



R. Fisher

Anglų statistikas ir genetikas Seras Ronaldas Ailmeras Fišeris (Sir Ronald Aylmer Fisher) gimė 1890 m. vasario 17 d. Londone. Pradėjęs 1909 m. studijas Kembri-džo universitete, jį baigė 1912 m. ir gavo astronomo diplomą. Į statistiką Fišerį atvedė jo susidomėjimas astronomijos duomenų paklaidų analize. Jau 1912 m. paskelbė ir savo pirmąjį darbą, kuris tapo maksimalaus tikėtimumo metodo pradžia. Nuo 1920 m. pradėjo domėtis genetika. S. R. A. Fišeris mirė 1962 m. liepos 29 d. Australijoje, Adelaidės miestelyje.

1933 m. I. Neimanas ir E. Š. Pirsonas paskelbė pagrindinius hipotezių tikrinimo principus.



J. Neyman

Iržis Neimanas (Jerzy Neyman) gimė 1894 m. balandžio 16 d. Moldavijoje. Studijavo Krokuvos universitete. Mokslų daktaru tapo Varšuvoje 1924 metais. Dirbo su K. Pirsonu, tačiau juo nusivylė dėl pastarojo skeptiško požiūrio į modernesnę matematiką. Statistika Neimaną sudomino E. Š. Pirsonas, su kuriuo 1928–1933 m. laikotarpiu parašė du svarbius darbus: „Efektyviausio statistinės hipotezės testo problema“ ir „Statistinių hipotezių tikrinimas remiantis a priori tikimybėmis“. 1938 m. Neimanas išvyko į JAV ir dirbo Berklyje. Jo darbai apie hipotezių tikrinimą ir pasikliautinuosius intervalus buvo labai reikšmingi statistikos teorijai. I. Neimanas mirė 1981 m. rugpjūčio 5 d. JAV, Kalifornijos valstijoje.



E. Pearson

Karlo Pirsono sūnus Egonas Šarpas Pirsonas (Egon Sharp Pearson) gimė 1895 m. rugpjūčio 11 d. Hampštete, netoli Londono. Baigęs 1914 m. Vinčesterio kolegiją, tais pačiais metais pradėjo studijas Kembri-džo universitete, tačiau prasidėjus I pasauliniam karui jas nutraukė ir niekad to universiteto nebaigė. Universiteto diplomą Pirsonui buvo įteiktas 1919 m. išlaikius specialius egzaminus, organizuotus nutraukusiems studijas dėl karo. 1921 m. pradėjo dirbti Londono universiteto koledžo Taikomosios statistikos departamente, kuriam vadovavo jo tėvas. Su Neimanu jų keliai susikirto 1925 m., kai pastarasis gavo dviejų metų stipendiją moksliniams tyrimams ir pirmus metus leido Londono universiteto koledže. Tačiau bendradarbiauti jie pradėjo tik 1926 metais. E. Š. Pirsonas mirė 1980 m. birželio 12 d. Anglijoje, Midkursto miestelyje.

Tas mokslas sparčiai vystėsi ir toliau. Naujų informacinių technologijų era atvėrė statistikai plačias galimybes. Pradėti plėtoti nauji, kompiuterių galimybes panaudojantys metodai. Iš jų pirmiausia išskirčiau *butstrepa* — imčių (duomenų) kompiuteriniu generavimu paremtą statistikinę analizę. Šį metodą plėtoti paskatino straipsnis, kurį Stenfordo universiteto profesorius Bredlis Efronas (*Bredley Efron*) paskelbė 1979 metais. Gali būti, kad pavadinimą *bootstrap* (angl. *batraiščiai*) Efronas sugalvojo prisiminęs baroną Miunchauzeną, kuris save iš pelkės traukė už plaukų. Mat tas žodis dažniausiai naudojamas su posakiu *by one's own bootstraps* (nuosavais batraiščiais), kuriuo nusakomos pastangos išsisukti iš padėties be kitų pagalbos. Iki šiol *butstrepiui* skiriamas didžiulis tiek statistikos teoretikų, tiek praktikų dėmesys. Vien mokslinių publikacijų galima suskaičiuoti daugiau nei pusanтро tūkstančio. Parašytos kelios monografijos. Tokį dėmesį lėmė ir metodo paprastumas, ir vis tobulėjančios kompiuterinės technikos galimybės.

Kita pakankamai nauja statistikos kryptis — funkcinė duomenų analizė. Jos tikslai iš esmės tokie pat kaip, ir kitų statistikos šakų:

- duomenų pateikimo būdų, patogių tolesnei jų analizei, paieška;
- pagrindinių duomenų charakteristikų išryškinimas;
- neapibrėžtumų šaltinių paieška ir jų tyrimas;
- dviejų ar daugiau duomenų grupių palyginimas.

