MONTE KARLO METODAS

Gediminas Stepanauskas

2008

Turinys

1 ĮVADAS. SISTEMOS, MODELIAI, MODELIAVIMAS IR MONTE-KARLO METODAS

Pirmiausia aptarsime sąvokas: sistema, modelis, modeliavimas ir Monte-Karlo metodas. Mes nepateiksime griežtų apibrėžimų, bet paaiškinti šias sąvokas dar reikia ir dėl to, kad literatūroje jos naudojamos nevisada ta pačia prasme.

1.1 Sistemos

Sistema suprantama kaip aibė susijusių objektų, vadinamų tos aibės *elementais*.

Pavyzdžiui ligoninė gali būti nagrinėjama kaip sistema. Šios sistemos elementai – gydytojai, seselės ir pacientai. Elementai (ar pati sistema) turi tam tikrus *atributus*, kurie įgyja logines ir skaitines reikšmes. Mūsų pavyzdyje tai galėtų būti lovų skaičius, Rentgeno spindulių aparatų skaičius, sugebėjimai, kokybė ir pan. Tarp elementų egzistuoja daug ryšių, ir, žinoma, elementai sąveikauja. Šie ryšiai sąlygoja pasikeitimus sistemoje. Pavyzdžiui. Ligoninė turi Rentgeno aparatus ir operatorių šiems aparatams. Jei ligoninėje operatoriaus nėra, tai gydytojai negali naudoti Rentgeno aparatų savo pacientams gydyti. Kitokios, abstrakčios, sistemos pavyzdžiu galėtų būti tiesinė erdvė su apibrėžtomis sudėties ir daugybos iš skaliaro operacijomis (vektorinė erdvė).

Ryšiai gali būti *vidiniai* ir *išoriniai*. Vidiniai ryšiai yra ryšiai tarp elementų sistemos viduje. Išoriniai ryšiai jungia elementus su aplinka (sistemos išorėje). Mūsų pavyzdyje vidiniai sistemos ryšiai yra ryšiai tarp gydytojų ir seselių arba tarp seselių ir pacientų. Išorinis ryšys, pavyzdžiui, yra kelias, kaip pacientas patenka į ligoninę. Sistemą galime pavaizduoti diagrama (žr. *1 pav.*). Sistema yra įtakojama aplinkos per įėjimą (input). Kai sistema sugeba reaguoti į savo būsenos pasikeitimus, ji vadinama sistema su *grįžtamuoju ryšiu* (feedback). Jei nesugeba, tai be grįžtamojo ryšio. Grįžtamasis ryšys mūsų pavyzdyje galėtų būti toks: kai ligoninėje pacientų skaičius viršija tam tikrą skaičių, ligoninė gali padidinti etatų skaičių.

1 pav. Sistemos grafinis vaizdas

Sistemos elementų atributai apibrėžia jos *būseną* (state). Mūsų pavyzdyje pacientų skaičius ligoninėje apibūdina jos būseną. Kai išvyksta pacientas arba atvyksta naujas, sistema įgyja naują būseną. Jei elementų elgesys negali būti tiksliai numatytas, tai tada reikėtų daryti atsitiktinius stebėjimus ir paimti dydžių tikimybinius vidurkius. Sakoma, sistemos būsena yra *pusiausvyroje* (equilibrium) arba *pastovi* (steady), jei tikimybė būti tam tikroje būsenoje nesikeičia laikui bėgant. Tai reiškia, kad yra sistemoje ryšiai, sistemos būsenos gali keistis, bet perėjimo iš vienos būsenos į kitą tikimybės nekinta, yra fiksuotos. Šios fiksuotos tikimybės yra ribinės tikimybės, kurios nusistovi (stabilizuojasi) po ilgo laiko tarpo ir jos nepriklauso nuo pradinės sistemos būsenos. Sistema vadinama *stabilia*, jei ji grįžta į pusiausvyros būseną, kai patiria išorinį šoką.

Sistemos gali būti klasifikuojamos įvairiai. Yra *natūralios* ir *dirbtinės* sistemos, yra *adaptyviosios* (prisitaikančios) ir *neadaptyviosios* (neprisitaikančios) sistemos. Adaptyviosios sistemos reaguoja į išorinius pasikeitimus, o neadaptyviosios nereaguoja. Tarkime, per tam tikrą laiką padidėjo pacientų skaičius. Jei ligoninė padidina etatų skaičių, tai ligoninė yra adaptyvioji sistema.

1.2 Modeliai

Pirmas etapas tyrinėjant sistemą yra modelio kūrimas. Rozenblatas (A.Rosenbluth) ir Vineris (N.Wiener) rašė:

> Jokia svarbi pasaulio sritis, dalis ar dalelė nėra tokia paprasta, kad galėtų būti suprasta ir kontroliuojama be abstrakcijos. Abstrakcija pakeičia ją panašios bet paprastesnės struktūros modeliu, kurį galima tyrinėti. Taigi modeliai yra mokslinių procedūrų būtinybė.

Mokslinis modelis gali būti apibrėžtas kaip kažkokios realios sistemos abstrakcija, kuri gali būti panaudota sistemos prognozei ir kontrolei. Mokslinio modelio tikslas yra tyrėjo galėjimas atsakyti į klausimą, kaip modeliuojamos sistemos elementų ar atributų pokyčiai įvairiais aspektais įtakoja kitus sistemos aspektus ar visą sistemą.

Esminis žingsnis, kontroliuojant modelį, yra *tikslo funkcijos* sudarymas. Tikslo funkcija tai matematinė funkcija, kurios kintamieji yra sprendimai, kurie gali būti priimti sistemoje.

Yra daug modelių tipų. Keletą išskirkime.

- *Portretiniai modeliai.* Tokie, kurie vaizdžiai aprašo tam tikrus sistemos aspektus.
- *Žodiniai modeliai.*
- *Fiziniai* (*materialūs*) *modeliai.*

• *Analoginiai modeliai.* Tokie, kurie, remdamiesi vienomis sistemos savybėmis, paaiškina (išveda) kitas savybes, kuriomis pasižymi sistema.

• *Simboliniai* (*abstraktūs*, *matematiniai*) *modeliai.* Tokie, kuriems reikalingos matematinės ar loginės operacijos ir jos gali būti panaudojamos formuluojant problemos sprendinį.

Mes nagrinėsime tik simbolinius (abstrakčiuosius siaurąja prasme, matematinius) modelius. Tokie modeliai, lyginant juos su kitais modeliais, turi daug pranašumų.

Pirma, leidžia tyrinėtojui atlikti empirinius sistemos tyrinėjimus, patikrinti teorinius įsitikinimus, daryti tyrinėjimų logines išvadas.

Antra, leidžia geriau suprasti visą sistemą, detaliau įsivaizduoti sistemos perspektyvą ir atnaujinimo reikalingumą, testuoti pageidaujamas sistemos modifikacijas.

Trečia, leidžia lengviau manipuliuoti, palyginus su pačia sistema. Galima kontroliuoti daug daugiau tai ką keičiame, palyginus kai betarpiškai keičiama pati sistema.

Ketvirta, matematiniai modeliai sistemą aprašo glausčiau, palyginus, pavyzdžiui, su žodiniais modeliais.

Penkta, bendrai paėmus, mažiau kainuoja negu pati sitema (sistemos fizinis modelis).

Yra ir keletas išlygų, kurias turime turėti galvoje kurdami modelį.

Pirma, nėra jokių garantijų, kad laikas ir pastangos, skirti modelio kūrimui, duos naudingą rezultatą. Kartais pritrūksta pastangų, juodo darbo, kartais – idėjų.

Antra, tyrėjas supranta problemą tendencingai, individualiai. Dažnai reikia daug pastangų ir praeina daug laiko, kol tyrėjas gali tikėtis gero rezultato.

Trečia, modelio naudingumas (prognoziškumas, pritaikomumas ir pan.) priklauso nuo tyrėjo kvalifikacijos.

Matematiniai modeliai gali būti įvairiai klasifikuojami. Vieni modeliai yra *statiniai*, kiti *dinaminiai*. Statiniai modeliai tiesiogiai nepriklauso nuo laiko, tuo tarpu kai dinaminiai tiesiogiai priklauso nuo laiko. Pavyzdžiui, Omo dėsnis yra statinio modelio pavyzdys, o Niutono judėjimo dėsnis yra dinaminis dėsnis.

Kita modelių klasifikacija būtų: *determinstiniai* ir *stochastiniai* modeliai. Determinstiniame modelyje matematiniai ir loginiai elementų ryšiai nekinta, yra fiksuoti. Iš šių ryšių kaip išvada išplaukia sprendiniai. Stochastiniame modelyje bent vienas kintamasis yra atsitiktinis.

Kad modelis būtų naudingas, jame reikia suderinti du svarbius bet vienas kitam prieštaraujančius dalykus: *realumą* ir *paprastumą*. Iš vienos pusės modelis turi atstovauti realią sistemą, atspindėti svarbiausius jos aspektus.

Iš kitos pusės modelis neturi būti tiek sudėtingas, kad negalėtume jo suprasti ir juo manipuliuoti. Taigi modelis būtinai yra abstrakcija. Kartais galvojame, kad kuo detalesnis modelis, tuo realiau ir tiksliau jis atspindi tikrovę. Bet detalumas padaro patį problemos sprendimą sudėtingu. Dažnai analizinį sprendimą reikia keisti skaitiniu. Naudojami apytikslio skaičiavimo metodai, prarandamas tikslumas, prarandama informacija. Galų gale pasidaro neaišku, ar detalumas davė naudos ar žalos.

Modelyje labai svarbu, kad būtų didelė *koreliacija* (ryšys) tarp prognozės ir to kas realiai atsitiktų su sistema. Norint įsitikinti ar šis reikalavimas yra išpildytas, svarbu modelį *testuoti*. Testavimas pradedamas pasitikrinant problemos formulavimą ir atskleidžiant su tuo susijusius galimus trūkumus. Kitas tikrinimo kriterijus: ar matematinės išraiškos yra dimensiškai tvarkingos. Trečias testas: keičiami įėjimo parametrai ir tikrinama ar išėjimo parametrai keičiasi leidžiamose ribose. Ketvirtas *retrospektyvos* testas: naudojant istorinius duomenis rekonstruojama praeitis ir tikrinama kaip gautas sprendimas atitinka tikrovę.

Kai modelis jau yra sukurtas ir testuotas, gaunamas modelio sprendinys ar sprendiniai. Yra *analiziniai* ir *skaitiniai* sprendiniai. Analiziniai sprendiniai paprastai pateikiami *formule*. Skaitinis sprendinys yra apskritai *apytikslis* ir gaunamas į modelio kintamųjų ir parametrų reikšmes įstačius skaitines reikšmes. Dauguma skaitinių metodų yra *iteraciniai*, t.y. kiekvienas paskesnis žingsnis sprendinyje naudoja ankstesnio žingsnio rezultatus. Specialūs skaitinių metodų tipai yra *modeliavimas* (simulation) ir *Monte-Karlo metodai*.

1.3 Modeliavimas ir Monte-Karlo metodas

Anglų kalboje yra du skirtingi žodžiai: modeling ir simulation. Pirmasis reiškia tai apie ką jau kalbėjome – tai modelio kūrimas. Simulation reiškia sukurto modelio bandymus, eksperimentus, sistemos imitaciją. Šiame skyrelyje kalbėsime apie modeliavimą (simulation) – sukurto modelio bandymą, sistemos imitaciją, naudojant sukurtą modelį.

Modeliavimas (simulation) tai skaičiavimai (bandymai, eksperimentai), dažniausiai naudojant kompiuterines priemones, kurie naudoja matematinius ir loginius modelius, aprašančius verslo, ekonominės ar kitokios sistemos elgesį, ir tie skaičiavimai atliekami realiame laiko tarpe. Apie modeliavimą Neiloras ir kt.[1] rašė (cituojame angliškai):

[1] T.J.Naylor, J.L.Balintfy, D.S.Burdick, K.Chu, Computer Simulation Techniques, Wiley, New York, 1966.

> The fundamental rationale for using simulation is man's unceasing quest for knowledge about the future. This search for knowledge and the desire to predict the future are as old as the history of mankind. But prior to the seventeenth century the pursuit of predictive power was limited almost entirely to the deductive methods of such philosophers as Plato, Aristotle, Euclid and others.

Modeliavimas naudojamas bandant ir *abstrakčiuosius* ir *fizinius* modelius. Į modeliavimo procesą kaip realūs proceso dalyviai gali būti įtraukti žmonės. Du tokio modeliavimo tipus verta paminėti. Tai *veiksmo žaidimai* (operational gaming) ir *žmogus-mašina* (man-machine).

Veiksmo žaidimų tipo modeliavimas charakterizuojamas taip. Modeliuojama tam tikra aplinka. Sudaromos konfliktinės situacijos tarp žaidėjų, priimančių galimus pasirenkamus sprendimus. Eksperimentatorius, bandydamas žaidėjus, gali įvertinti, prognozuoti jų elgesį arba net visos sistemos elgesį. Tokie veiksmo žaidimai plačiai naudojami treniruojant kariškius, verslo vadybos specialistus ir pan.

