Pitkä matematiikka, todennäköisyys ja tilastot

Satu Juhala, Petri Juutinen, Anni Laitinen, Erkki Luoma-Aho, Terhi Raittila ja *Tommi Tikka* : **Juuri 10. Todennäköisyys ja tilastot.** 206 s. Otava 2018. Hinta tammikuussa 2019 eri verkkokaupoissa 20,80 – 27,75 euroa.

Paavo Heiskanen, Päivi Kaakinen, Jukka Lehtonen, Mika Leikas ja *Jorma Tahvanainen*: **Tekijä. Pitkä matematiikka 10. Todennäköisyys ja tilastot.** 202 s. Sanoma Pro 2018. Hinta tammikuussa 2019 eri verkkokaupoissa 21,10 – 29,55 euroa.

Lukion pitkän matematiikan kymmenenteen kurssiin on pakattu paljon asiaa. Opetussuunnitelma luettelee peräti yhdeksän keskeistä sisältöä: "diskreetti ja jatkuva tilastollinen jakauma, jakauman tunnusluvut, klassinen ja tilastollinen todennäköisyys, kombinatoriikka, todennäköisyyksien laskusäännöt, diskreetti ja jatkuva todennäköisyysjakauma, diskreetin jakauman odotusarvo, normaalijakauma". Tarkasteltavat oppikirjatkin ovat sarjojensa edellisiä osia paksumpia. Lievästi sivumäärää kasvattavat näissä niin kuin sarjojen aikaisemmissakin osissa kaiketi vain kauneuden vuoksi lisätyt kuvat, **Tekijässä** sinne tänne sijoitellut, **Juuressa** kunkin pääluvun aloitussivun peittävät.

Kymmenennen kurssin sisältö on osaksi sitä arkilaskentoa, jota monesti peräänkuulutetaan: numeerisen ja graafisen tilastoesityksen perusteita. Opetussuunnitelma ja kirjat jättävät kuitenkin – ilmeisesti ajan ja tilan puutteen vuoksi – asian harmittavantuntuisesti kesken. Aineistosta estimoidun parametrin luottamusvälistä ei sanota mitään, vaikka tämä käsite tulee esimerkiksi mielipidetutkimusten yhteydessä tiedotusvälineissäkin tavan takaa vastaan.

Kilpailevat kirjasarjat esittelevät asiat jälleen hiukan eri järjestyksessä. Kun oppimäärä on pilkottu lyhyehköiksi kursseiksi, niin tästä ei varmaan aiheudu ongelmia koulua ja ehkä kirjasarjaa vaihtavalle oppilaalle.

Tekijä ei noudata opetussuunnitelman antamaa järjestystä, vaan lähtee aluksi katselemaan klassista todennäköisyyttä. "Tapahtuma" on jokin joukko, jonka alkiot ovat alkeistapahtumia. Lukija saattaa kyllä joutua ymmälleen, kun kerrotaan, että tapahtuma "noppaa heitettäessä saadaan silmäluku, joka on ≥ 9" onkin tyhjä joukko. Todennäköisyys määritellään tietysti suotuisien alkeistapahtumien lukumäärän osuutena kaikkien alkeistapahtumien lukumäärästä. Muuta todennäköisyyden määritelmää ei esitetä, mutta marginaalissa kerrotaan "Kolmogorovin aksioomat" ja kohta esimerkissä käsitellään tilannetta, jossa puhutaan tapahtumasta, jonka todennäköisyys on 64 %. Toisessa alaluvussa **Tekijä** esittelee "geometrisen todennäköisyyden", joka on

$$P(A) = \frac{m(A)}{m(E)},$$

missä $m(A)$ on "tapahtumalle suotuisan geometrisen kuvion mitta" ja "$m(E)$ perusjoukkoa kuvaavan geometrisen kuvion mitta". Olisi ollut suotavaa jotenkin ilmaista, että tässä määrittelyssä E:n eri osat ovat "yhtä todennäköisiä". Asianomaisessa kohdassa **Juuren** esitys on samalla tavalla puutteellinen.