Skirtumas tik tas, kad duomenys interpretuojami kaip atsitiktinių funkcijų trajektorijos, o jų analizei taikomi abstrakčios statistikos ir funkcinės analizės metodai. Kaip ir butstrepas, funkcinė duomenų analizė imli kompiuterių galimybėms.

Akivaizdu, kad statistikos reikšmė šiame modernių technologijų amžiuje nuolat didės. Jau dabar duomenų aibės gali būti daug didesnės, nei įprasta — siekti šimtus gigabaitų ar terabaitų. Beje, nenuostabu, kad nemažai naujovių tokiems duomenų masyvams apdoroti siūlo ne tik statistikai. Štai kompiuterių inžinerijos specialistai pradėjo plėtoti neuroninius tinklus, kurti duomenų sandėliavimo technologijas, duomenų vaizdavimo, ekspertines sistemas. Visa tai kartu su statistika dažnai vadinama *data mining* (duomenų kalnakasyba). Skirtingai nei statistika, neuroniniai tinklai ar ekspertinės sistemos kur kas mažesnę dėmesį skiria modelių paieškai ar asimptotinei didelių imčių teorijai ir siūlo tam tikrą „mokymo“ filosofiją.

Ar statistika tik mokslas?

Statistikos negalime priskirti tai pačiai mokslų grupei kaip matematiką, fiziką, chemiją ar biologiją. Dažniausiai ji tarnauja kitų, iš jų ir minėtų, disciplinų problemoms spręsti. Būtent iš to gimsta nauji metodai ar net naujos kryptys.

Štai Gosetas *t* testą atrado tyrinėdamas nedideles alaus daryklų duomenų imtis.

*Viljamas Sili Gosetas (William Sealey Gosset) gimė 1876 m. birželio 13 d. Anglijoje, Kanterburi miestelyje. Naujojoje Oksfordo kolegijoje studijavo chemiją ir matematiką. Darbą gavo kaip chemikas Gineso alaus darykloje Dubline, rūpinosi alaus kokybės kontrole. Čia jis ir atrado garsųjį *t* testą, kuris dažnai vadinamas Studento testu. Mat Gosetas savo darbus pasirašinėjo Studento slapyvardžiu. Nuo 1922 metų Gosetas gavo pagalbininką statistikai, o vėliau įkūrė ir statistikos skyrių, kuris veikė iki pat 1934 metų. V. S. Gosetas mirė 1937 m. spalio 16 d. Anglijoje.*

Neparametriniai statistikos metodai pradėti plėtoti, kai vaistų pramonei prireikė atlikti daug statistikinių testų, o šiuolaikiškų kompiuterių dar nebuvo. O štai parametrinius laiko eilučių modelius Dž. U. Jule rado tirdamas kviečių kainų ciklus.



G. Jule

Džordžas Udni Jule (Georg Udny Yule) gimė 1871 m. vasario 18 d. Škotijoje, Morhame, netoli Hadingtono. Londono universiteto koledžą lankyti pradėjo 1887 m., studijavo inžineriją ir 1890 m. gavo inžinieriaus diplomą. Jau 1895 m. paskelbė pirmąjį savo mokslinį darbą, kuriame matyti koreliacinės analizės užuomazgos. Didžiausi jo nuopelnai statistikai siejami su koreliacine analize ir regresija. Jo darbų ciklas „Koreliacijos teorija“ pasirodė 1947 metais. Dž. U. Jule mirė 1951 m. birželio 26 d. Anglijoje.

Galėtume rasti daugybę tokių praktinių pavyzdžių, davusių pradžią įvairiems statistikos metodus ar kryptims. Pastaraisiais metais daugelio statistikų žvilgsniai krypta į finansų ir draudimo matematiką, naująsias ekonomikos sritis, biotechnologijas ir pan.

Esminis statistikos, kaip mokslo, objektas — neapibrėžtumai, slypintys surinktuose duomenyse. Tokiems neapibrėžtumams tirti yra įvairių metodų. Tačiau jie negali būti taikomi savaime. Vartotojui juos reikia modifikuoti, priderinti prie konkrečių poreikių. Statistikos metodus taip pat galima pritaikyti įvairioms technologijoms kurti. Geriausias to pavyzdys — kokybės kontrolės sistemos. Garsusis statistikas Rao apie jas atsiliepia taip:

Retai pasitaiko tokių technologinių išradimų kaip kokybės kontrolė, kuri taip paplitusi, teoriškai tokia paprasta, efektyvi ir lengvai pritaikoma, kurios graža tokia didelė ir kuri reikalauja tokių mažų investicijų.