Žmogaus-mašinos tipo modeliavime nėra jokių žaidimų. Žmogus, bendraudamas su kompiuteriu, sutvarko gautus duomenis ir atlieka jų analizę.

Dabar paminėsime keletą situacijų, kuriose modeliavimas gali būti sėkmingai panaudojamas.

Pirma, kai neįmanoma arba labai brangu gauti duomenis iš tam tikro realaus proceso. Tai gali būti kokia nors ekonominė sistema ar raketa kosmose ir pan. Tokiais atvejais modeliavimo duomenys yra būtini prognozių apie sistemą formulavimui.

Antra, tiriama sistema gali būti tokia sudėtinga, kad aprašyti matematinėmis lygtimis, kurios turi analizinius sprendinius, neįmanoma. Dauguma ekonominių sistemų kaip tik tokios ir yra.

Trečia, kai sistema aprašyta matematiniu modeliu, bet, naudojant analizinę techniką, neįmanoma gauti sprendinio.

Ketvirta, kai neįmanoma arba labai brangu atlikti matematinio sistemos modelio eksperimentus. Tokiais atvejais modeliavimo duomenys gali būti panaudojami testuojant alternatyvias hipotezes.

Visuose minėtuose atvejuose modeliavimas yra vienintelis praktiškas įrankis tinkamiems atsakymams (problemos sprendimams) gauti.

Paminėkime keletą priežasčių, dėl kurių turėtų būti atliekama modeliavimo analizė.

Pirma, galima tyrinėti sistemą ar jos posistemį tik su tam tikrų vidinių ryšių rinkiniu, pasirinktu priklausomai nuo kylamų klausimų ar problemų.

Antra, galima tyrinėti vidinių ir išorinių pasikeitimų įtaką sistemos elgesiui.

Trečia, detalus sistemos modeliavimo nagrinėjimas leidžia geriau suprasti pačią sistemą ir pasiūlyti pagerinti ją kai kuriais aspektais.

Ketvirta, modeliavimas gali būti naudojamas pedagoginiais tikslais, ir teorinių dalykų žinių gilinimui ir praktinių disciplinų mokymui. Ypač taikoma verslo administravime, ekonomikoje, medicinoje, teisėje, karyboje.

Penkta, galima nustatyti kurie sistemos kintamieji yra svarbesni už kitus ir kaip šie kintamieji sąveikauja.

Šešta, kartais galima sudėtingą sistemą suskaidyti į dalis ir modeliuoti tas dalis atskirai, padedant specialistams iš tų sričių, kurioms tos sistemos dalys priklauso.

Septinta, dinamines sistemas galima tyrinėti realiame laike, sutrumpintame (suspaustame) laike, prailgintame (išplėstame) laike.

Kompiuterinis modeliavimas leidžia *kopijuoti* eksperimentą. Kopijavimas reiškia eksperimento pakartojimą, padarius sistemos parametrų ar sąlygų pakeitimus. Kompiuterinis modeliavimas dažnai leidžia nustatyti *koreliaciją* tarp atsitiktinių skaičių sekų ir dėl to pagerinti modeliavimo išėjimo duomenų statistinę analizę. Pavyzdžiui, neigiama koreliacija yra gerai, kai dviejų kopijų rezultatai sudedami, o teigiama yra geriau, kai rezultatai yra atimami ar lyginami.

Daug programavimo kalbų yra naudojama modeliavime. Yra sukurtos ir specialios kalbos modeliavimo tikslams: GPSS, SIMSCRIPT, SIMULA.

Problemose, kuriose analizinė technika yra netinkama, modeliavimas yra neįkainojamas, niekuo nepakeičiamas. Modeliavimas yra *netiksli* technika (priemonė). Jis pateikia ne tikslius rezultatus bet tik *statistinius įverčius*; jis dažnai tik palygina rezultatus, kai parametrai yra skirtingi, bet nesuranda optimalaus. Modeliavimas taip pat yra *lėtas* ir *brangus* problemos nagrinėjimo būdas. Modeliavimas pateikia tik *skaitinius duomenis* apie sistemą, o detali duomenų analizė yra gana brangi.

Mes modeliavimą (simulation) apibrėžėme kaip sistemos modelio *pavyzdžių ėmimą* (sampling experiments). Šis gana bendras apibrėžimas dažnai naudojamas modeliavimui *plačiąja prasme*. Modeliavimas *siaurąja prasme* arba *stochastinis modeliavimas* apibrėžiamas kaip eksperimentavimas su modeliu laike; jis naudojasi stochastinių atsitiktinių dydžių parinkimu.

Stochastinis modeliavimas dar vadinamas ir *Monte-Karlo modeliavimu* arba modeliavimu *Monte-Karlo metodu*. Bet Monte-Karlo modeliavimas ir stochastinis modeliavimas suprantami šiek tiek skirtingai. Stochastinis modeliavimas yra bendresnė sąvoka. Taigi Monte-Karlo metodas tai priemonės, kai naudojami atsitiktiniai ar pseudoatsitiktiniai skaičiai. Jie gali būti naudojami modeliuojant sistemą ar ieškant sistemos modelio sprendinio, ar dar kaip nors kitaip. Visas procesas susideda iš trijų dalių: pirma, atsitiktinio dydžio su duotu skirstiniu modeliavimas; antra, realios sistemos tikimybinio mod-

elio sudarymas; trečia, statistinės įvertinimo teorijos uždaviniai, leidžiantys įvertinti sudarytą uždavinį įvairiais aspektais, atlikti modelio statistinę analizę.

Atsitiktinių procesų modeliavimo idėja labai sena ir, kai kurių autorių (pavyzdžiui Haltono[2]) nuomone, siekia net Senovės Babilono ir Senojo Testamento laikus.

Vienas iš seniausių uždavinių, susijusių su Monte-Karlo metodu, yra įžymusis Biufono (G.L.L.Buffon) uždavinys. l ilgio adata yra atsitiktinai metama ant popieriaus lapo, padalinto lygiagrečiomis tiesėmis į d pločio juostas. Tikimybė, kad adata kirs ar lies tiesę, yra lygi $2l/\pi d$. Ši tikimybė gali būti panaudota apytiksliam skaičiaus π radimui. Šį uždavinį jau 1873 m. aprašė Holas[3].

Praeito šimtmečio pradžioje Monte-Karlo metodas buvo naudojamas Bolcmano (L.Bolzmann) lygčiai tyrinėti. 1908 m. įžymus statistikas Stjudentas (Student – W.S.Gosset) naudojo Monte-Karlo metodą savo t-skirstinio koreliacijos koeficientui įvertinti.

Pats Monte-Karlo pavadinimas įvestas Noimano (J.von Neumann) ir Ulamo (S.Ulam) Antrojo pasaulinio karo metu. Buvo kuriama atominė bomba. Sprendžiami atominės fizikos uždaviniai. Juose buvo naudojamas stochastinis modeliavimas. Užslaptinti darbai buvo pavadinti Monte-Karlo (Monako miestas, garsus savo lošimo namais) vardu, ir tas vardas prigijo.

Monte-Karlo metodas gali būti naudojamas ne tik stochastiniams bet ir deterministiniams uždaviniams spręsti. Deterministinis uždavinys gali būti sprendžiamas Monte-Karlo metodu, jei jo formali išraiška yra tokia pati kaip kokio nors stochastinio proceso. Naudojant Monte-Karlo metodą skaičiuojami daugialypiai integralai, masinio aptarnavimo uždavinių parametrai, sprendžiamos integralinės ir diferencialinės lygtys.

Monte-Karlo metodas yra nepakeičiamas, kai reikia spręsti sudėtingus uždavinius. O šiandien realybė tokia, kad uždaviniai yra labai sudėtingi, priklausantys nuo gausybės parametrų. Monte-Karlo metodo pritaikymų sritis vis plečiasi.

[2] J.H.Halton, A retrospective and prospective survey of the Monte-Carlo method, SIAM Rev.(1970) **12**, No 1, 1-63.

[3] A.Hall, On an experiment determination of π, Messeng.Math.(1873) **2**, 113-114.

2 ATSITIKTINIAI SKAIČIAI

Monte-Karlo metodas naudoja atsitiktinius skaičius (toliau naudosime trumpinį a.s.), tiksliau atsitiktinius dydžius, kurie pasiskirstę pagal reikiamą pasiskirstymo dėsnį. Kadangi iš nepriklausomų, tolygiai pasiskirsčiusių intervale $[0, 1]$ a.s. galima sukonstruoti kitokius a.s., tai kalbėsime apie a.s. iš intervalo $[0, 1]$.

Panagrinėkime trijų tipų a.s.: *tikruosius atsitiktinius skaičius*, *pseudoatsitiktinius skaičius* ir *kvaziatsitiktinius skaičius*.

2.1 Tikrieji atsitiktiniai skaičiai

Tikrieji a.s. yra atsitiktiniai statistine prasme. Jokia a.s. serija nepriklauso nuo anksčiau gautų a.s. Tikrųjų a.s. serijos nepasikartoja antrą kartą. Jie negali būti atspėjami.

Tikrieji a.s. gali būti generuoti atsitiktinio fizinio proceso metu. Pavyzdžiui, su ruletės pagalba (žr. *2 pav.*). Tarkime, ruletės rodyklė gali parodyti skaitmenis $0, 1, 2, 3, 4, 5, 6, 7, 8, 9$ su vienodomis tikimybėmis, lygiomis $1/10$. Norimo tikslumo a.s. iš intervalo $[0, 1]$ gausime taip: pirma vieta po kablelio – pirmas ruletės nurodytas skaičius, antra vieta – skaičius po antro ruletės pasukimo ir t.t. Galima generuoti tikruosius a.s. mėtant monetą. Šiuo atveju būtų patogu naudoti skaičiaus dvejetainį užrašą. Generuoti a.s. galima panaudoti ir radioaktyviųjų medžiagų skilimo procesą, skaičiuojant skilimų skaičių per laiko vienetą.

2 pav. Atsitiktinių skaičių generavimas su ruletės pagalba

Anksčiau mokslininkai, kuriems reikėjo a.s., konstruodavo juos maišydami kortas, mesdami lošimo kauliuką arba traukdami iš dėžės rutulius,

prieš tai juos gerai sumaišę. Vėliau buvo sukonstruotos specialios mašinos, su kurių pagalba buvo gaunami a.s. Pirmą tokią mašiną 1939 m. panaudojo Kendalas (M.G.Kendall) ir Babingtonas-Smitas (B.Babington-Smith). Jie sudarė 100 000 a.s. Dar vėliau buvo sugalvota įvairių a.s. gavimo mechanizmų, ir gana greitų. Sukurti elektroniniai a.s. generatoriai (pavyzdžiui, ERNIE) buvo net tiesiogiai prijungiami prie kompiuterio.

Tikrųjų a.s. trūkumai:

• Jų generavimas reikalauja specialių priemonių.

• Juos generuoti reikia daug laiko.

• Jų negalima pakartoti. O modeliavime kartais prireikia pakartoti eksperimentą su ta pačia a.s. seka.

• Generatoriai gali turėti sisteminių (konstravimo ir pan.) klaidų. Gali pasirodyti, kad generuoti skaičiai jau ne visai atsitiktiniai, t.y. nėra tikrieji a.s.

2.2 Pseudoatsitiktiniai skaičiai

Pseudoatsitiktiniai skaičiai (naudosime trumpinį p.a.s.) yra generuoti su kokio nors algoritmo pagalba. Taigi kiekvienas paskesnis skaičius priklauso nuo ankstesnių skaičių. Bet ši priklausomybė tokia, kad p.a.s. turi tas pačias svarbias statistines savybes kaip ir tikrieji a.s.

Aišku, tos savybės negali būti visiškai tos pačios. Bet bet kurios neilgos p.a.s. sekos statistinės savybės daugeliu aspektų turi būti labai panašios į tikrųjų a.s. savybes. Taikymuose paprastai to ir pakanka, jei prieš einantis skaičius pakankamai sujaukiamas.

Reikalavimai, keliami a.s. generatoriams:

• Generuoti skaičiai turi būti tolygiai pasiskirstę intervale $[0, 1]$, nes taip pasiskirstę yra tikrieji a.s. Kitaip pasiskirstę a.s. gali būti gauti iš pastarųjų.

• Generuoti skaičiai turi būti statistiškai nepriklausomi, kadangi atsitiktinėje sekoje vieno skaičiaus reikšmė neturi turėti įtakos kito skaičiaus reikšmei.

• Generuojamų skaičių seka turi būti atstatoma. Tai leidžia kartoti eksperimentus.

• Bet kokio norimo ilgio seka neturi kartotis. Tai teoriškai neįmanoma, bet praktinėms reikmėms užtenka ilgų besikartojančių ciklų.

• Skaičių generavimas turi būti greitas. Modeliavimo procese paprastai rekia daug p.a.s. Jei generatorius lėtas, tai modeliavimo procesui gali prireikti labai daug laiko, ir todėl pats modeliavimas pasidarys brangus.

• P.a.s. generatoriuje naudojamas metodas turi naudoti kiek galima mažiau atminties. Pats modeliavimas paprastai užima daug atminties, o atmintis yra ribota.

