

Sissejuhatus

Miks ma vihkasin algebrat, aga armastan statistikat

Mul on matemaatikaga alati ebamugav suhe olnud. Ma ei armasta numbriteid numbrita enda pärast. Mind ei vaimusta peened valemid, millel pole reaalses maailmas rakendust. Eriti vastumeelne oli mulle keskkooliaegne algebra, sest keegi ei vaevunud seletama, miks ma peaksin seda õppima. Kui suure pindala katab parabool? Keda see huvitab?

Tõtt-öelda leidis üks mu elu parimaid hetki aset keskkooli lõpuklassis, algebrakursuse esimese semestri lõpus. Murdsin pead lõpuksami kallal, milleks olin tõesti vähem õppinud, kui oleksin pidanud. (Olin mõni nädal varem oma esimesena valitud ülikooli sisse saanud ning seepärast minetanud viimsegi huviraasu selle kursuse vastu.) Lõpueksamil küsimusi vaadates tundusid need mulle täiesti võõrad. Asi polnud selles, et mul oluks raske neile küsimustele vastata, vaid ma ei saanud isegi sellest aru, mida küsiti. See polnud esimene kord, kui ma ette valmistamata eksamile läksin, ent Donald Rumsfeldi parafraseerides, tavaliselt ma vähemalt teadsin, mida ma ei tea. See eksam oli minu silmis veel hullem hiina keel kui muud ained. Lappasin mõnda aega küsimustikku ning andsin alla. Astusin klassi ette, kus mu algebraõpetaja (nimetagem teda siinkohal Carol Smithiks) eksamineeritavil

silma peal hoidis. „Proua Smith,“ ütlesin ma, „mulle ei tule siin paljud asjad üldse tuttavad ette.“

Piisab sellest, kui ütlen, et proua Smith ei sallinud mind sugugi rohkem kui mina teda. Jah, praegu võin ma tunnista, et õpilasesinduse presidendina kasutasin mõnikord oma piiratud võimu ja ajastasin ülekoolilised koosolekud nimme sedasi, et proua Smithi algebratund ära jääks. Jah, mina oma sõpradega lasin tunni ajal proua Smithile „sala-jaselt austajalt“ lilli tuua, ise tagapingis naerda luksudes, kui õpetaja piinlikkusest punastades ringi vaatas. Jah, ma lakkasin koduseid töid tegemast niipea, kui ülikooli sisse sain.

Nii et kui ma eksami ajal proua Smithi ette astusin ja ütlesin, et teema mulle üldse tuttav ei tundu, ei äratanud ma temas erilist kaastunnet. „Charles,“ ütles ta valjusti, kõnetades pealtnäha mind, kuid hoolitsedes ühtlasi selle eest, et kogu klass kuuleks, „kui sa oleksid õppinud, oleks teema sulle hulga tuttavam.“ See oli mõjuv argument.

Nii ma siis löntsisin oma pinki tagasi. Mõni minut hiljem astus klassi ette Brian Arbetter, kes oli minust algebras palju parem, ja soistas midagi proua Smithile kõrva. Õpetaja soistas vastu ja siis juhtus midagi täiesti enneolematut. „Klass, palun tähelepanu,“ kuulutas proua Smith. „Selgub, et ma olen teile kogemata andnud teise semestri eksamiküsimused.“ Eksami algusest oli juba nii palju aega möödas, et see tuli tervenisti tühistada ja uus aeg määrata.

Minu eufooria oli kirjeldamatu. Hilisemas elus abiellusin ma võr-ratu naisega, meil on kolm tervet ja toredat last, ma olen avaldanud mitu raamatut ning käinud vaatamas Taj Mahali ja Angkor Wati. Ja ometi on päev, mil minu algebraõpetaja sai oma teenitud tasu, üks mu elu viiest tippphetkest. (See, et ma viimaks lõpueksamil peaaegu läbi kukkusin, ei varjutanud seda imelist elamust kuigivõrd.)

See juhtum algebraeksamil kõneleb teile mõndagi mu suhtest matemaatikaga, ent mitte kõike. Kummalisel kombel armastasin ma keskkoolis füüsikat, kuigi see toetub väga olulisel määral tollele samale

algebrale, mida ma proua Smithi klassis keeldusin õppimast. Miks siis? *Sest füüsikal on selge eesmärk.* Ma mäletan väga eredalt, kuidas keskkooli füüsikaõpetaja näitas meile pesapallimeistrivõistluste ajal, kuidas kiirrenduse põhivalemi abil välja arvutada, kui kaugele löödud pall lendab. See on väga lahe ning selsamal valemil on veel palju sotsiaalse kandes-pinnaga rakendusi.