Po šių žodžių drąsiai galiu teigti, kad statistika yra ne vien tik mokslas, bet ir inžinerija (technologija). Labai taikliai mokslo ir inžinerijos skirtumą apibūdino F. Brooksas: „Mokslas kuria tam, kad studijuotų, o inžinerija studijuoja tam, kad sukurtų.“

Esu giliai įsitikinęs, kad statistika yra ir menas. Aišku, ne vieno piliečio sąmonėje statistika dar asocijuojasi su trečiaja melo rūšimi: *paprastas melas, didelis melas, statistika*. Bet prisiminkime garsiąją P. Picasso frazę: „*Menas yra melas, leidžiantis suvokti tiesą.*“

Statistika leidžia tiesą suprasti. Jos metodologija remiasi induktyviais argumentais ir joje nemažai kontraversijų. Dirbdami su tais pačiais duomenimis, skirtingi statistikai gali gauti nevienodus rezultatus. Mat, norint priversti skaičius pasakyti visą tiesą, reikia patirties, įkvėpimo, pagaliau ir sėkmės. Ta prasme statistika yra menas. Visi, kas kuria statistikinius modelius, algoritmus ar procedūras, gali save vadinti menininkais.

Statistikos ir statistikų poreikis

Be jokių abejonių, statistikos ateitis daugiausia priklauso ne tiek nuo jos pačios, kiek nuo tų, kuriems statistikos reikia. O jų yra daug: vyriausybė, mokslas, inžinerija, medicina, žemės ūkis ir kt., net literatūra ir archeologija.

Bet kurios valstybės vyriausybei statistika yra įrankis daryti trumpalaikes ar ilgalaikes prognozes, planuoti konkrečią priemonę socialiniams ar ekonominiams tikslams siekti. Statistika taikoma nagrinėjant demografinę situaciją, tiriant vartojimo reikmių bei paslaugų paklausą ir pan. Sakoma, kuo labiau išsivysčiusi šalis, tuo geresnė jos statistika. Lietuvoje turime Statistikos departamentą, kurio 2002–2004 m. veiklos strategija yra:

Rengti ir laiku teikti visuomenei kokybišką statistikinę informaciją apie valstybės ekonominius, demografinius procesus, socialinius veiksnius ir visuomeninius bei aplinkos pokyčius.

Mano nuomone, didžiausias statistikos naudotojas vis dėlto yra mokslas. Eksperimentų planavimas, įvairiausių praktinių hipotezių tikrinimas, nežinomų parametrų vertinimas, rezultatų interpretavimas — tai tik dalelė poreikių, kuriuos gali patenkinti statistika. Rezus faktoriaus kraujo grupėse atradimas, prie kurio prisidėjo statistikai, — puikus statistikos ir mokslo sąveikos pavyzdys. Šiuolaikiškas mokslas tampa vis labiau matematinis ir statistinis. Ir tai būdinga ne tik fizikai ar inžinerijai, bet ir biologijai, geofizikai, aplinkosaugai, psichologijai ir ekonomikai. Didėjantį statistikos poreikį rodo besiformuojančios ar jau susiformavusios naujos mokslo šakos: ekonometrija, technometrija, psichometrija, chemometrija. Kiekviena jų naudoja baziniais statistikos principais, suteikdama jiems tam tikros specifikos.

Ekonometrija — mokslas, statistikos metodais tiriantis kiekybinius ekonominių objektų ir ūkinių procesų sąryšius. Daugelis ekonominių dydžių išreiškiami skaičiais (nedarbo lygis, bendrasis vidaus produktas, mokesčiai, palūkanos ir kt.), todėl juos galima analizuoti, tirti jų dinamiką, sąryšius. Tam tikslui pasitelkiami statistikos metodai bei šiuolaikiškos informacinės technologijos. Ekonometras yra savotiškas ekonomikos analitikas. 1930 metais susikūrė tarptautinė ekonometrų draugija (detaliau su ja galima susipažinti pabuvojus jų svetainėje adresu <http://www.econometricsociety.com/es/>) ir pradėtas leisti žurnalas „Econometrics“. Dabar ekonometrijai skirtų mokslinių žurnalų yra daugiau kaip dešimt. Pripažinimo ekonometrija susilaukė jau 1998 m., kai Nobelio premijos laureatais ekonomikos srityje tapo Mertonas (*Merton*) ir Šoulas (*Scholes*), kuriuos drąsiai galima priskirti prie ekonometrų. Jų darbai daugiausia lėmė finansų ekonometrijos plėtojimą, o jų akcijų kainos modelį privalo žinoti kiekvienas save gerbiantis finansų makleris. Ekonomikos Nobelio premiją 2000 m. pasidalijo du „grynieji“ ekonometrai Makfadenas (*McFaden*) ir Hekmanas (*Heckman*). Taip buvo pripažinta jų sukurta ir išplėta mikroekonometrijos teorija.