Kompiuterio atmintyje laikyti generuotą p.a.s. seką nėra tikslo. Reikėtų sugaišti daug laiko nuskaitant p.a.s., užimtų daug atminties vietos. Be to, priklausomai nuo sprendžiamų uždavinių, reikėtų daug p.a.s. pavyzdžių. Geriausia kompiuteryje turėti programą, kuri, naudodama tam tikrus algoritmus, generuotų p.a.s. tada, kai jų prireikia pačiame modeliavimo procese.

Kai naudojami p.a.s., reikia ypač būti įsitikinusiam, kad p.a.s. seka sprendžiamam uždaviniui yra tikrai tinkama, t.y. pakankamai atsitiktinė, pakankamo ilgio nepasikartojantis ciklas ir kt. P.a.s. yra tokie svarbūs modeliavime, kad daug darbo turi būti atlikta testuojant p.a.s. algoritmus. Apie šiuos dalykus pakalbėsime kituose skyreliuose.

2.3 Kvaziatsitiktiniai skaičiai

Tai visiškai neatsitiktiniai skaičiai. Ilgos tokių skaičių sekos kai kuriems uždaviniams spręsti yra geriau už tikruosius a.s.

Yra daug skaičiavimų, naudojančių Monte-Karlo metodą, kurių rezultatai indiferentiški gretimų a.s. koreliacijai. Kvaziatsitiktiniai skaičiai svarbūs integruojant Monte-Karlo metodu. Jame naudojant a.s., kurie yra atsitiktiniai statistine prasme, rezultato klaida yra proporcinga $1/\sqrt{N}$. Naudojant kvaziatsitiktinius skaičius, gaunamas rezultatas, kurio klaida proporcinga $1/N$. Čia N – a.s. skaičius.

Yra kvaziatsitiktinių skaičių sekų, kurių gretimų skaičių skirtumai yra pastovūs. Be jokios abejonės, koreliacija tarp šių skaičių yra, bet jie pasiskirstę tolygiau už tikruosius a.s.

2.3.1 Richtmajerio formulė

i-asis a.s. j-ojoje serijoje paskaičiuojamas pagal formulę[4]:

$$r_{ij} = iS_j \mod 1.$$

Čia S_j yra kvadratinė šaknis iš j-ojo pirminio skaičiaus. Taigi skirtumas tarp dviejų gretimų skaičių yra S_j. Tai reiškia, kad koreliacija stipri.

2.3.2 Van der Korputo formulė

Šiame skyrelyje pateiksime dar vieną kvaziatsitiktinių skaičių gavimo algoritmą[5].

Procedūra tokia. Imami iš eilės einantys natūralieji b-tainės skaičiavimo sistemos skaičiai. Pirmasis skaičius parenkamas. Po to skaitmenų tvarka

[4] Ši formulė vadinama Richtmajerio (R.D.Richtmyer) formule.

[5] Algoritmas vadinamas van der Korputo (J.G.van der Corput) vardu.

skaičiuose pakeičiama: skaitmenys surašomi priešinga tvarka. Padaromos trupmenos, pridedant kairėje trupmenos tašką. Pagaliau gautos trupmenos paverčiamos dešimtainėmis. *3 pav.* pateiktas pavyzdys, kai naudojama dvejetainė skaičiavimo sistema ir pirmas parinktas skaičius yra lygus 1.

Dešimtainis skaičius	Dvejetainis skaičius	Dvejetainė trupmena	Dešimtainė trupmena
1	1	0.1	0.5
2	10	0.01	0.25
3	11	0.11	0.75
4	100	0.001	0.125
5	101	0.101	0.625
6	110	0.011	0.375
7	111	0.111	0.875
8	1000	0.0001	0.0625
...	...	...	...

3 pav. Van der Korputo formulė

3 PSEUDOATSITIKTINIŲ SKAIČIŲ GENERAVIMAS

P.a.s. gaunami skaitinių algoritmų pagalba. Bendriausias p.a.s. gavimo algoritmas turi formą:

$$x_{n+1} = f(x_1, x_2, \ldots, x_n).$$

Naudojant tokį algoritmą, atmintyje reikėtų laikyti visus generuotus skaičius, pradedant pirmuoju. Tokia procedūra užimtų per daug kompiuterio atminties.

Praktinėms reikmėms paprastai užtenka paprastesnių algoritmų. Dauguma jų turi tokią formą:

$$x_{n+1} = f(x_n). \tag{1}$$

Funkcija f turi būti labai atidžiai parinkta. Funkcija, pavaizduota *4 pav.*, aiškiai bloga. Taškai, kurių koordinatės yra gretimi skaičiai

$$(x_1, x_2),\ (x_3, x_4),\ (x_5, x_6),\ \ldots,$$

yra kreivės taškai, ir aišku nėra tolygiai pasiskirstę vienetiniame kvadrate $\{(x, y) | 0 \leq x, y < 1\}$. (1) funkcijos grafikas turi padengti vienetinį kvadratą kiek galima tolygiau (žr. *5 pav.*). Tegul

$$y = \{gx\}; \tag{2}$$

čia g yra didelis skaičius, o skaičiaus z trupmeninė dalis yra žymima $\{z\}$. Iš *5 pav.* matyti, kad (2) funkcijos, kai $g = 19$, grafikas vienetinį kvadratą padengia tolygiai.

4 pav. Ši funkcija vienetinį kvadratą uždengia netolygiai

5 pav. Ši funkcija vienetinį kvadratą uždengia tolygiai

3.1 Kvadrato vidurio metodas

Kvadrato vidurio metodas tai Noimano pasiūlytas ir pirmas plačiai naudotas algoritminis metodas p.a.s. generuoti. Metodo idėja tokia. n-asis p.a.s. yra gaunamas paėmus vidurinius $(n-1)$-ojo p.a.s. kvadrato skaitmenis. Žiūrėkite *6 pav.*

Kvadrato vidurio metodas nėra geras. Yra skaičių, kurie užsiciklina. Forsaitas (G.E.Forsythe) patikrino 16 keturženklių skaičių. 12 iš jų baigėsi ciklu 6100, 2100, 4100, 8100, 6100, ... Dvi sekos tapo nulinėmis. Žiūrėkite *7 pav.* Metropolis (N.Metropolis), tyrinėdamas šį metodą, ėmė dvejetainius 20 ženklų skaičius. Jis parodė, kad egzistuoja 13 ciklų. Didžiausio iš jų periodas 142. Tyrinėdamas 38 ženklų dvejetainius skaičius, surado 750 000 periodo seką.

$\gamma_0 = 0.1234$	$\gamma_0^2 = 0.01522756$	
$\gamma_1 = 0.5227$	$\gamma_1^2 = 0.27321529$	
$\gamma_2 = 0.3215$	$\gamma_2^2 = 0.10336225$	$24\underline{13}$
$\gamma_3 = 0.3362$	$\gamma_3^2 = 0.11303044$	$1\underline{68}1$
$\gamma_4 = 0.3030$	$\gamma_4^2 = 0.09180900$	$4\underline{62}4$
$\gamma_5 = 0.1809$	$\gamma_5^2 = 0.03272481$	$3\underline{84}4$
$\gamma_6 = 0.2724$	$\gamma_6^2 = 0.07420176$	$7\underline{05}6$
$\gamma_7 = 0.4201$	$\gamma_7^2 = 0.17648401$	$0\underline{02}5$
$\gamma_8 = 0.6484$	$\gamma_8^2 = 0.42042256$	$0\underline{00}4$
$\gamma_9 = 0.0422$	$\gamma_9^2 = 0.00178084$	$0\underline{00}0$
.....	.....	.....

6 pav. Noimano kvadrato vidurio metodas

7 pav. Bloga kvadrato vidurio metodu gauta seka

Užfiksavus tokių skaičių ciklų pabaigą, galima procesą pradėti iš naujo. Floidas (R.Floyd) sugalvojo metodą, kaip tai padaryti. Šis metodas reikalauja nedaug mašinos atminties ir triskart daugiau laiko a.s. generuoti.

Taigi Noimano metodas nėra geras. Pirma, jis labai nepatogus statistinei analizei atlikti. Antra, sekos linkę užsiciklinti; seka labai priklauso nuo pirmojo skaičiaus parinkimo; jei sekoje pasitaiko nuliai, seka baigiasi. Trečia, skaičiai generuojami ne pakankamai greitai.

Atsiradus geresniems generatoriams, o ypač pradėjus naudoti tiesinį kongruentinį metodą, kvadrato vidurio metodas nebenaudojamas. Jis liko istorijoje kaip pirmas algoritminis p.a.s. generatorius.

3.2 Tiesinis kongruentinis metodas

Šiandien p.a.s. dažniausiai gaunami taikant Lemerio (D.H.Lehmer) 1948 m. pasiūlytos schemos dalinius atvejus. Metodas vadinamas *tiesiniu kongruentiniu metodu*. Šiame skyrelyje jį ir panagrinėsime.

P.a.s. seka apibrėžiama lyginiu:

$$X_{n+1} \equiv (aX_n + c) \bmod m, \quad n \geq 0. \tag{3}$$

Čia skaičiai:

$$\begin{array}{lll} X_0 & - \textit{pradinė reikšmė}, & X_0 \geq 0, \\ a & - \textit{daugiklis}, & a \geq 0, \\ c & - \textit{prieauglis}, & c \geq 0, \\ m & - \textit{modulis}, & m > X_0,\ m > a,\ m > c, \end{array} \tag{4}$$

yra parenkami.

Tarkime, $X_0 = a = c = 7,\ m = 10$. Gausime seką

$$7,\ 6,\ 9,\ 0,\ 7,\ 6,\ 9,\ 0, \ldots .$$

Gauta seka nėra gera.

Toliau panagrinėsime principus, kaip gauti kuo ilgesnio ciklo ("geras") atsitiktines sekas priklausomai nuo pradinių parametrų. Pasikartojanti p.a.s. sekos dalis vadinama jos *periodu*. Tislas – gauti kuo ilgesnio periodo sekas. Kai $c = 0$, tiesinės sekos gavimo metodas vadinamas *multiplikatyviuoju*. Multiplikatyviuoju atveju p.a.s. generavimo procesas vyksta greičiau. Apribojimas $c = 0$ sumažina sekos periodo ilgį, bet ir šiuo atveju galima gauti gana ilgo periodo sekas. Taip pat žymėsime $b = a - 1$ (bus patogiau). Atvejų, kai $a = 0$ ir $a = 1$, nenagrinėsime, nes sekos gaunasi labai skurdžios ir, aišku, mažai "atsitiktinės". Taigi turėsime galvoje, kad $a \geq 2$, $b \geq 1$.

Iš (3) lyginio turime, kad

$$\begin{aligned} X_{n+k} &\equiv aX_{n+k-1} + c \equiv a(aX_{n+k-2} + c) + c \\ &= a^2 X_{n+k-2} + c(a+1) \equiv a^2(aX_{n+k-3} + c) + c(a+1) \\ &= a^3 X_{n+k-3} + c(a^2 + a + 1) \equiv \cdots \equiv a^k X_n + c(a^{k-1} + a^{k-2} + \cdots + 1) \\ &= a^k X_n + \frac{c(a^k - 1)}{a - 1} \mod m \end{aligned}$$

arba

$$X_{n+k} \equiv a^k X_n + \frac{c(a^k - 1)}{b} \mod m, \quad k \geq 0,\ n \geq 0.$$

Paėmę $n = lk$, gausime

$$X_{(l+1)k} \equiv a^k X_{lk} + \frac{c(a^k - 1)}{b} \mod m, \quad l \geq 0. \tag{5}$$

Taigi seka $X_0, X_k, X_{2k}, \ldots$ – tai nauja tiesinė kongruentinė seka su daugikliu a^k ir prieaugliu $c(a^k - 1)/b$.

3.2.1 Modulio parinkimas

Kadangi periodas negali būti didesnis už m, tai m reikėtų imti gana didelius. Netgi jei mums reikia atsitiktinės sekos iš nuliukų ir vienetukų, nereikia imti $m = 2$, nes šiuo atveju daugiausiai gausime $\ldots, 0, 1, 0, 1, \ldots$ arba blogiau, tik vienetukus arba tik nuliukus.

Kitas faktorius, nulemiantis m parinkimą, tai sekos elementų paskaičiavimo greitis. Skaičiuojant kompiuteriu, patogu imti m, lygų žodžio ilgiui (vienetu daugiau negu kompiuterio žodyje telpantis didžiausias sveikas

skaičius). Tegul w – toks maksimalus sveikas skaičius. Šiuo atveju labai paprasta atlikti operacijas moduliu w, nes rezultatas gaunamas paskutinėse žodžio skiltyse ir kas netelpa tiesiog galima išstumti kairėn.

