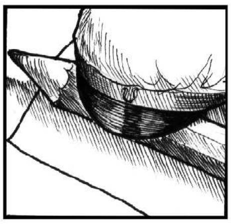

Millä todennäköisyydellä satunnainen luku on neljällä jaollinen?

Juho Niemelä
Kustannustoimittaja
Tammi Oppimateriaalit

Tämä pohdiskelu juontaa juurensa syksyn 2009 pitkän matematiikan ylioppilastehtävään numero 7, jota tarkasteltiin lukion *Pitkä Sigma* -kirjasarjan osaa *Todennäköisyys ja tilastot* (MAA6) tehtäessä:

> *A, B, C ja D aikovat jakaa keskenään korillisen omenoita siten, että kukin vuorollaan ottaa aina yhden omenan. Korissa olevien omenoiden lukumäärää ei tiedetä. A ehdottaa vedonlyöntiä: jos jokaiselle tulee yhtä monta omenaa, A maksaa kolmelle muulle kullekin 50 euroa. Muussa tapauksessa kukin kolmesta maksaa A:lle 25 euroa. Laske A:n saaman rahamäärän odotusarvo.*

Jotta odotusarvon saa laskettua, on ensin ratkaistava, millä todennäköisyydellä omenoiden määrä on neljällä jaollinen. Omenoiden lukumäärää "ei tiedetä", joten oletus lienee, että omenoita voi olla mikä tahansa määrä, vähintään nolla tietysti.

Jos mahdollinen omenoiden määrä olisi rajattu esimerkiksi välille 1–100 ja lisäksi tiedettäisiin, että jokainen määrä on yhtä todennäköinen, neljällä jaollisen määrän todennäköisyys voitaisiin laskea ongelmitta. Suotuisia tapauksia ovat määrät $4, 8, \ldots, 100$ $(= 4 \cdot 25)$, joita on 25, joten todennäköisyys on $\frac{25}{100} = \frac{1}{4}$.

Entä jos omenoita voikin olla mikä tahansa määrä? Mahdollisuuksia on äärettömän monta, joista suotuisia tapauksia ovat kaikki neljällä jaolliset luvut, joita on myös äärettömän paljon. Äärellisen tilanteen tapaan yrittämällä todennäköisyydeksi tulisi $\frac{\infty}{\infty}$ eli ei mitään järkevää.

Yritetään kiertää ongelmaa. Ajatellaan, että olipa omenoiden lukumäärä mikä tahansa, niin neljällä jaettaessa jakojäännökseksi tulee varmasti $0, 1, 2$ tai 3. Ovatko nämä jakojäännökset yhtä todennäköisiä? Jos ovat, niin neljällä jaollisen luvun (eli jakojäännöksen 0) todennäköisyys on selvästi $\frac{1}{4}$. Voisi ajatella, että kyllä kai jakojäännökset ovat yhtä todennäköisiä, koska luvut 0–3 ovat keskenään yhtä todennäköisiä ja loput peräkkäiset nelikot (4–7, 8–11 jne.) ovat olennaisesti täysin samanlaisia kuin nelikko 0–3: jokaisessa nelikossa on yksi kutakin jakojäännöstä edustava luku. Toisaalta matematiikassa ei mielellään haluta sanoa "kai". Täytyy siis yrittää vielä keksiä jokin raudanluja perustelu.

Oletetaan aluksi, että korissa voi olla omenia 0–n ja että jokainen näistä määristä on yhtä todennäköinen. Luvun n jakojäännös neljällä jaettaessa on jälleen 0, 1, 2 tai 3, joten voidaan kirjoittaa $n = 4m + k$, jossa $k = 0, 1, 2$ tai 3. Suotuisia tapauksia ovat $0, 4, 8, \ldots, 4m$, ja niiden lukumäärä on $m + 1$. Klassisen todennäköisyyden mukaan neljällä jaollisen luvun todennäköisyys on nyt

$$\frac{m+1}{n+1} = \frac{m+1}{4m+k+1} = \frac{1+\frac{1}{m}}{4+\frac{k+1}{m}}.$$

Kun n kasvaa rajatta, niin myös m kasvaa rajatta. Luku k vaihtelee mutta pysyy koko ajan rajoissa 0–3. Näin ollen osoittaja lähestyy lukua 1 ja nimittäjä lukua 4, joten osamäärä lähestyy lukua $\frac{1}{4}$. Joko tämä riittää osoittamaan, että todennäköisyys on $\frac{1}{4}$?