Ülikoolis õppima hakates nautisin ma tohutult tõenäosust – jällegi seepärast, et see aitas mõista tõsielus ette tulevaid põnevaid olukordi. Tagasivaates saan nüüd aru, et algebratunnis ei häirinud mind mitte matemaatika, vaid see, et keegi ei vaevunud kunagi selle mõtet selgitama. Kui meid ei lumma pelgalt valemite elegants (ja mind see päris kindlasti ei lumma), on algebra lihtsalt hunnik üksluisseid ja mehaanilisi valemuid – vähemalt nõndaviisi seda mulle õpetati.

Siit jõuame statistika juurde (mis selle raamatu huvides hõlmab ka tõenäosust). Ma armastan statistikat. Statistika abil saab seletada kõike alates DNA-testidest kuni lotopiletite ostmise mõttetuseni. Statistika võib aidata tuvastada vähi ja südamehaigustega seotud tegureid ning paljastada riigieksamitel tehtavaid pettusi. Statistika võib olla abiks isegi telemängude võitmisel. Minu lapsepõlves jooksis telekas kuulus saade nimega „Let's Make a Deal“ („Lööme käed“), mida vedas sama kuulus saatejuht Monty Hall. Iga saate lõpus seisis mängu võitja koos Montyga kolme suure ukse ees: nimetagem neid üks nr 1, üks nr 2 ja üks nr 3. Monty Hall selgitas mängijale, et ühe ukse taga on väga ihaldusväärne auhind – näiteks uus auto –, ülejäänud kahe ukse taga aga kits. Asja mõte oli väga lihtne: mängija valib ühe ukse ja saab selle taga olnud auhinna endale.

Mängijal, kes koos Monty Halliga nende uste ees seisis, oli üks võimalus kolmest valida üks, mille taga on väärtuslik auhind. Kuid saates oli nõks, mis on sellest ajast peale statistikuid vaimustanud (ja kõiki teisi segadusse ajanud). Kui mängija oli oma valiku teinud, avas Monty Hall ühe kahest ülejäänud uksest, mille tagant tuli alati lagedale kits. Sel-

guse mõttes oletagem, et mängija valib ukse nr 1. Seejärel avab Monty ukse nr 3 ning laval seisab elus kits. Kaks ust (nr 1 ja nr 2) jäävad siiski suletuks. Kui väärtuslik auhind on nr 1 taga, siis mängija võidab; kui see on nr 2 taga, siis kaotab. Kuid siis läks asi põnevamaks: Monty pöördus mängija poole ja küsis, kas ta tahab ümber mõelda ja teise ukse valida (antud juhul valida nr 1 asemel nr 2). Pidage meeles, et mõlemad ukсед on ikka veel kinni ning mängijal on vaid üks uus teadmine: et ühe ukse taga, mida ta ei valinud, oli kits.

Kas ta peaks ümber mõtlema?

Vastus on jah. Miks? Sellest räägime peatükis 5 ½.

Statistika paradoks seisneb selles, et see on kõikjal olemas – pesapallist kuni valimiseelsete küsitlusteni –, kuid eriala ennast peetakse üleüldiselt igavaks ja arusaamatuks. Paljud statistikaõpikud ja -tunnid kubitsevad matemaatikast ja segastest formuleeringutest. Uskuge mind, tehnilised üksikasjad on üliolulised (ja huvitavad), kuid see kõik on puhaskiina keel, kui õiget tunnetust ei teki. Võib-olla see tunnetus eriti oluline ei paistagi, kuni te pole veendunud, et selle kõige õppimisel üldse mingit mõtet on. Ma luban, et iga peatükk selles raamatus vastab põhiküsimusele, mille ma ise esitasin (tulutult) oma keskkooliaegsele algebraõpetajale: *milleks seda vaja on?*

See raamat räägibki intuitsioonist. Siin leidub vähe matemaatikat, võrrandeid ja graafikuid; kui neid siiski ette tuleb, siis ma luban, et neil on selge ja valgestav eesmärk. See-eest leidub raamatus ohtralt näiteid, mis veenavad lugejat, et statistika õppimiseks on suurepäraseid põhjusi. *Statistika võib olla äärmiselt huvitav ja enamik sellest pole kuigi keeruline.*

Selle raamatu idee sündis üsna varsti pärast mu õnnetut läbielamist proua Smithi algebratunnis. Ülikoolis hakkasin õppima majandust ja politoloogiat. Enne kui õppetöö algas, määrati mind (mõistetaval põhjusel) koos mitme klassikaaslasega „matemaatikalaagrisse“, et valmistada meid ette eesseisvaiks katsumusiks. Istusime kolm nädalat ühes

keldriklassis, kus ei olnud ainsatki akent, ja õppisime matemaatikat (tõsijutt).