Bet šitas metodas galima taikyti ne visuomet. Jis nėra labai geras. Ir štai kodėl. Paskutiniai skaičiaus X_n skaitmenys yra daug mažiau "atsitiktiniai" negu pirmieji. Sakykime,

$$d|w \quad \text{ir} \quad Y_n \equiv X_n \bmod d.$$

Tada Y_n yra paskutinieji skaičiaus X_n skaitmenys. Kadangi

$$X_{n+1} \equiv aX_n + c \bmod m_1 d,$$

tai

$$Y_{n+1} \equiv X_{n+1} \equiv aX_n + c \equiv aY_n + c \bmod d.$$

Vadinasi paskutiniųjų skaitmenų periodas neilgesnis kaip d. Pavyzdžiui, jeigu $w = 2^l$, tai paskutiniojo X_n skaitmens periodas lygus tik 2 (jeigu kodas dvejetainis). Taigi periodiškai keičiasi 0 su 1, arba tik 0, arba tik 1. Paskutiniųjų dviejų skaitmenų periodas būtų $2^2 = 4$, paskutinių trijų – $2^3 = 8$ ir t.t.

Situacija visai kita, kai vietoje m pasirenkame didžiausią pirminį skaičių, mažesnį už w. Čia problemos iškyla (tiesa, jos nedidelės, nugalimos) su to pirminio skaičiaus suradimu ir su sekos elementų greitu paskaičiavimu. Kuo tas pirminis skaičius artimesnis w, tuo greitesni skaičiavimai.

Pakanka imti $m = w \pm 1$, ir situacija taip pat žymiai pagerėja. Paprastai laikoma $c = 0$. Kaip gaunamas algoritmas? Sakykime, $m = w+1$. Visuomet galima užrašyti:

$$aX = qw + r, \qquad 0 \le q, r < w.$$

Paprastai dauginant a iš X, liekana r užims vieną žodį, o q – kitą žodį. Bet

$$aX = q(w+1) + r - q.$$

Taigi

$$aX \equiv \begin{cases} r - q & \bmod (w+1), \quad \text{jei } r - q \ge 0, \\ r - q + w + 1 & \bmod (w+1), \quad \text{jei } r - q < 0. \end{cases}$$

Užrašytas procedūras atlikti su kompiuteriu labai paprasta. Kai $aX = w$, nesunku numatyti programoje ir išmesti tokią reikšmę. Šiuo atveju gauname perpildymą (skaičius w netelpa viename žodyje). Taigi aišku, kaip skaičiuoti seką

$$X_{n+1} \equiv aX_n \bmod (w+1).$$

3.2.2 Daugiklio parinkimas

Šiame skyrelyje parodysime kaip parinkti a, kad gautume maksimalaus periodo p.a.s. Didelis periodas – tai tik vienas iš būtinų atsitiktinumo elementų. Pavyzdžiui, kai $a = c = 1$, turėsime maksimalaus ilgio visiškai neatsitiktinę seką

$$X_{n+1} \equiv X_n + 1 \bmod m.$$

Kadangi seka gali įgyti tik m skirtingų reikšmių, tai maksimalaus periodo ilgis yra, aišku, nedidesnis kaip m. Kad tokio ilgio gali būti, rodo ką tik pateiktas pavyzdys.

Pastaba. *Kai periodo ilgis maksimalus* $(= m)$*, kiekvienas skaičius nuo* 0 *iki* $m - 1$ *periode sutinkamas vieną kartą. Po to periodas pasikartoja. Todėl visiškai nesvarbu kokį* X_0 *bepasirinksime.*

Pateiksime reikalingus faktus iš skaičių teorijos.

1 lema *Tegul* $p \in \mathbb{P}$, $l \in \mathbb{N}$, $p^l > 2$. *Jeigu*

$$x \equiv 1 \bmod p^l, \quad x \not\equiv 1 \bmod p^{l+1},$$

tai

$$x^p \equiv 1 \bmod p^{l+1}, \quad x^p \not\equiv 1 \bmod p^{l+2}.$$

Įrodymas. Turime

$$x = qp^l + 1, \quad (q, p) = 1.$$

Iš čia

$$x^p = q^p p^{pl} + C_p^1 q^{p-1} p^{(p-1)l} + \cdots + C_p^{p-1} q p^l + 1.$$

Niutono binomo koeficientas

$$C_p^k = \frac{p(p-1)\ldots(p-k+1)}{k!}$$

yra natūralusis skaičius. Kadangi $k < p$, tai po suprastinimo būtinai liks pirminis daugiklis p. Vadinasi $p|C_p^k$, ir todėl

$$x^p = 1 + qp^{l+1}\left(1 + \frac{1}{p} C_p^{p-2} q p^l + \frac{1}{p} C_p^{p-3} q^2 p^{2l} + \ldots\right.$$

$$\left. + \frac{1}{p} C_p^1 q^{p-2} p^{(p-2)l} + \frac{1}{p} q^{p-1} p^{(p-1)l}\right).$$

Reiškinio skliaustuose kiekvienas dėmuo, išskyrus pirmąjį, yra p kartotinis. Paskutinis dėmuo dalijasi iš p, nes $p^l > 2$ ir todėl $(p-1)l > 1$. Taigi

$$x^p = 1 + q' p^{l+1}, \quad (q', p) = 1.$$

Lema įrodyta. ◄

2 lema *Tegul*

$$m = p_1^{l_1} \dots p_t^{l_t}.$$

Tiesinės kongruentinės sekos (X_0, a, c, m) *periodas* λ *yra lygus sekų*

$$\left(X_0 \bmod p_j^{l_j}, a \bmod p_j^{l_j}, c \bmod p_j^{l_j}, p_j^{l_j}\right), \quad 1 \le j \le t,$$

periodų λ_j *mažiausiam bendram kartotiniui.*

Įrodymas. Pakanka įrodyti lemą, kai $(r, s) = 1$, λ – sekos (X_0, a, c, rs) periodas, λ_1, λ_2 – sekų $(X_0 \bmod r, a \bmod r, c \bmod r, r)$, $(X_0 \bmod s, a \bmod s, c \bmod s, s)$ periodai. Pasinaudojus indukcija, iš to jau išplauktų lema.

Pažymėkime minėtų sekų elementus atitinkamai X_n, Y_n, Z_n. Kadangi $Y_0 \equiv X_0 \bmod r$, tai

$$Y_1 \equiv aY_0 + c \equiv aX_0 + c \equiv X_1 \bmod r.$$

Pasinaudoję indukcija, gautume, kad $Y_n \equiv X_n \bmod r$. Analogiškas lyginys teisingas ir sekai Z_n. Taigi

$$Y_n \equiv X_n \bmod r, \quad Z_n \equiv X_n \bmod s, \quad \forall n. \tag{6}$$

Įrodysime, kad

$$X_n = X_k \iff Y_n = Y_k \text{ ir } Z_n = Z_k. \tag{7}$$

Iš tikrųjų, tegul $X_n = X_k$, tada iš (6) gausime, kad $Y_n = Y_k$ ir $Z_n = Z_k$. Tegul dabar $Y_n = Y_k$ ir $Z_n = Z_k$. Iš (6) turėsime, kad

$$X_n = X_k + ru = X_k + sv, \quad ru = sv,$$

ir kadangi $(r, s) = 1$, tai $u = su_1$, o $v = rv_1$. Taigi $X_n = X_k + rsu_1$. Kadangi visuomet imame $\bmod rs$, tai iš čia išplaukia, kad $X_n = X_k$. (7) įrodyta.

Tegul $\lambda' = MBK(\lambda_1, \lambda_2)$. Kai n pakankamai didelis, $n \ge \mu$, turime $X_n = X_{n+\lambda}$, o iš (7) $Y_n = Y_{n+\lambda}$ ir $Z_n = Z_{n+\lambda}$. Taigi λ yra λ_1 ir λ_2 kartotinis. Vadinasi $\lambda \ge \lambda'$. Iš kitos pusės $Y_n = Y_{n+\lambda'}$ ir $Z_n = Z_{n+\lambda'}$ visiems pakankamai dideliems n. Iš (7) išplaukia, kad $X_n = X_{n+\lambda'}$ visiems pakankamai dideliems n. Taigi $\lambda' \ge \lambda$. Gavome, kad $\lambda = \lambda'$. Lema įrodyta. ◀

3 lema (Mažoji Ferma teorema, 1640) *Tegul* $p \in \mathbb{P}$. *Tuomet*

$$a^p \equiv a \bmod p.$$

Įrodymas. Jeigu $a \equiv 0 \bmod p$, tai lemos tvirtinimas akivaizdus.

Tegul $a \not\equiv 0 \bmod p$. Tuomet $(a,p) = 1$ ir ax, $x = 0, 1, \ldots, p-1$, perbėga visas likinių klases $\bmod p$. Iš tikrųjų, jei ax_1 ir ax_2 priklauso tai pačiai likinių klasei, tai $ax_1 = k_1 p + l$, $ax_2 = k_2 p + l \implies a(x_1 - x_2) = (k_1 - k_2)p \implies x_1 - x_2 = kp \implies x_1 = x_2$. Taigi seka

$$0 \bmod p,\ a \bmod p,\ \ldots,\ (p-1)a \bmod p$$

sudaryta iš skirtingų skaičių: $0, 1, \ldots, p-1$. Todėl

$$a \cdot 2a \ldots (p-1)a \equiv 1 \cdot 2 \ldots (p-1) \bmod p.$$

Pastarąjį lyginį padauginę iš a, gausime

$$a^p((p-1)!) \equiv a((p-1)!) \bmod p,$$

arba

$$a^p \equiv a \bmod p.$$

Lema įrodyta. ◄

4 lema *Tegul $p \in \mathbb{P}$. Tuomet*

$$a^{p^l} \equiv a \bmod p.$$

Įrodymas. Iš 3 lemos išplaukia, kad $a^p = a + pt$. Todėl

$$a^{p^2} = (a^p)^p = a^p + C_p^1 a^{p-1} pt + \cdots + C_p^{p-1} a(pt)^{p-1} + (pt)^p \equiv a^p \equiv a \bmod p.$$

Panaudoję indukciją gausime lemos įrodymą. ◄

5 lema *Tegul $1 < a < p^l$, $p \in \mathbb{P}$, o λ – mažiausias natūralusis skaičius, kuriam*

$$\frac{a^\lambda - 1}{a - 1} \equiv 0 \bmod p^l.$$

Tuomet

$$\lambda = p^l \iff \begin{cases} a \equiv 1 \bmod p, & \text{kai } p > 2, \\ a \equiv 1 \bmod 4, & \text{kai } p = 2. \end{cases}$$

Įrodymas. Būtinumas. Tegul $\lambda = p^l$. Jeigu $a \not\equiv 1 \bmod p$, tai

$$\frac{a^n - 1}{a - 1} \equiv 0 \bmod p^l \iff a^n - 1 \equiv 0 \bmod p^l.$$

Jeigu

$$a^{p^l} \equiv 1 \bmod p^l,$$

tai ir

$$a^{p^l} \equiv 1 \bmod p.$$

Iš 4 lemos turėsime

$$a^{p^l} \equiv a \bmod p.$$

O iš paskutiniųjų dviejų lyginių gausime

$$a \equiv 1 \bmod p.$$

Gauta prieštara įrodo, kad

$$a \equiv 1 \bmod p.$$

Jeigu $p = 2$ ir $a \equiv 3 \bmod 4$, tai

$$a - 1 = 2(2n - 1). \tag{8}$$

Be to

$$a^2 = (4n + 3)^2 = 16n^2 + 24n + 9 \equiv 1 \bmod 8.$$

Kadangi

$$x \equiv 1 \bmod 2^l \quad \Longrightarrow \quad x^2 = (n2^l + 1)^2 = n^2 2^{2l} + n2^{l+1} + 1 \equiv 1 \bmod 2^{l+1},$$

tai

$$a^4 \equiv 1 \bmod 16,\ a^8 \equiv 1 \bmod 32,\ \ldots,\ a^{2^{l-1}} \equiv 1 \bmod 2^{l+1};$$

$$a^{2^{l-1}} - 1 \equiv 0 \bmod 2^{l+1},$$

$$\frac{a^{2^{l-1}} - 1}{2} \equiv 0 \bmod 2^l.$$

Iš (8) turime, kad $2|(a-1)$ ir $4 \nmid (a-1)$. Todėl

$$\frac{a^{2^{l-1}} - 1}{a - 1} \equiv 0 \bmod 2^l.$$

Gavome prieštarą λ apibrėžimui. Taigi, kai $\lambda = p^l$, turėsime, kad

$$a = 1 + qp^t, \quad p^t > 2, \quad (q, p) = 1.$$

Pakankamumas. Tegul

$$a = 1 + qp^t, \quad p^t > 2. \tag{9}$$

Iš 1 lemos turėsime

$$\begin{array}{ll} a^p \equiv 1 \bmod p^{t+1}, & a^p \not\equiv 1 \bmod p^{t+2}; \\ a^{p^2} \equiv 1 \bmod p^{t+2}, & a^{p^2} \not\equiv 1 \bmod p^{t+3}; \\ \cdots\cdots\cdots & \cdots\cdots\cdots \\ a^{p^s} \equiv 1 \bmod p^{t+s}, & a^{p^s} \not\equiv 1 \bmod p^{t+s+1}. \end{array}$$

Iš pastarųjų lyginių ir (9) gausime

$$a^{p^s} - 1 \equiv 0 \bmod p^{t+s}, \quad a - 1 \equiv 0 \bmod p^t.$$

Dabar aišku, kad

$$\frac{a^{p^s} - 1}{a - 1} \equiv 0 \bmod p^s, \quad \frac{a^{p^s} - 1}{a - 1} \not\equiv 0 \bmod p^{s+1}.$$

Vietoje s įstatome l. Tuomet

$$\frac{a^{p^l} - 1}{a - 1} \equiv 0 \bmod p^l, \quad \frac{a^{p^l} - 1}{a - 1} \not\equiv 0 \bmod p^{l+1}.$$

Imkime tiesinę kongruentinę seką $(0, a, 1, p^l)$. Iš (5) formulės, paėmę $l = 0$ ir $k = n$, gausime, kad

$$X_n \equiv \frac{a^n - 1}{a - 1} \bmod p^l.$$

Pagal λ apibrėžimą, šios sekos periodo ilgis lygus λ, t.y. mažiausiam skaičiui, kuriam

$$\frac{a^\lambda - 1}{a - 1} \equiv 0 \bmod p^l,$$

nes tik šiuo atveju vėl pasikartos X_0. Jei λ – periodas, tai jis turi dalyti p^l, nes ir p^l periodas. Taigi $\lambda = p^s$.

