TILASTOTIETEEN JATKOKURSSI, 5+5 OP, 15.1.–23.4.2015. Kirjallisuus: Ilkka Mellinin Johdatus tilastotieteeseen, 2. kirja. Luennoi: yliopistonlehtori Pekka Pere.

Galtonin kone

Francis Galton (1877) kuvasi sittemmin Galtonin koneena (*Galton's machine*, *Galton board*, *Galton box*, *Quincunx* tai *bean machine*) tunnetun mekaanisen laitteen. Sillä voidaan havainnollistaa hyvin konkreettisella tavalla keskeistä raja-arvolausetta. Ohessa on Galtonin luonnos koneesta (1889, 63), kuva Galtonille 1873 tehdystä koneesta ja kuva uudemmasta suuresta Galtonin koneesta.[1]

Koneen yläosasta tippuu suppilosta kuulia, jotka päätyvät koneen alaosassa oleviin numeroituihin ("0","1",...,"n") laareihin ($n+1$ kappaletta). Välissä on pyramidin muodossa n riviä pinnejä (i. rivillä i pinniä, $i = 1, \dots, n$) niin, että 1. pinni on suppilon suun keskipisteen alapuolella.

Kuula, joka on päätynyt laariin "k" ($k = 0, \dots, n$), on pompannut k kertaa oikealle todennäköisyydellä p ja $n-k$ kertaa vasemmalle todennäköisyydellä $1-p$. Kunkin laariin "k" johtavan polun todennäköisyys on $p^k(1-p)^{n-k}$.

Jos $p = 1 - p = 0,5$ — kuten alla oletetaan — niin jokaisen polun mihin tahansa laariin todennäköisyys on $(0,5)^n$. Keskimmäisiin laareihin päätyy kuitenkin enemmän kuulia kuin laitimmaisiin, koska edellisiin johtaa useampia polkuja mutta laitimmaisiin vain yksi.

Kuinka monta polkua johtaa laariin "k"? Polkujen lukumäärä voidaan päätellä seuraavasti.

Tarkastellaan oransseja ja vihreitä alkioita ("o" ja "v"), joita on k ja $n-k$ kappaletta (yhteensä n alkiota). Kuinka moneen erilaiseen n alkiota sisältävään jonoon ne voidaan järjestää? Merkitään näiden jonojen lukumäärää N:llä. Mikäli voitaisiin erotella o-alkiot toisistaan, olisi jonoja $N \times k!$ kappaletta, sillä o-alkiot voidaan järjestää $k!$ eri tavalla yhdessä jonossa (kirjan s. 12; lause 1). Mikäli lisäksi voitaisiin erotella v-alkiotkin, olisi erilaisia jonoja $N \times k! \times (n-k)!$ kappaletta, sillä v-alkiot voidaan järjestää $(n-k)!$ eri tavalla yhdessä jonossa. Tällöin pystyttäisiin erottelemaan kaikki alkiot, jolloin erilaisia jonoja on $n!$ (kirjan s. 12; lause 1). Näin ollen täytyy päteä

$$N \times k! \times (n-k)! = n!$$

eli

$$N = \frac{n!}{k! \times (n-k)!} = \binom{n}{k}.$$

Korvataan o- ja v-alkiot edellä pompuilla oikealle ja vasemmalle. Polku laariin "k" koostuu k pompusta oikealle ja $n-k$ pompusta vasemmalle. Päättelyn edellä mukaan tällaisia järjestyksiä eli polkuja laariin "k" on $\binom{n}{k}$ kappaletta.

[1]Ensimmäinen kuva on Stephen M. Stiglerin ottama. Kaksi ensimmäistä kuvaa ovat artikkelista Burnett (2009). Galtonin koneen R-kielisen emulaattorin ovat tehneet Xie, Yu ja O'Rourke (2013). Koneen "parannettu" versio on patentoitu Yhdysvalloissa 1990 (`http://www.google.nl/patents/US4900255`; viitattu 29.1.2014).

Polut ovat toisensa poissulkevia, ja jokainen polku on yhtä todennäköinen. Todennäköisyys, että kuula päätyy laariin "k", saadaan siten summaamalla kaikkien laariin "k" johtavien polkujen todennäköisyys eli kertomalla yhden polun todennäköisyys polkujen lukumäärällä $\binom{n}{k}$ $(k = 0, \ldots, n)$:

$$\begin{aligned} \mathsf{P}(\text{kuula päätyy laariin "}k\text{"}) &= p^k(1-p)^{n-k} + \ldots + p^k(1-p)^{n-k} \\ &= \binom{n}{k} p^k(1-p)^{n-k}. \end{aligned}$$

Kuulan päätyminen laariin "k" noudattaa binomijakaumaa. Lisäämällä luonteva oletus $p = 0{,}5$ todennäköisyydeksi saadaan

$$\begin{aligned} \mathsf{P}(\text{kuula päätyy laariin "}k\text{"}) &= \binom{n}{k}(0,5)^k(1-0,5)^{n-k} \\ &= \binom{n}{k}(0,5)^n. \end{aligned}$$

Galtonin kone osoittaa konkreettisesti, kuinka kuulien lukumäärän kasvaessa laareihin muodostuu normaalijakauman kuvio, vaikka jakauma on binomijakauma. Normaalijakauma approksimoi binomijakaumaa. Rakennetut Galtonin koneet havainnollistavat tilannetta $p = 0{,}5$ (kuula pomppaa samalla todennäköisyydellä vasemmalle tai oikealle), mutta approksimaatio toimii muillakin p:n arvoilla.

Lähteet

Burnett, D. Graham (2009): Games of Chance. *Cabinet*. `Http://www.cabinetmagazine.org/issues/34/burnett.php` (viitattu 5.2.2014).

Galton, Francis (1877): Typical Laws of Heredity. *Nature*, 15, 492–495, 512–514 ja 532–533.

Galton, Francis (1889): *Natural Inheritance*. McMillan, Lontoo.

Stigler, Stephen M. (1986): *The History of Statistics. The Measurement of Uncertainty before 1900*. Harvard University Press, Cambridge MA.

Xie, Yihui, Lijia Yu ja Keith O'Rourke (2013): Demonstration of the Quincunx (Bean Machine/Galton Box). `Http://www.rforge.net/doc/packages/animation/quincunx.html` (viitattu 10.2.2015.)