Seuraavaksi **Tekijä** esittelee kombinatoriikkaa. Niin **Tekijä** kuin **Juurikin** havainnollistavat useampiportaisia valintatilanteita *puumalleiksi* kutsumillaan kaavioilla käsitettä mitenkään määrittelemättä. Molemmat kirjat käyttävät vanhahtavaa terminologiaa nimittäessään k-alkioisia joukkoja *k-kombinaatioiksi*. Vain **Juuri** kertoo merkinnän $\binom{n}{r}$ nimityksen *binomikerroin* sekä nimityksen syyn ja esittelee Pascalin kolmion.

Sivulla 37 **Tekijä** antaa, ehkä tahattomasti, aseita niiden käteen, jotka varoittelevat liiasta luottamisesta sähköisiin laskulaitteisiin. Kertoman määrittämistä laskimella esittelevässä kappaleessa on nimittäin esimerkkinä ”laskimen antama” $4! = 120$. Toinen kohta, jossa **Tekijän** lukija joutuu ymmälleen, on sivun 64 esimerkki, jossa annetaan kahden kahvinkeittimen ehjänä olemisen todennäköisyydet ja sitten ilmoitetaan, että ”molemmat keittimet ovat toisistaan riippumatta yhtä aikaa rikki 9 koulupäivänä lukuvuodessa”. ”Toisistaan riippumatta” johtaa ajatukset siihen, että keittimien kunnossa tai rikki olo ei riippuisi toisesta keittimestä, toisin kuin esimerkin kertomukseen liittyvät luvut osoittavat.

Sekä **Tekijä** että **Juuri** kertovat, mitä tarkoittaa laskuesimerkeissä usein esiintyvä ”tavallinen korttipakka”. Hyvä näin, sillä omienkin kokemusteni mukaan pakka ei ole kaikille selviö – onhan korttipeli monille ollut ”syntiä”.

Juuri seuraa nytkin opetussuunnitelman järjestystä aloittamalla esityksensä tilastollisten jakaumien peruskäsitteiden esittelyllä. **Tekijä** säästää nämä asiat toiseksi viimeiseen lukuun, ennen todennäköisyysjakaumien käsittelyä. Suurta eroa asioiden käsittelyssä ei kuitenkaan näytä olevan. Vain **Tekijä** näyttää hiukan sivuavan tilastollisen aineiston graafisen havainnollistamisen sudenkuoppia, joita esimerkiksi tiedotusvälineissä kovin usein tapaa.

Yksi syy sijoittaa todennäköisyyslaskennan ja tilastoasioiden esittely kurssiin 10 lienee satunnaismuuttujan kertymäfunktion F ja tiheysfunktion f yhteyden palauttaminen integraaliin. Paha kyllä peruskaavaa

$$F(x) = \int_{-\infty}^{x} f(t)\,dt$$

ei oikein voida käyttää, kun aikaisemmat kurssit eivät mainitse epäoleellisia integraaleja. **Juuri** näyttää kerta kaikkiaan välttävän integraalimerkkiä, jossa alarajan paikalla on $-\infty$. **Tekijä** antaa sentään jatkuvan jakauman odotusarvon ja keskihajonnan epäoleellisina integraaleina. **Juuri** tuntee nämä jakauman tunnusluvut vain äärellisellä välillä määritellyn jatkuvan jakauman tapauksessa, joskin se käyttää termejä normaalijakauman yhteydessä. **Tekijä** kirjoittaa myös marginaaliin ”syventävän tiedon”

$$\Phi(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-\frac{1}{2}x^2}\,dx.$$

Vain **Juuri** kertoo lukijoilleen usein esimerkeinkin, mikä on Poissonin jakauma. Luotettavuusteorian eksponenttijakaumakin vilahtaa kirjan esimerkissä, vaikkei sen merkitystä erikseen korosteta.