Jeigu $s < l$, tai

$$\frac{a^{p^s} - 1}{a - 1} \equiv 0 \bmod p^l,$$

bet

$$\frac{a^{p^s} - 1}{a - 1} \not\equiv 0 \bmod p^{s+1},$$

o $l \geq s + 1$. Ši prieštara įrodo, kad

$$\lambda = p^l.$$

Lema įrodyta. ◀

1 teorema *Tiesinės kongruentinės sekos periodo ilgis lygus* $m \iff$

- $(c, m) = 1$,
- $p|m \quad \Rightarrow \quad p|(a - 1)$,
- $4|m \quad \Rightarrow \quad 4|(a - 1)$.

Įrodymas. Iš 2 lemos išplaukia, kad teoremą pakanka įrodyti, kai $m = p^l$.
Kai $a = 1$, teoremos įrodymas akivaizdus. Šiuo atveju

$$X_1 \equiv X_0+c \bmod m,\ \ X_2 \equiv X_0+2c \bmod m,\ \ \ldots,\ \ X_m \equiv X_0+mc \equiv X_0 \bmod m.$$

Kadangi $(c, m) = 1$, tai sandaugos ct, $t = 0, \ldots, m-1$, perbėga visą likinių klasę $\bmod m$. Todėl gausime skirtingus $\bmod m$ skaičius. Iš viso jų yra m.

Sakykime, $a > 1$. Periodo ilgis lygus $m \Leftrightarrow$ kai kiekvienas skaičius x, $0 \le x < m$, sutinkamas periodo ilgio sekos dalyje lygiai vieną kartą. Vadinasi periodas lygus $m \Leftrightarrow$ kai sekos, su $X_0 = 0$, periodo ilgis lygus m. Paėmę $l = 0$, $k = n$, iš (5) lyginio turėsime

$$X_n \equiv \frac{a^n - 1}{a - 1}\, c \bmod m.$$

Kai $(c, m) \neq 1 \Rightarrow X_n \neq 1$ (skaičius 1 periodo ilgio sekos dalyje nebus sutinkamas). Dėl to sąlyga $(c, m) = 1$ būtina. Periodas lygus $m \Leftrightarrow$ kai mažiausias teigiamas skaičius n, kuriam $X_n = X_0 = 0$, yra lygus m. Dabar, kadangi $(c, m) = 1$, teoremos įrodymas išplaukia iš 5 lemos. ◀

1 pavyzdys *Pateiksime tiesinės kongruentinės sekos su maksimaliai galimu periodu, lygiu* 48, *pavyzdį.*

Kadangi $m = 48 = 2^4 \cdot 3$, tai iš 1 teoremos išplaukia, kad a ir c gali būti parinkti taip: $a = 13$, $c = 23$. Tegul $X_0 = 0$. Visa atsitiktinė seka atrodo taip:

$$\begin{array}{c} 0,\ 23,\ 34,\ 33,\ 20,\ 43,\ 6,\ 5,\ 40,\ 15,\ 26,\ 25,\ 12,\ 35,\ 46,\ 45, \\ 32,\ 7,\ 18,\ 17,\ 4,\ 27,\ 38,\ 37,\ 24,\ 47,\ 10,\ 9,\ 44,\ 19,\ 30,\ 29, \\ 16,\ 39,\ 2,\ 1,\ 36,\ 11,\ 22,\ 21,\ 8,\ 31,\ 42,\ 41,\ 28,\ 3,\ 14,\ 13. \end{array} \tag{10}$$

◀

Pateiktas pavyzdys nėra geras. Mažas atsitiktinumo laipsnis. Kai kurie sekos dėsningumai lengvai pastebimi. Autorius mėgino parinkti kitokius a ir c, bet dėsningumai vis tiek buvo akivaizdūs. Parinkti gerą atsitiktinę seką nėra paprasta. O ir parinkus reikia naudoti įvairius testus, ir įsitikinti, kad seka tikrai gera ir tinka modeliavimui.

Iš generuotos atsitiktinės natūraliųjų skaičių sekos lengvai galima sudaryti tolygiai intervale $[0, 1]$ pasiskirsčiusią pseudoatsitiktinę seką. Tai galima padaryti, pavyzdžiui, su formulės

$$U_n = \frac{X_n}{m} \tag{11}$$

pagalba. Tolygiai pasiskirsčiusios intervale $[0, 1]$ sekos naudojamos kitaip pasiskirsčiusioms sekoms sudaryti. Apie tai kalbėsime kituose skyriuose.

2 pavyzdys *Iš natūraliųjų atsitiktinių skaičių sekos* (10) *sukonstruokime atsitiktinių skaičių, tolygiai pasiskirsčiusių intervale* $[0,1]$, *seką.*

Naudodami (11) formulę (10) sekai, gausime tokią tolygiai intervale $[0,1]$ pasiskirsčiusią seką (apvaliname iki 4 vietų po kablelio tikslumu):

0	0,4792	0,7083	0,6877	0,4167	0,8958
0,125	0,1042	0,8333	0,3125	0,5417	0,5208
0,25	0,7292	0,9583	0,9375	0,6667	0,1458
0,375	0,3542	0,0833	0,5625	0,7917	0,7708
0,5	0,9792	0,2083	0,1875	0,9167	0,3958
0,625	0,6042	0,3333	0,8125	0,0417	0,0208
0,75	0,2292	0,4583	0,4375	0,1667	0,6458
0,875	0,8542	0,5833	0,0625	0,2917	0,2708

Šios sekos statistinės savybės kaip ir (10) natūraliųjų skaičių sekos nėra geros. Per mažas atsitiktinumo laipsnis. ◀

3.2.3 Multiplikatyvusis kongruentinis metodas

Tiesinė kongruentinė seka, kai $c = 0$, vadinama *multiplikatyviąja kongruentine seka.* Iš 1 teoremos išplaukia, kad šiuo atveju maksimalaus periodo ilgio negausime. Bet sekos generavimas yra greitesnis.

Pirma įrodysime pagalbinį rezultatą, o vėliau išsamiau panagrinėsime multiplikatyviąsias kongruentines sekas.

Tegul $\phi(n)$ – Eulerio funkcija, lygi skaičiui skaičių $0, 1, 2, \ldots, n-1$, kurie yra tarpusavyje pirminiai su n:

$$\phi(n) = \#\{i \,|\, (i,n) = 1,\ i \in \{0,1,\ldots,n-1\}\}.$$

6 lema (Eulerio teorema) *Jei* $(a,m) = 1$, *tai*

$$a^{\phi(m)} \bmod m = 1.$$

Įrodymas. Jeigu $(a_1, m) = 1$ ir $(a_2, m) = 1$, tai $(a_1 a_2 \bmod m, m) = 1$. Tegul

$0 \le x_1, \ldots, x_{\phi(m)} < m$ – skirtingi tarpusavy pirminiai su m skaičiai.

Tada

$ax_1 \bmod m, \ldots, ax_{\phi(m)} \bmod m$ – tie patys skaičiai, išdėstyti, gal būt, kita tvarka.

Taigi

$$ax_1 \bmod m \cdot \cdots \cdot ax_{\phi(m)} \bmod m = x_1 \cdot \ldots \cdot x_{\phi(m)},$$

arba

$$a^{\phi(m)} \bmod m = 1.$$

Lema įrodyta. ◀

Jei $d|m$ ir $d|X_n$, tai $d|X_{n+j}$, $j = 1, 2, \ldots$ Todėl, kai $c = 0$, būtų gerai, kad $(X_n, m) = 1$, $n = 1, 2, \ldots$ Tai, žinoma, taip pat riboja periodo ilgį.

Turėdami omenyje 2 lemą, nemažindami bendrumo galime nagrinėti atvejį

$$X_n \equiv a^n X_0 \bmod p^l.$$

Jeigu $p|a$, tai periodo ilgis nedidesnis už l. Todėl tegul $(a, p) = 1$. Periodas lygus mažiausiam sveikam λ, kuriam

$$X_0 \equiv a^\lambda X_0 \bmod p^l. \tag{12}$$

Sakykime,

$$(X_0, p^l) = p^s,$$

tuomet (12) ekvivalentu

$$a^\lambda \equiv 1 \bmod p^{l-s}.$$

Iš Eulerio teoremos (6 lema) turime, kad

$$a^{\phi(p^{l-s})} \equiv 1 \bmod p^{l-s}.$$

Taigi

$$\lambda | \phi(p^{l-s}) = p^{l-s} - p^{l-s-1} = p^{l-s-1}(p-1).$$

Dabar pateiksime keletą skaičių teorijos apibrėžimų. Tegul $(a, m) = 1$. Mažiausias natūralusis λ, kuriam

$$a^\lambda \equiv 1 \bmod m,$$

vadinamas *rodikliu* $\bmod m$. Skaičius a, kurį atitinka maksimaliai galimas rodiklis $\bmod m$, vadinamas *primityviuoju elementu*[6] $\bmod m$.

Tegul $\lambda(m)$ – primityvaus elemento rodiklis, t.y. maksimaliai galimas rodiklis $\bmod m$. Tuomet

$$\lambda(p^l) | p^{l-1}(p-1), \quad \text{kai } (X_0, p) = 1.$$

Galima tiksliai paskaičiuoti, kad

$$\begin{cases} \lambda(2) = 1, \\ \lambda(4) = 2, \\ \lambda(2^l) = 2^{l-2}, & \text{kai } l \geq 3, \\ \lambda(p^l) = p^{l-1}(p-1), & \text{kai } p > 2, \end{cases} \tag{13}$$

[6] Primityvaus elemento nereikia painioti su primityviąja šaknimi. Primityvios šaknys egzistuoja ne visiems m.

ir

$$\lambda(p_1^{l_1} \dots p_t^{l_t}) = \text{MBK}\Big(\lambda(p_1^{l_1}),\ \dots,\ \lambda(p_t^{l_t})\Big). \tag{14}$$

Dabar galime suformuluoti Karmaiklo[7] teoremą.

2 teorema (Karmaiklo teorema) *Multiplikatyviosios kongruenčios sekos maksimalus periodas yra lygus $\lambda(m)$, apibrėžtam (13) ir (14) formulėmis. Toks periodas gaunamas, kai*

- $(X_0, m) = 1$,
- a *– primityvusis elementas* $\bmod m$.

Pastebėkime, jei m – pirminis skaičius, tai galima gauti periodo ilgį lygų $m - 1$.

Kaip rasti primityviuosius elementus $\bmod m$? Teisinga tokia teorema.

3 teorema *Skaičius a yra primityvusis elementas*[8] $\bmod p^l \iff$

- $p^l = 2$, a *– nelyginis skaičius;*
- $p^l = 4$, $a \bmod 4 = 3$*;*
- $p^l = 8$, $a \bmod 8 = 3, 5, 7$*;*
- $p = 2$, $l \geq 4$, $a \bmod 8 = 3, 5$*;*
- $p > 2$, $l = 1$, $a \not\equiv 0 \bmod p$, $a^{(p-1)/q} \not\equiv 1 \bmod p \ \ \forall q \in \mathbb{P},\ q|(p-1)$*;*
- $p > 2$, $l > 1$, a *tenkina ankstesnę sąlygą ir* $a^{p-1} \not\equiv 1 \bmod p^2$.

Jeigu reikia rasti primityvųjį elementą $a \bmod m$, $m = p_1^{l_1} \dots p_t^{l_t}$, tai, pasirodo, egzistuoja vienintelis toks a, kad $a \equiv a_j \bmod p_j^{l_j}$, $j = 1, \dots, t$. Čia a_j – primityvusis elementas $\bmod p_j^{l_j}$.

Kai $m = 2^l$, $l \geq 4$, tai $a = 3, 5 \bmod 8$. Šiuo atveju ketvirtoji dalis visų galimų daugiklių duoda maksimalų periodą.

Antras svarbus atvejis, kai $m = 10^l$. Šiuo atveju teisinga tokia teorema.