Joillekin riittää, joillekin ei. Edellä laskettiin, kuinka suuri on neljällä jaollisten lukujen suhteellinen osuus luvuista 0–n, ja havaittiin suhteellisen osuuden lähestyvän neljäsosaa. Entä jos neljällä jaollisten lukujen mukaan otetaan kaikki luvut yhdestä miljoonaan, mikä on näin saatavan ”paljon isomman” joukon todennäköisyys? Itse asiassa se on jälleen neljäsosa, jos käytetään samanlaista raja-arvoperustelua: lukujen 0–n joukossa on uuden, isomman joukon lukuja korkeintaan miljoona enemmän kuin alkuperäisessä joukossa, ja kun miljoona jaetaan luvulla n ja annetaan n:n kasvaa rajatta, osamäärä lähestyy nollaa. Lisäys ei muuta raja-arvoa, jos lisäys on niin pieni, että sen suhteellinen osuus lähestyy nollaa. Olisi siis yhtä todennäköistä, että poikien korissa on omenia mikä tahansa määrä yhdestä miljoonaan tai jokin neljällä jaollinen määrä kuin että korissa on pelkästään jokin neljällä jaollinen määrä. Sama todennäköisyys saataisiin myös silloin, kun neljällä jaollisiin lisättäisiin vaikkapa kaikki luvun 3 potenssit, sillä niidenkin suhteellinen osuus (noin $\log_3 n/n$) lähestyy nollaa. Ei kuulosta kovin hyvältä.

Tässä kierretään kehää ongelman ympärillä, kunnes lopulta huomataan, että ongelmalla ei ole kunnollista ratkaisua. Selitys on se, että luonnollisten lukujen joukossa ei yksinkertaisesti ole olemassa kaivatunlaista todennäköisyyden mittaria. Haluaisimme, että koko luonnollisten lukujen joukon todennäköisyys on 1. Lisäksi haluaisimme, että jokainen luonnollinen luku on yhtä todennäköinen. Ja vielä haluaisimme, että joukon todennäköisyys saadaan laskemalla yhteen osien todennäköisyydet. Jos jokaisen luonnollisen luvun todennäköisyys on p, niin pitäisi siis olla

$$\begin{aligned} P(\mathbb{N}) &= P(0) + P(1) + P(2) + \cdots \\ &= p + p + p + \cdots = \begin{cases} 0, & \text{jos } p = 0 \\ \infty, & \text{jos } p > 0, \end{cases} \end{aligned}$$

joten ehto $P(\mathbb{N}) = 1$ ei tällöin voi toteutua.

Miten tehtävä pitäisi siis ymmärtää ja ratkaista? Ratkaisussa täytyy olettaa, että jakojäännökset $0, 1, 2$ ja 3 ovat yhtä todennäköisiä. Edellä nähtiin, että ei ole olemassa hyvää mittaria sen toteamiseksi, että luonnollisten lukujen joukon osajoukot ovat yhtä todennäköisiä, joten oletuksen järkevyys täytyy perustella jollakin muulla tavalla. Voitaneen vedota jonkinlaiseen symmetriaan, mutta loppujen lopuksi kysymys on vain sopimuksesta, joka joko hyväksytään lähtökohdaksi tai sitten ei. Perimmäisiä lähtökohtia ei toki yleensäkään matematiikassa edes yritetä perustella (koska niiden perusteleminen on matematiikan keinoin mahdotonta), mutta ehdottomasti ne täytyy kertoa lukijalle, ellei lukijan voida olettaa niitä ilman mainitsemistakin tietävän.

Tämä ylioppilastehtävä ei ollut onnistunein mahdollinen. Ongelmat olisivat poistuneet, jos tehtävänannossa olisi mainittu, että omenoita voi olla esimerkiksi 1–100 ja että jokainen määrä on yhtä todennäköinen. Eihän nimittäin siitä, että omenien määrää ei tiedetä, suinkaan seuraa, että jokainen määrä on yhtä todennäköinen. Voisihan olla, että omenakoreja täyttää vaikkapa kone, joka syöttää peräjälkeen 100 omenaa, joista kukin osuu koriin 95 %:n todennäköisyydellä. Tällöin omenoiden määrä noudattaa binomijakaumaa eikä tasaista jakaumaa. Jos omenien määrää ei haluta rajata, mikä toisaalta vastaa hyvin huonosti todellisuutta (omenakoreihin tuskin mahtuu ainakaan kovin monta tuhatta omenaa), täytyisi kertoa jo tehtävänannossa, että kaikki jakojäännökset oletetaan yhtä todennäköisiksi. Tällöin kuitenkin tehtävän luonne muuttuisi olennaisesti, joten omenien määrän rajaaminen on selvästi parempi vaihtoehto. Sen sijaan on mieletöntä sanoa, että omenoita voi olla mikä tahansa määrä, joista jokainen on yhtä todennäköinen, koska luonnollisten lukujen joukossa ei ole todennäköisyyskäsitettä tällaisen sanonnan perustaksi.