset arvot ovat "yllättävän" todennäköisiä.



KUVA 1: Lottorivien numeroiden summan jakauma.

Lotto alkoi vuonna 1970 ja ensimmäinen rivi arvottiin 3.1.1971. Kaikki arvotut rivit löytyvät Veikkauksen nettisivuilta. Kiinnostunut lukija voi testata, mitenkä lottorivien numeroiden summat ovat jakautuneet tähän mennessä.

Generoivat funktiot kuuluvat lähinnä kombinatoriikan alaan. Esimerkkeinä alan oppikirjoista mainittakoot [1, 3]. Myös diskreetin matematiikan yleisteoksissa esitellään usein generoivat funktiot, ks. esim. [2].

Viitteet

[1] R. A. Brualdi, *Introductory Combinatorics.* Second edition. North-Holland Publishing Co., New York, 1992.

[2] R. P. Grimaldi, *Discrete and Combinatorial Mathematics.* Fourth edition. Addison Wesley Publishing Co., 1998.

[3] A. Tucker, *Applied Combinatorics.* Third edition. John Wiley & Sons, Inc., New York, 1995.