Just seal kogesin ma ühel päeval midagi karjääriilmutuse sarnast. Juhendaja üritas selgitada asjaolusid, mille juures lõpmatu jada liikmete summa läheneb lõplikule arvule. Varuge nüüd hetk kannatust, see põhimõte saab teile kohe selgeks. (Praegusel hetkel on teil küllap samasugune tunne nagu minul tolles akendeta klassis.) Lõpmatu jada on arvurida, mis jätkub igavesti, näiteks $1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} \dots$. Kolm punkti tähendavadki seda, et jada jätkub lõpmatuseni.

Tol ajal tekitas see meile tõsist peamurdmist. Mingi tõestuse abil, mille olen ammu unustanud, üritas juhendaja meid veenda, et arvude jada võib jätkuda lõputult ja ometi võib nende summa (ligikaudu) võrrelda lõpliku arvuga. Mu klassikaaslane Will Warshauer keeldus seda uskumast, hoolimata mõjuvast matemaatilisest tõestusest. (Ausalt öeldes olin minagi pisut skeptiline.) Kuidas võib millegi lõputu summa olla lõplik?

Siis tabas mind inspiratsioon, või täpsemalt intuitsioon selle kohta, mida juhendaja üritas seletada. Pöördusin Willi poole ja rääkisin, mis ma välja olin mõelnud. Kujutage ette, et seisate seinast täpselt kahe meetri kaugusel.

Nüüd vähendage kaugust seinaga poole maa ehk ühe meetri võrra, nii et jääte sellest ühe meetri kaugusele.

Meetri kauguselt astuge seinale veel poole vahemaa (poole meetri) võrra lähemale. Ja poole meetri kauguselt tehke jälle sama: liikuge edasi 25 cm ehk veerand meetrit. Siis liikuge 12,5 cm ehk ½ meetrit. Ja nii edasi.

Tasapisi jõuate te seinale paganama lähedale. (Näiteks siis, kui te seisate seinast 1/1024 cm kaugusel, peate vähendama vahemaad poole ehk 1/2048 cm võrra.) Kuid vastu seinat te ei pörkagi, sest iga samm viib teid edasi vaid poole võrra järelejäänud vahemaast. Teisisõnu, te jõuate seinale lõpmata lähedale, kuid mitte kunagi seinani välja. Kui mõõta

teie samme meetrites, moodustubki säärane jada: $1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} \dots$

Selles ongi asja iva: kuigi jätkate liikumist lõpmatuseni, kusjuures iga samm lähendab teid seinale poole võrra järelejäänud vahemaast, ei saa kogu läbitud teepikkus kuidagi ületada kahte meetrit ehk esialgset vahemaad seinani. Matemaatika huvides võib teie läbitava vahemaa ümardada kahe meetrini, mis on arvutamiseks väga mugav. Matemaatik ütleks, et lõpmatu jada summa $1\text{ m} + \frac{1}{2}\text{ m} + \frac{1}{4}\text{ m} + \frac{1}{8}\text{ m} \dots$ läheneb kahele meetrile, ning just seda üritaski juhendaja meile tol päeval selgeks teha.

Oluline on, et ma veensin selles Willi ja iseennastki. Ma ei mäleta matemaatilist tõestust väitele, et lõpmatu jada liikmete summa võib läheneda lõplikule arvule, kuid seda saan alati internetist vaadata. Ja kui ma seda vaatan, siis küllap sellest ka aru saan. Minu kogemuse põhjal teeb intuitsioon matemaatikat ja muid tehnilisi üksikasju arusaadavamaks, aga vastupidi mitte tingimata.

Selle raamatu eesmärk on muuta kõige olulisemad statistika mõisted jõukohasemaks ja arusaadavamaks – mitte üksnes meiesugustele, kes pidid neid akendeta klassis tuupima, vaid igaühele, kellele pakub huvi arvude ja andmete uskumatu vägi.

Nüüd, kui ma olen just rääkinud, et statistika põhivahendid on raskemini tajutavad ja vähem arusaadavad, kui need võiksid olla, tulen lagedale esmapilgul vasturääkiva väitega: statistika võib olla *liigagi arusaadav* selles mõttes, et igaüks, kel on olemas mingid andmed ja arvuti, võib mõne klahvivajutusega sooritada keerukaid statistilisi protseduure. Häda on selles, et kui andmeid on vähe või kui statistilisi võtteid kasutatakse valesti, võivad järeldused olla pööraselt eksitavad ja kogu ni ohtlikud. Võtame näiteks kujuteldava internetiuudise pealkirjaga „Inimestel, kes teevad tööl lühikesi puhkepause, on suurem oht vähki surra“. Kujutage ette, et see pealkiri torkab teile veebis ringi liikudes silma. 36 000 kontoritöötaja seas (see on tohutu andmekogum!) tehtud

pealtnäha usaldusväärse uuringu kinnitusel oli neil töötajail, kes lahkuvad tööpäeva vältel kontorist regulaarselt kümnest minutisteks pausideks, 41% suurem tõenäosus järgmise viie aasta jooksul vähki haigestuda kui töötajail, kes tööpäeva jooksul kordagi väljas ei käinud. Ilmselt peaks sellisele avastusele kuidagi reageerima – näiteks korraldama riikliku hoiatuskampania, et vähendada tööl tehtavaid puhkepause.