Todennäköisyysjakaumien esittely on kummassakin kirjassa, niin kuin yleinen käytäntö on, jaettu diskreetteihin ja jatkuviin. **Tekijälle** diskreetti satunnaismuuttuja näyttää saavan vain äärellisen monta arvoa, **Juuri** tuntee myös, kuten yllä mainittiin, Poissonin jakauman, mutta diskreetin jakauman tunnuslukujen kaavoissa esiintyy **Juuressakin** vain äärellisiä summia. Kumpikin kirja kertoo jatkuvaksi jakauman, jossa satunnaismuuttujalla voi saada arvokseen minkä tahansa reaaliluvun joltakin väliltä, ja kirjojen mukaan kullakin arvolla on tällöin pistetodennäköisyys nolla. Kaikki satunnaismuuttujat eivät kuitenkaan sovi näihin kahteen kategoriaan: kirjojen esimerkeissäkin tavataan satunnaismuuttujia, joiden arvojoukko on reaalilukuväli, mutta joilla jollakin yksittäisellä arvolla on positiivinen todennäköisyys. Ehkäpä tämänkin asian olisi voinut mainita.

Molemmissa kirjoissa on runsaasti mielenkiintoisia harjoitustehtäviä, vaativiakin. Kumpikin esittää harjoituksissa mm. *de Mérén* kaksoiskuutosongelman, laatikkoleikin *Monty Hall* -ongelman ja kysymyksen siitä, miten isossa ihmisryhmässä ainakin kahden saman syntymäpäivän esiintyminen on tapahtuu todennäköisyydellä $\geq \frac{1}{2}$. Kurssin aikana tehtävistä kuitenkin ehdittäneen käsitellä vain murto-osa.

Todennäköisyystehtävän muotoilu sellaiseksi, että kaikki lukijat sen samalla tavalla ymmärtäisivät, on joskus vaikeaa, ja sanamuotojen merkityksistä voi saivarrella. Mitä oikeastaan tarkoittaa **Tekijän** ilmoitus, että ”laatikossa on kaikki suomen kielen 29 aakkosta”? Ei myöskään **Juuren** ilmoitus, että ”permutaatio on joukon alkioista muodostettu jono” vaikuta kovin täsmälliseltä.

Todennäköisyystehtävien ratkaisu on usein yksinkertainen murtoluku. Huomio kiinnittyy siihen, että kummankaan kirjan ratkaisuosastolle eivät tyyppiä $\frac{1}{3}$ olevat vastaukset riitä, vaan ne on aina ilmoitettu myös desimaalilukulikiarvoina. – Siihen, että esimerkeissä oleva yksinkertainenkin algebra annetaan koneen tehtäväksi, on lukija jo sarjojen aikaisemmissa osissa turtunut. Ei käytäntöä silti hyvänä opi pitämään.

Molemmat kirjat, etenkin **Tekijä**, käyttävät esimerkeissä ja harjoitustehtävissä melko paljon erisnimiä. Miten mahtavat tekijät näitä valita? Pikaisen selauksen mukaan **Tekijässä** esiintyy ainakin 62 eri nimeä, näistä ehkä 32 maskuliinisia. **Juuressa** nimiä on ainakin 31, poikia 17. Kumpaakaan kirjaa ei voine erityisesti syyllistää jommankumman sukupuolen suosimisesta. **Juuren** lievän poikavaltaisuuden aiheuttaa esimerkki, jossa ovat mukana kaikki Aleksis Kiven seitsemän veljestä.

Jälleen kerran voi todeta kilpailevien sarjojen päätyneen tasapeliin. **Juuressa** saattaa olla enemmän asioita, mutta onko se vain hyvä, kun ottaa huomioon kurssiin tarjolla olevan varsin rajallisen ajan? Kummallakin kirjalla on puutteensa, mutta kokonaisvaikutelma on myönteinen.