4 teorema *Tegul $m = 10^l$, $l \geq 5$, $(X_0, 10) = 1$. Multiplikatyviosios kongruentinės sekos periodas lygus $5^{l-1}2^{l-2} = 5 \cdot 10^{l-2} \iff a \bmod 200$ yra lygus vienam iš 32 skaičių:*

$$3, 11, 13, 19, 21, 27, 29, 37, 53, 59, 61, 67, 69, 77, 83, 91, 109, 117,$$

$$123, 131, 133, 139, 141, 147, 163, 171, 173, 179, 181, 187, 189, 197.$$

[7] R.D.Carmichael, *Bull. Amer. Math. Soc.*, **16**, 1910, 232-238.

[8] Jeigu modulis yra 2, arba 2^2, arba p^l $(p > 2)$, tai primityvieji elementai bus ir primityviosiomis šaknimis.

3.2.4 Tiesinės kongruentinės sekos galingumas

Jau žinome, kad maksimalus tiesinės kongruentinės sekos periodas gaunamas, kai $b = a - 1$ yra visų pirminių m daliklių kartotinis ir 4 kartotinis, jei m dalijasi iš 4 (žr. 1 teoremą). Tegul $m = z^l$. Tada daugiklis a tenkina minėtus reikalavimus, pavyzdžiui, kai

$$a = z^k + 1, \quad 2 \leq k < l. \tag{15}$$

Iš 1 teoremos išplaukia, kad galime paimti $c = 1$. Tuomet

$$X_{n+1} \equiv (z^k + 1)X_n + 1 \bmod z^l.$$

Ši formulė patogi skaičiavimuose, nes galima išvengti daugybos (kai žodžio ilgis lygus z^l), pakeičiant ją postūmiu ir sudėtimi.

Ir vis tik (15) tipo daugiklių reikia vengti. Gautos sekos nėra labai "atsitiktinės". Kodėl?

Tiesinės kongruentinės sekos su maksimaliu periodu *galingumu* vadinsime mažiausią natūralųjį skaičių s, kuriam

$$b^s \equiv 0 \bmod m.$$

Toks s visuomet egzistuoja, kai a tenkina 1 teoremos reikalavimus.

Nemažindami bendrumo galime paimti $X_0 = 0$. Tuomet iš (5) turėsime

$$X_n \equiv \frac{(a^n - 1)c}{b} \bmod m.$$

Išskleidę $a^n - 1 = (1+b)^n - 1$ pagal Niutono binomo formulę, (kai n pakankamai dideli, $n > s$) gausime

$$X_n \equiv c\Big(n + C_n^2 b + \cdots + C_n^s b^{s-1}\Big) \bmod m.$$

Narius su b^s, b^{s+1} ir t.t. praleidžiame, nes jie yra m kartotiniai.

Jeigu $a = 1$, galingumas $s = 1$, $X_n \equiv cn \bmod m$. Seka, aišku, neatsitiktinė. Tegul $s = 2$. Tuomet

$$X_n \equiv cn + cbC_n^2 \bmod m.$$

Ir šiuo atveju seka mažai atsitiktinė:

$$X_{n+1} - X_n \equiv c + cbn \bmod m.$$

Jeigu $s = 3$, seka, atrodo, labiau atsitiktinė, bet X_n, X_{n+1}, X_{n+2} dar vis stipriai susiję. Priimtini rezultatai gaunami, kai $s = 4$, bet dar ginčytini. Reikia siekti, kad $s \geq 5$.

Pateikti samprotavimai ir paaiškina sekos galingumo prasmę. Galingumas yra tik vienas iš kriterijų parenkant daugiklį. Pabaigai pateiksime keletą pavyzdžių.

Paimkime $m = 2^{35}$ ir $a = 2^k + 1 \Rightarrow b = 2^k$. Kai $k \geq 18$, $b^2 = 2^{2k}$ yra m kartotinis $\Rightarrow s = 2$. Kai $k = 17, \ldots, 12 \Rightarrow s = 3$. Kai $k = 11, 10, 9 \Rightarrow s = 4$. Taigi reikėtų imti $k \leq 8$. Tada $a \leq 257$. Bet šiuo atveju daugiklis a nedidelis. Vėl gaunamos sekos, kurių reikia vengti (įrodyta, kad sekos labiau "atsitiktinės", kai a dideli).

Kai $m = w \pm 1$ (w – žodžio ilgis), tai, bendrai imant, m neišsiskaido į aukšto laipsnio pirminius daugiklius ir s nedidelis. Todėl šiais atvejais nereikėtų naudotis maksimalaus periodo metodu, o imti $c = 0$.

Vėliau nagrinėsime spektrinį testą. Su jo pagalba bus galima įsitikinti, kad daugiklis $2^{23} + 2^{14} + 2^2 + 1$, kai $m = 2^{35}$, gana geras. Narys 2^{23} padaro daugiklį gana didelį, 2^2 užtikrina didelį galingumą, 2^{14} naudojamas, kad daugiklis nebūtų labai jau paprastas, o seka būtų pakankamai "atsitiktinė".

3.3 Tiesinis rekurentinis metodas

Tiesine rekurentine seka vadinama dvejetainė seka X_n, gaunama algoritmo

$$X_n \equiv c_1 X_{n-1} + c_2 X_{n-2} + \cdots + c_p X_{n-p} \bmod 2 \tag{16}$$

pagalba. Kad algoritmas galėtų veikti, turi būti parinkti pirmieji sekos nariai: $X_1, X_2, \ldots, X_p \in \{0, 1\}$ ir konstantos $c_1, c_2, \ldots, c_p \in \{0, 1\}$. Kai konstantų seka $c_1, \ldots, c_p$ yra fiksuota, X_n priklauso tik nuo p paskutiniųjų sekos narių $X_{n-1}, \ldots, X_{n-p}$. Todėl tiesinės rekurentinės sekos periodo ilgis yra mažiausias skaičius (sekos nario indeksas), kai rinkinys $X_{n-1}, \ldots, X_{n-p}$ pasikartoja. Jei rinkinį iš eilės einančių p narių sudaro nuliukai, tai visi sekantys nariai irgi bus nuliai. Todėl sekos periodas negali būti didesnis už $2^p - 1$.

5 teorema *Tiesinės rekurentinės sekos* (16) *periodas yra maksimalus* (*lygus* $2^p - 1$) $\Longleftrightarrow$ *daugianaris*

$$f(x) = 1 + c_1 x + c_2 x^2 + \cdots + c_p x^p$$

yra primityvusis polinomas virš baigtinio polinomų, su koeficientais 0 *arba* 1, *kūno* (*virš Galua kūno* $GF(2)$).

3 pavyzdys *Imdami primityvųjį polinomą* $f(x) = 1 + x^3 + x^5$, *sukonstruokime tiesinę rekurentinę seką.*

Atitinkama tiesinė rekurentinė seka yra

$$X_n = X_{n-3} + X_{n-5} \bmod 2.$$

Paimkime pradinę seką: $X_1 = 1,\ X_2 = 1,\ X_3 = 0,\ X_4 = 1,\ X_5 = 0$. Tuomet generuota seka atrodys taip:

(17) 11010 10000 10010 11001 11110 00110 1.

Toliau seka vėl kartosis. Sekos periodas yra lygus $2^5 - 1 = 31$. ◀

Yra daug būdų iš dvejetainės sekos gauti natūraliųjų skaičių sekas. Tarkime, b_i, $i = 1, 2, \ldots$, yra dvejetainė seka. Sekos

$$X_i = (b_{ik} b_{ik+1} \ldots b_{ik+L})_2, \tag{18}$$

$$Y_i = (b_i b_{i-l} \ldots b_{i-lt})_2, \tag{19}$$

su pasirinktais $k,\ L,\ l,\ t$, yra natūraliųjų skaičių sekos. Seka Y_i gauta naudojant metodą su vėlavimais.

4 pavyzdys *Dvejetainę seką* (17) *paverskime natūraliųjų skaičių sekomis pagal* (18) *ir* (19) *formules.*

Pasirinkime $k = 5, L = 4$. Iš (18) formulės turėsime

$$X_1 = (b_5 b_6 b_7 b_8 b_9)_2 = 01000_2 = 8,$$

$$X_2 = (b_{10} b_{11} b_{12} b_{13} b_{14})_2 = 01001_2 = 9,$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$X_7 = (b_{35} b_{36} b_{37} b_{38} b_{39})_2 = (b_4 b_5 b_6 b_7 b_8)_2 = 10100_2 = 20,$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$X_{31} = (b_{31\cdot 5} b_{31\cdot 5+1} b_{31\cdot 5+2} b_{31\cdot 5+3} b_{31\cdot 5+4})_2 = (b_{31} b_1 b_2 b_3 b_4)_2 = 11101_2 = 29.$$

Visa seka atrodo taip:

$$8, 9, 12, 31, 3, 28, 20, 8, 22, 15, 17, 23, 10, 2, 11, 7,$$

$$24, 27, 21, 1, 5, 19, 28, 13, 26, 16, 18, 25, 30, 6, 29.$$

Parinkę $l = 6,\ t = 4$, iš (19) formulės gausime

$$Y_i = (b_i b_{i-6} b_{i-12} b_{i-18} b_{i-24})_2.$$

Paskaičiavę, turėsime

$$Y_1 = (b_1 b_{-5} b_{-11} b_{-17} b_{-23})_2 = (b_1 b_{26} b_{20} b_{14} b_8)_2 = 10110_2 = 22,$$

$$Y_2 = (b_2 b_{27} b_{21} b_{15} b_9)_2 = 10100_2 = 20$$

ir t.t. Visa seka atrodo taip:

$$22, 20, 14, 31, 8, 24, 11, 10, 7, 15, 18, 12, 5, 21, 3,$$
$$23, 25, 6, 2, 26, 17, 27, 28, 19, 1, 13, 8, 29, 30, 9, 16.$$

◀

Iš atsitiktinės binarinės sekos lengva padaryti atsitiktinę seką iš intervalo $[0, 1]$. Jei binarinė seka atsitiktinė, tai gauta seka intervale $[0, 1]$ turėtų būti jame tolygiai pasiskirsčiusi. Yra daug būdų. Pavyzdžiui, tolygiai pasiskirsčiusi seka galėtų būti gauta naudojant formulę:

$$U_i = (0, b_{ik} b_{ik+1} \dots b_{ik+L})_2, \tag{20}$$

su pasirinktais k ir L.

5 pavyzdys *Iš dvejetainės sekos* (17) *gaukime tolygiai intervale* $[0, 1]$ *pasiskirsčiusią seką.*

Naudodami (17) seką, (20) formulėje parinkę $k = 5,\ L = 4$, gautume

$$U_1 = (0, b_5 b_6 b_7 b_8 b_9)_2 = 0,01000_2 = 0,25,$$
$$U_2 = (0, b_{10} b_{11} b_{12} b_{13} b_{14})_2 = 0,01001_2 = 0,28125,$$
$$\dots\dots\dots\dots\dots\dots\dots\dots$$
$$U_{31} = (0, b_{31} b_1 b_2 b_3 b_4)_2 = 0,11101_2 = 0.90625.$$

Visa atsitiktinė seka iš intervalo $[0, 1]$ atrodytų taip:

0,25	0,28125	0,375	0,96875	0,09375	0,875
0,625	0,25	0,6875	0,46875	0,53125	0,71875
0,3125	0,0625	0,34375	0,21875	0,75	0,84375
0,65625	0,03125	0,15625	0,59375	0,875	0,40625
0,8125	0,5	0,5625	0,78125	0,9375	0,1875
0,90625					

◀

Tiesinio rekurentinio metodo, lyginant jį su tiesiniu kongruentiniu, privalumai yra tokie. Pirma, tiesinė rekurentinė seka yra labai greitai generuojama. Čia nereikia atlikti jokios daugybos tik sudėtį $\bmod 2$ ir postūmį, rekurentiškai skaičiuojant. Atminties irgi nereikia daug. Antra, periodo ilgis nėra susijęs su kompiuterio žodžio ilgiu. Galima generuoti labai ilgo periodo sekas (pavyzdžiui, tokio ilgio $2^{251} - 1 > 10^{156}$ ar dar ilgesnes). Šio metodo trūkumas, lyginant jį su tiesiniu kongruentiniu metodu, yra tai, kad jis gerokai mažiau ištirtas. Dėl šios priežasties kongruentinis metodas patikimesnis ir plačiau taikomas.

3.4 Kiti metodai

Yra daug p.a.s. generavimo metodų. Bet tiesinis kongruentinis metodas geras tuo, kad žinoma teorija ir galima garantuoti atsitiktinumą. Pavyzdžiui, tegul

$$X_{n+1} \equiv (aX_n \bmod (m+1) + c) \bmod m.$$

Atrodo, kad seka labiau sujaukta negu tiesinio kongruentinio metodo. Atsakymas toks: mažiau atsitiktinė. Apskritai, jeigu generuojame $X_{n+1} = f(X_n)$, tai funkcija f turi būti tiksliai apibrėžta ir ne ypač sudėtinga, kad galėtume kurti teoriją. Yra daug tiesinio kongruentinio metodo apibendrinimų.

A)

$$X_{n+1} \equiv dX_n^2 + aX_n + c \bmod m$$

– metodas geras, yra ir 1 teoremos apibendrinimas, nusakantis kada gauname maksimalų periodą lygų m; sunkiau realizuojamas kompiuteryje; reikia daugiau laiko generavimui.