Ekonomikos teorija kartu su matematine ekonomika nustato įvairių ekonominių objektų bei procesų tarpusavio sąveikas ir išreiškia jas matematiniais simboliais. Taip sudaromos teorinės prielaidos ekonometriniam modeliui parinkti. Paprastai nagrinėjami trijų tipų modeliai: regresiniai, simultaninių lygčių ir laiko eilučių. Simultaninių lygčių modeliai ir jų analizės metodai

— kertinis ekonometrijos akmuo. Ekonometriniai modeliai sukurti beveik visoms ekonomikos sritims, įskaitant vartojimo prekių rinką, energetiką, sveikatos apsaugą, namų ūkio ekonomiką, pramonės ekonomiką, tarptautinę prekybą, finansus, darbo ekonomiką, monetarinę ekonomiką, transportą, regioninę ekonomiką ir netgi nusikaltimų tyrimą.

Chemometrija statistikos metodus taiko chemijos duomenų analizei. Organinės ir analizinės chemijos, maisto bei aplinkos tyrimų duomenims charakteringas kolinearumas, todėl jiems tirti reikia specialių metodų. Chemometrija išplėtojo dalinį mažiausiųjų kvadratų metodą, įvairius pagrindinių komponentų metodo variantus.

Pagrindinis psichometrijos tikslas — plėtoti psichologiją kaip kiekybinį racionalų mokslą. Jos sukurta testavimo teorija praturtina faktoriinę analizę. Daugiau nei šešis dešimtmečius veikia tarptautinė psichometrijos draugija, leidžianti savo žurnalą „Psychometrika“ (detaliau su jos veikla galima susipažinti interneto svetainėje <http://www.psychometricsociety.com>).

Technometrijos objektas — fizikos, chemijos, mechanikos bei inžineriniai duomenys. Čia dominuoja regresiniai modeliai ir įvairūs jų analizės metodai. Nuo 1959 metų leidžiamas žurnalas „Technometrics“ (internetu svetainės adresas <http://www.amstat.org/publications/technometrics/>).

Dažnai pramonė vadinama didžiausiu statistikos vartotoju. Žinoma, kad gamykloje įdiegus statistinius metodus produktyvumas padidėja nuo 10 iki 100 nuošimčių ir tam nereikia jokių papildomų investicijų. Ne veltui statistikinė kokybės kontrolės sistema laikoma vienu iš svarbiausių praėjo šimtmečio technologinių išradimų. Prognozuojant produktų paklausą, naudojami ir laiko eilučių bei regresiniai modeliai.

Medicinoje statistikos prireikia planuojant eksperimentus, prognozuojant epidemijas, klinikiškiems vaistų tyrimams. Literatūroje — stiliaus kiekybinei analizei, kas praverčia ginčuose dėl autorystės. Archeologijoje statistika padeda nustatyti radinio priklausomybę.

Taigi akivaizdu, kad šiame modernių technologijų amžiuje matematinio ir statistinio raštinumo poreikis vis didėja. Statistikos specialistus paprastai rengia universitetai, siūlydami įvairias programas. Lietuvoje man žinomos tik trys statistikos bakalauro programos: *statistika, finansų ir aktuarinė matematika* bei *ekonometrija*. Visos jos vykdomos Vilniaus universiteto Matematikos ir informatikos fakultete. Ypač populiarios pastarosios dvi programos. Mano nuomone, statistikos mokymas turi būti siejamas su kuria nors sritimi: ekonomika, biologija, geologija ar kita. Ateityje labai perspektyvios turėtų būti ekonometrija, technometrija, biometrija ar chemometrija.

2003 01 28. 12,5 sp. l. Tiražas 1000 egz. Užs. Nr. 1422

Leidykla TEV, Akademijos g. 4, LT-2021 Vilnius

Spausdino UAB „Mokslo aidai“

A. Goštauto g. 12, LT-2600 Vilnius

Viršelį spausdino P. Kalibato II „Petro ofsetas“

Žalgirio g. 90, LT-2600 Vilnius