Või äkki peaks hoopis järele mõtlema, mida paljud töötajad selle kümnest minutise pausi ajal teevad. Minu töökogemus ütleb, et mitmed neist töötajaist, kes natukeseks ajaks kontorist lahkuvad, kogunevad hoone ette suitsetama (tekitades seejuures suitsupilve, millest me kõik peame sisse-välja käies läbi marssima). Sellest järeldaksin, et vähki ei põhjusta mitte lühikesed puhkepausid, vaid hoopis sigaretid. Ma mõtlesin meelega välja eriti absurdse näite, kuid uskuge mind, paljud ehtsad statistilised soperdised on peaaegu sama absurdsed, kui need lahti harutada.

Statistika on justkui suure kaliibriga tulirelv – õigesti kasutades tarvilik abivahend, kuid valedes kätes äärmiselt ohtlik. See raamat ei tee teist statistikaeksperti, aga õpetab tundma selle valdkonna vastu piisavalt austust ja lugupidamist, et te ei käsitseks statistikat taparelvana.

See raamat pole õpik. Siin ei läbita kohustuslikku programmi ega seletata teemasid rangelt ettenähtud moel. *See raamat on koostatud selleks, et tutvustada teile igapäevase eluga kõige vahetumalt seotud statistikamõisteid.* Mille põhjal järeldavad teadlased, et üks või teine asi põhjustab vähki? Kuidas tehakse arvamusuuringuid (ja mis võib seejuures viltu minna)? Kes „valetavad statistika abil“ ja kuidas nad seda teevad? Kuidas ennustab pank teie krediitkaardiga tehtud ostude põhjal, kas teil võib olla kalduvus maksetega hilineda? (Tõsijutt, seda nad oskavad.)

Kui te tahate saada aru uudiste taga olevaist arvudest ja hinnata õigesti andmete tohutut (ning üha kasvavat) väge, on teil hea seda kõike teada. Lõpuks loodan ma teid veenda tähelepanekus, mille esimesena tegi Rootsi matemaatik ja kirjanik Andrejs Dunkels: statistika abil on

lihtne valetada, aga raske on ilma selleta tõtt rääkida.

Kuid minu lootused on veelgi suuremad. Ma arvan, et te võite statistikat lausa nautima hakata. Ideed, millel see põhineb, on pööraselt põnevad ja olulised. Kogu loo võti seisneb selles, et olulised ideed tuleb eraldada müstilistest tehnilistest üksikasjadest, mis kipuvad teele ette jääma. See ongi „Alasti statistika“.

Tänuavaldused

See raamat sündis austusavaldusena ühele kirjastuse W. W. Norton varasemale menuraamatule, Darrell Huffi 1950. aastail kirjutatud teosele „How to Lie with Statistics“ („Kuidas statistika abil valetada“), mida on müüdud üle miljoni eksemplari. Nagu käesolev raamat, oli seegi kirjutatud eesmärgiga demüstifitseerida statistikat ja veenda lugejaid tõigas, et kui nad oluliste väidete taga peituvaid arve ei mõista, võib see neile kahju tuua. Ma loodan, et ei tee selle väitega isand Huffi meistriteosele kurja. Igal juhul oleks puhas rõõm, kui minu raamatut oleks 50 aasta pärast miljon eksemplari müüdud!

Ma olen jätkuvalt tänulik kirjastusele W. W. Norton ja eriti Drake McFeelyle võimaluse eest kirjutada raamatuid, mis käsitlevad olulisi teemasid tavalugejale arusaadaval moel. Drake on nüüdseks juba üle kümne aasta olnud mu väga hea sõber ja toetaja.

Jeff Shreve on kirjastuses W. W. Norton see mees, kes raamatu ilmale tõi. Jeffiga esimest korda kohtudes võiks arvata, et nii malbe inimene ei suuda mingil juhul kõiki säärase raamatu kirjastamiseks vajalikke toiminguid õigeks tähtajaks läbi suruda. Vale puha. Jah, ta on tõesti malbe, ent kuidagi teevad tema leebed meeldetuletused ometi tarviliku töö ära. (Näiteks peavad need tänuavaldused homme hommikuks valmis saama.) Ma olen väga tänulik sellise lahke järelevaataja eest, kes asjad joonde ajab.

Suurima tänu võlgnen ma neile paljudele inimestele, kes viivad