B)

$$X_{n+1} \equiv X_n(X_n + 1) \bmod 2^l, \quad X_0 \equiv 2 \bmod 4,$$

– geras. Pasinaudojus šia seka galima gauti seką, kurios gavimas sutampa, jei ją skaičiuotume kvadrato vidurio metodu.

Imant $X_{n+1} = f(X_n, X_{n-1}) \bmod m$, periodo ilgis gali būti $\leq m^2$.

C)

$$X_{n+1} \equiv X_n + X_{n-1} \bmod m$$

– Fibonači seka; nepakankamai atsitiktinė; periodas šiek tiek $> m$.

D)

$$X_{n+1} \equiv X_n + X_{n-k} \bmod m$$

– kai $k \leq 15$, seka nėra pakankamai atsitiktinė; kai $k = 16$, atrodo, viskas gerai. Periodas gali būti $> m$. Kitas privalumas – greitis; nėra daugybos, tik sudėtis.

E)

$$X_n \equiv a_1 X_{n-1} + \cdots + a_k X_{n-k} \bmod p$$

– tai ir tiesinio kongruentinio ir tiesinio rekurentinio metodų apibendrinimas. Didžiausias periodo ilgis $p^k - 1$. Kai $k = 1$, gauname jau nagrinėtą multiplikatyviąją kongruentinę seką. Atveju $p = 2$, gauname jau nagrinėtą tiesinę rekurentinę seką. Konstantų $a_1, \ldots, a_k$ parinkimas tik tuomet duoda laukiamą rezultatą, kai polinomas $f(x) = x^k - a_1 x^{k-1} - \cdots - a_k$ yra primityvusis daugianaris $\bmod p$. Yra sukurta teorija kaip konstantas $a_1, \ldots, a_k$ atskirais atvejais parinkti, bet tai gana sudėtingas reikalas.

F)

$$Z_n \equiv X_n + Y_n \bmod m$$

– gaunami geri razultatai, tik reikia, kad sekų X_n ir Y_n periodai būtų tarpusavyje pirminiai. Galimi ir kitokie dviejų tiesinių kongruentinių sekų X_n ir Y_n sumaišymo variantai.

4 STATISTINIAI TESTAI

Labai svarbu, kad kokiu nors būdu gauta p.a.s. seka būtų panaši į atsitiktinę. Jau žinome kaip gauti didelio periodo sekas, kad praktiniuose uždaviniuose nebūtų sekos pasikartojimo. Bet ilgas periodas dar nereiškia, kad seka pasižymi atsitiktinės sekos savybėmis. Ryšį su sekos atsitiktinumu turi anksčiau aptartas sekos galingumas. Bet pats galingumas apibrėžiamas ne statistiniais terminais. Galingumas apsprendžia gretimų tiesinės kongruentinės sekos narių susijimo laipsnį. Didelis galingumas gerai, bet vis tiek seka gali nebūti panaši į atsitiktinę. Tai kaipgi nuspręsti, ar seka pakankamai atsitiktinė?

Iš dalies atsakyti į iškeltą klausimą leidžia statistinė teorija. Tam yra daugybė statistinių testų. Šiame skyriuje susipažinsime su testų teorijos pagrindais. Nemažai konkrečių testų panagrinėsime. Skaitytojas, kuriam prireiks kitokių testų, susipažinęs su testų pagrindais, jau galės pats taikyti tam sukurtus testus. Daug aprašytų statistinių testų galima rasti [7, 9, 14].

Jeigu p.a.s. seka sėkmingai praėjo testų $T_1, T_2, \ldots, T_n$ patikrinimą, tai dar negalima daryti išvados, kad ji bus gerai įvertinta ir testo T_{n+1}, ir juo labiau, kad ji pasižymi visomis atsitiktinės sekos savybėmis. Tačiau kiekvienas testas mus vis labiau įtikina, kad seka atsitiktinė. Paprastai p.a.s. sekos tikrinamos daugybe testų (priklausomai nuo sprendžiamų uždavinių svarbumo, nuo sekos taikymo intensyvumo ir t.t.). Jeigu testų rezultatai teigiami, laikome seką atsitiktine (atsitiktine ji laikoma tol, kol su kokio nors testo pagalba įrodomas jos "kaltumas").

Yra kelios testų rūšys.

Empiriniai testai – testai, kai kompiuteris, manipuliuodamas sekos skaičių grupėmis ir naudodamas konkrečius statistinius kriterijus, įvertina seką.

Teoriniai testai – testai, kai, naudojant skaičių teorijos metodus ir rekurentinį pačių p.a.s. generavimo sąryšį, nustatomos kai kurios sekos charakteristikos.

4.1 Universalieji testai

Panagrinėsime tris dažniausiai naudojamus statistinius testus.

4.1.1 χ^2 kriterijus

χ^2 statistinis kriterijus pasiūlytas 1900 m. Pirsono (C.Pearson). Jis yra gana paprastas ir turbūt plačiausiai taikomas.

Teorinis χ^2 testo pagrindimas yra 6 teorema.

Tarkime, yra duotas atsitiktinis dydis ξ su reikšmėmis aibėje Ξ. Aibė Ξ yra išskaidyta į tarpusavyje nesikertančius poaibius $\Xi_1, \Xi_2, \ldots, \Xi_r$, $\Xi_i \bigcap \Xi_j = \emptyset$, jei $i \neq j$. Toliau, tegul tikimybė, kad atsitiktinis dydis ξ įgis reikšmę iš poaibio Ξ_i yra lygi p_i:

$$P\{\xi \in \Xi_i\} =: p_i, \quad i = 1, 2, \ldots, r, \quad p_1 + p_2 + \cdots + p_r = 1.$$

Nagrinėkime n nepriklausomų atsitiktinio dydžio ξ realizacijų $\xi_1, \xi_2, \ldots, \xi_n$. Raide ν_i pažymėkime skaičių tų realizacijų ξ_i, kurios pakliūva į poaibį Ξ_i. Aišku, kad ν_i yra atsitiktinis dydis, o jo vidurkis $E\nu_i = np_i$. Kai turime konkrečias atsitiktinio dydžio ξ realizacijas $\xi_1, \xi_2, \ldots, \xi_n$, tai santykiniai tankiai ν_i/n skirsis nuo teorinių tikimybių p_i. Išmatuoti šiam skirtumui galima naudoti dydžius $(\nu_i - np_i)^2$, arba tiksliau jų sumą su svoriais. Teisinga tokia teorema.

6 teorema (Pirsono) *Atsitiktinis dydis*

$$\chi_n^2 := \sum_{i=1}^{r} \frac{(\nu_i - np_i)^2}{np_i}\,, \tag{21}$$

kai $n \to \infty$, konverguoja į χ^2 skirstinį su $r-1$ laisvės laipsniu:

$$\lim_{n\to\infty} P(\chi_n^2 < u) = \int_0^u g_{r-1}(x)\,dx\,.$$

Čia

$$g_{r-1}(x) = \left(x^{\frac{r-3}{2}} e^{-\frac{x}{2}}\right) \Bigg/ \left(2^{\frac{r-1}{2}} \Gamma\left(\frac{r-1}{2}\right)\right),$$

o Γ žymi Oilerio gama funkciją.

6 teoremos įrodymą galima rasti [9, ?].

Naudodami lygybes

$$(\nu_i - np_i)^2 = \nu_i^2 - 2np_i\nu_i + n^2p_i^2\,,$$

$$\nu_1 + \nu_2 + \cdots + \nu_r = n\,,$$

$$p_1 + p_2 + \cdots + p_r = 1\,,$$

(21) formulę galime užrašyti tokiu būdu:

$$\chi_n^2 = \frac{1}{n}\sum_{i=1}^{r} \frac{\nu_i^2}{p_i} - n\,. \tag{22}$$

Daugeliu atvejų (22) formulė palengvina skaičiavimus.

Kaip praktiškai taikyti Pirsono teoremą? Tarkime, kokio nors generatoriaus pagalba gauti p.a.s. $x_1, x_2, \ldots, x_n$, priklausantys intervalui $[a, b]$. Mums reikia patikrinti savo spėjimus, kad jie yra atsitiktiniai, vienodai pasiskirstę, o jų tankio funkcija yra $f(x)$. Išskaidome intervalą $[a, b]$ į r dalių

$$[a, c_1], [c_1, c_2], \ldots, [c_{r-1}, b]\,.$$

Tuomet ν_i yra generuotų p.a.s. $x_1, x_2, \ldots, x_n$ skaičius intervale $[c_{i-1}, c_i]$, o $p_i = \int_{c_{i-1}}^{c_i} f(x)\,dx$. Paskaičiuokime statistiką

$$k_n^2 = \sum_{i=1}^{r} \frac{(\nu_i - np_i)^2}{np_i} = \frac{1}{n}\sum_{i=1}^{r} \frac{\nu_i^2}{p_i} - n\,.$$

Jeigu p.a.s. $x_1, x_2, \ldots, x_n$ pasiskirstę pagal dėsnį su tankio funkcija $f(x)$, tai iš 6 teoremos išplaukia, kad statistika k_n^2 asimptotiškai pasiskirsčiusi pagal χ^2 dėsnį su $r-1$ laisvės laipsniu.

Pasirinkime *patikimumo lygmenį* β (paprastai $\beta = 0.95$). Dydis $\alpha = 1-\beta$ vadinamas kriterijaus *reikšmingumo lygmeniu* arba I *rūšies klaidos tikimybe.* Jeigu mūsų spėjimai teisingi, tai dydis

$$\int_{k_n^2}^{\infty} g_{r-1}(x)\,dx$$

neturėtų būti labai mažas. Turėdami reikšmingumo lygmenį α, apibrėžkime dydį $\chi^2(r-1, \alpha)$, vadinamą *kritine reikšme*, pagal formulę:

$$P(\chi^2 \geq \chi^2(r-1, \alpha)) = \int_{\chi^2(r-1,\alpha)}^{\infty} g_{r-1}(x)\,dx = \alpha\,.$$

Geometrinė dydžių α ir $\chi^2(r-1, \alpha)$ interpretacija pateikta *8 pav.*

8 pav. Skirstinio χ^2_{r-1} tankio funkcija

Jei teisinga nelygybė

$$k_n^2 < \chi^2(r-1,\alpha)\,,$$

tai galime tvirtinti, kad eksperimento duomenys neprieštarauja pradinei hipotezei, t.y. kad atsitiktiniai dydžiai $x_1, x_2, \ldots, x_n$ pasiskirstę pagal dėsnį su tankio funkcija $f(x)$.

Antraip, jei

$$k_n^2 \geq \chi^2(r-1,\alpha)\,,$$

tai hipotezė (su patikimumo lygmeniu β) atmetama.

Lentelėse paprastai būna pateikta reikšmingumo lygmenys α ir juos atitinkančios kritinės reikšmės $\chi^2(r-1,\alpha)$. χ^2 kriterijaus lenteles galima rasti [**?**]. Šias lenteles pateikia ir programų paketas STATISTIKA.

6 pavyzdys *Sekos iš intervalo* $[0,1]$ *tolygumo testas.*

Reikia patikrinti hipotezę, kad generuota p.a.s. seka $x_1, x_2, \ldots, x_n$, priklausanti intervalui $[0,1]$, yra pasiskirsčiusi tolygiai.

Padaliname intervalą, pavyzdžiui, į 10 vienodų intervaliukų ($r = 10$):

$$[0, 0.1],\ [0.1, 0.2],\ \ldots,\ [0.9, 1]\,. \tag{23}$$

Tarkime, kad mūsų seką sudaro 1000 narių ($n = 1000$). Tegul sekos $x_1, x_2, \ldots, x_{1000}$ narių, pakliūvančių į intervalus (23), skaičiai eilės tvarka yra tokie:

$$83,\ 98,\ 104,\ 120,\ 117,\ 78,\ 95,\ 106,\ 89,\ 110.$$

Kadangi tikriname hipotezę, kad seka pasiskirsčiusi tolygiai, tai, aišku, tikimybės $p_i = 0.1$, $i = 1, 2, \ldots, 10$. Paskaičiuojame statistiką

$$k^2_{1000} = \sum_{i=1}^{10} \frac{(x_i - 1000 \cdot 0.1)^2}{1000 \cdot 0.1}$$

$$= \frac{289 + 4 + 16 + 400 + 289 + 484 + 25 + 36 + 121 + 100}{100} = 17.64\,.$$

Pasirinkime reikšmingumo lygmenį $\alpha = 0.05$. Tuomet klaidos tikimybė bus lygi 0.05.

χ^2 kriterijaus su 9 laisvės laipsniais duomenų eilutė, paimta iš lentelės, atrodo taip:

	$\alpha = 0.99$	$\alpha = 0.95$	$\alpha = 0.75$	$\alpha = 0.50$	$\alpha = 0.25$	$\alpha = 0.05$	$\alpha = 0.01$
$r = 9$	2.088	3.325	5.899	8.343	11.39	16.92	21.67

Iš lentelės matome, kad kritinė reikšmė $\chi^2(9, 0.05) = 16.92$. Kadangi $17.64 > 16.92$, tai iškeltą hipotezę, kad seka $x_1, x_2, \ldots, x_{1000}$ yra tolygiai pasiskirsčiusi intervale $[0, 1]$ (su patikimumo lygmeniu 0.95) atmetame. ◀

Jeigu χ^2 kriterijus naudojamas patikrinti kokiam nors

generatoriui daug kartų ir generuotos sekos skirtingos, tai normalu, kad kartais dydžiai k^2_n būna dideli. Hipotezė gali būti priimta, jei maždaug devyniolikai atvejų iš dvidešimties $k^2_n < \chi^2(r - 1, \alpha)$ (su patikimumo lygmeniu $\beta = 0.95$).

Pabaigai keletas pastabų.

χ^2 kriterijus duoda geriausius rezultatus. Pirma, kai patekimo į poaibius tikimybės p_i yra lygios arba beveik lygios. Antra, kai realizacijų skaičius pakankamai didelis: $np_i > 20$.

Tuo atveju, kai χ^2 kriterijus duoda labai gerus rezultatus (t.y. kai $\alpha = 0.99$ ar panašiai), hipotezę irgi reikėtų atmesti. Atmetimo argumentai tokie. Maždaug 99% atvejų dydžio, kuris turi χ^2 skirstinį, reikšmė yra didesnė už kritinę reikšmę $\chi^2(r - 1, 0.99)$. Pavyzdžiui, jeigu gauname seką, pasiskirsčiusią labai tolygiai, tai irgi nenormalu (tokių realizacijų tikimybė nedidelė, vargu ar tokia seka atsitiktinė). Geriausia, kai reikšmingumo lygmuo α, atitinkantis gautą k^2_n reikšmę, pakliūva į intervalą $[0.05, 0.95]$. Jei eksperimentas kartojamas daug kartų, tai atsilenkimai nuo normos turėtų būti. Tų statistikų k^2_n, kurioms α nepakliūva į intervalą $[0.05, 0.95]$, skaičius turėtų apytikriai sudaryti vieną dešimtąją dalį visų statistikų.

4.1.2 Kolmogorovo-Smirnovo kriterijus

χ^2 kriterijus taikomas tais atvejais, kai eksperimento rezultatai yra suskirstyti į baigtinį skaičių r grupių. Jeigu taip nėra, tai reikia dirbtinai skirstyti

eksperimento rezultatus į grupes. (6) pavyzdyje taip ir darėme. Dirbtinai skirstydami į grupes netenkame dalies informacijos. Kai atsitiktiniai dydžiai gali įgauti be galo daug reikšmių, dažnai hipotezių tikrinimui naudojami kiti kriterijai, tarp jų ir Kolmogorovo-Smirnovo (KS) kriterijus.

KS kriterijus *palygina empirinę ir tikrąją pasiskirstymo funkcijas.* Tegul nepriklausomų bandymų metu gautos atsitiktinio dydžio ξ realizacijos ξ_1, $\xi_2, \ldots, \xi_n$. Tuomet funkcija

$$F_n(x) := \frac{1}{n} \#\{i \mid \xi_i \le x,\ i = 1, 2, \ldots, n\}$$

vadinama *empirine atsitiktinio dydžio ξ pasiskirstymo funkcija.* Jeigu atsitiktinis dydis ξ iš tikrųjų pasiskirstęs pagal dėsnį $F_\xi(x)$, tai dideliems n skirtumas tarp tikrosios pasiskirstymo funkcijos $F_\xi(x)$ ir jos statistinio įverčio $F_n(x)$ neturėtų būti didelis, t.y. dydis $|F_n(x) - F_\xi(x)|$ neturėtų būti didelis.

Tegul

$$D_n := \sup_x |F_n(x) - F_\xi(x)|. \tag{24}$$

Teisinga tokia teorema.

7 teorema (Kolmogorovo-Smirnovo) *Kiekvienai tolydžiajai pasiskirstymo funkcijai $F_\xi(x)$*

$$\lim_{n\to\infty} P(\sqrt{n}\, D_n \le x) = H(x).$$

Čia

$$H(x) = \begin{cases} 0, & \text{jei } x \le 0, \\ 1 - 2\sum\limits_{i=1}^{\infty} (-1)^{i-1} \exp(-2ix^2), & \text{jei } x > 0. \end{cases}$$

Pasiskirstymo funkcijos $H(x)$ reikšmės yra tabuliuotos ir praktinėms reikmėms paprastai pakanka tikslumo, kai $n > 35$. KS kriterijų, skirtingai nuo χ^2 kriterijaus, galima taikyti visiems n, pradedant $n = 1$. Pats KS kriterijus pritaikomas pagal tą pačią schemą kaip ir χ^2 kriterijus. Daroma prielaida, kad atsitiktinis dydis X pasiskirstęs pagal tolydųjį dėsnį $F_X(x)$ ((24) formulėje sutampa su $F_\xi(x)$). Pasirenkamas reikšmingumo lygmuo α (paprastai $\alpha = 0.05$). Iš lentelių surandamas kritinis taškas x_α. Naudojant (24) formulę paskaičiuojama statistika D_n. Jei

$$\sqrt{n}\, D_n > x_\alpha,$$

tai hipotezę, kad atsitiktinis dydis X pasiskirstęs pagal dėsnį $F_X(x)$, (su patikimumo lygmeniu $\beta = 1 - \alpha$) atmetame. Jei

$$\sqrt{n}\, D_n \le x_\alpha,$$

tai galima tvirtinti (su patikimumo lygmeniu β), kad stebėjimo duomenys neprieštarauja hipotezei.

Remdamiesi šiais rezultatais, galime sudaryti KS kriterijų tikrinti hipotezei, kad atsitiktinių dydžių seka $X_1, X_2, \ldots, X_n$ yra pasiskirsčiusi pagal dėsnį $F(x)$. Tarkime, kad $X_1^*, X_2^*, \ldots, X_n^*$ yra seka $X_1, X_2, \ldots, X_n$ išdėstyta didėjimo tvarka. (24) D_n statistika nėra labai patogi skaičiavimuose. Bet, remdamiesi tuo, kad $F(x)$ yra nemažėjanti funkcija, o $F_n(x)$ turi baigtinį skaičių šuoliukų, (24) statistiką galime paskaičiuoti tokiu būdu. Skaičiuojame statistikas

$$D_n^+ = \max_{1\le i\le n}\left(\frac{i}{n} - F(X_i^*)\right), \tag{25}$$

$$D_n^- = \max_{1\le i\le n}\left(F(X_i^*) - \frac{i-1}{n}\right). \tag{26}$$

Tada

$$D_n = \max(D_n^+, D_n^-).$$

7 pavyzdys *Patikrinkime hipotezę, kad* 50 *narių seka* $X_1, X_2, \ldots, X_{50}$ *yra tolygiai pasiskirsčiusi intervale* $[0, 1]$. *Tegul ši seka, užrašyta didėjimo tvarka, yra tokia*

0.01	0.02	0.08	0.08	0.12	0.15	0.18	0.18	0.19	0.25
0.25	0.27	0.27	0.29	0.32	0.35	0.36	0.37	0.39	0.39
0.40	0.41	0.41	0.49	0.49	0.49	0.51	0.54	0.56	0.58
0.60	0.63	0.66	0.68	0.69	0.71	0.74	0.75	0.76	0.77
0.81	0.82	0.82	0.88	0.89	0.90	0.91	0.91	0.98	0.99

Šiuo atveju pasiskirstymo funkcija $F(x)$ (25) ir (26) formulėse yra lygi x. Paskaičiavę gauname, kad

$$D_{50}^+ = \max\left(\frac{1}{50} - 0.01,\ \frac{2}{50} - 0.02,\ \ldots,\ \frac{50}{50} - 0.99\right) = 0.05,$$

$$D_{50}^- = \max\left(0.01 - \frac{0}{50},\ 0.02 - \frac{1}{50},\ \ldots,\ 0.99 - \frac{49}{50}\right) = 0.07.$$

Taigi $D_{50} = 0.07$, o $\sqrt{50}\,D_{50} \approx 0.5$. Pasirinkę reikšmingumo lygmenį $\alpha = 0.05$, gausime, kad kritinė reikšmė $x_\alpha \approx 13.3$. Kadangi statistika

$$\sqrt{50}\,D_{50} \approx 0.5 < 13.3,$$

tai darome išvadą (su patikimumo lygmeniu $\beta = 0.95$), kad KS kriterijaus rezultatai neprieštarauja iškeltai hipotezei, jog atsitiktinė seka $X_1, X_2, \ldots, X_{50}$ pasiskirsčiusi tolygiai intervale $[0, 1]$. ◀

Taigi χ^2 kriterijus taikomas tada, kai atsitiktiniai dydžiai yra suskirstyti į grupes. Paprastai tolydiesiems dydžiams χ^2 kriterijus netaikomas, bet, dirbtinai suskaidžius atsitiktinius dydžius į grupes, jį galima taikyti ir tolydiesiems dydžiams. Norint gauti gerus rezultatus su χ^2 kriterijumi, imties elementų skaičius n turi būti didelis.

KS kriterijus taikomas bet kokiems n. Žinoma, kai n didesnis – geriau. Skirtingai nuo χ^2 kriterijaus, KS kriterijus taikomas tik tolydiesiems atsitiktiniams dydžiams. Juo tikrinama hipotezė, ar atsitiktiniai dydžiai pasiskirstę pagal dėsnį su spėjama tolydžiąja pasiskirstymo funkcija.

4.1.3 ω^2 kriterijus

ω^2 arba Kramero-fon Mizeso (C.H.Cramér ir R.E.von Mises) testas, panašiai kaip ir KS testas, pagrįstas teorinės ir empirinės pasiskirstymo funkcijų artumu. Tegul atsitiktinis dydis ξ yra pasiskirstęs pagal tolyduųjį dėsnį $F_\xi(x)$, o $F_n(x)$ yra jo n nepriklausomų realizacijų pasiskirstymo funkcija. ω^2 teste naudojama statistika

$$\omega_n^2 := n \int_{-\infty}^{+\infty} \Big(F_\xi(x) - F_n(x)\Big)^2 \, dF_\xi(x). \tag{27}$$

Teisinga tokia teorema.

8 teorema *Jei atsitiktinis dydis ξ yra pasiskirstęs pagal tolyduųjį dėsnį $F_\xi(x)$, tai*

$$\lim_{n\to\infty} P(\omega_n^2 < x) = a_1(x).$$

Funkcija $a_1(x)$ yra tabuliuota. Jeigu generuotų atsitiktinių dydžių $X_1, X_2, \ldots, X_n$ pasiskirstymo funkcija yra $F(x)$, tai jų empirinė pasiskirstymo funkcija $F_n(x)$ turėtų nedaug skirtis nuo $F(x)$, kai n didelis. Analogiškai kaip ir taikant KS kriterijų, reikia pasirinkti reikšmingumo lygmenį α, iš lentelių surasti kritinę reikšmę x_α, suskaičiuoti statistiką ω_n^2. Jei $\omega_n^2 > x_\alpha$, tai hipotezė, kad seka $X_1, X_2, \ldots, X_n$ pasiskirsčiusi pagal dėsnį $F(x)$ (su patikimumo lygmeniu $\beta = 1 - \alpha$), atmetama. Jei $\omega_n^2 \le x_\alpha$, tai daroma išvada (su patikimumo lygmeniu β), kad eksperimento duomenys neprieštarauja hipotezei, jog atsitiktinė seka $X_1, X_2, \ldots, X_n$ pasiskirsčiusi pagal dėsnį $F(x)$.

(27) formulė nėra patogi skaičiavimuose, todėl dažnai naudojama tokia jos forma:

$$\omega_n^2 = \frac{1}{12n} + \sum_{i=1}^{n} \left(F(X_i^*) - \frac{2i-1}{2n}\right)^2.$$

Čia $X_1^*, X_2^*, \ldots, X_n^*$ yra atsitiktinių dydžių seka $X_1, X_2, \ldots, X_n$ išrašyta didėjimo tvarka.

Su ω^2 kriterijumi gaunami pakankamai geri rezultatai, kai $n \geq 50$.

ω^2 privalumai ir trūkumai yra faktiškai tie patys kaip ir KS kriterijaus. Skirtumas tarp šių kriterijų toks. KS kriterijus naudoja statistiką D_n, kuri paremta skirtumo (tarp tikrosios ir empirinės pasiskirstymo funkcijų) modulio maksimumu. Tuo tarpu ω^2 kriterijus naudoja statistiką ω_n^2, kuri paremta skirtumo (tarp tikrosios ir empirinės pasiskirstymo funkcijų) kvadrato vidurkiu.

Yra ir daugiau universaliųjų hipotezių tikrinimo testų. Nemažai jų galite rasti [?, ?].

4.2 Empiriniai testai

Visi testai, kuriuos čia pateiksime gali būti panaudoti p.a.s. sekos

$$U_0, U_1, \ldots, U_n, \ldots \tag{28}$$

tolygumui intervale $[0, 1]$ patikrinti. Kai kurie testai tiesiogiai bus naudojami sveikų p.a.s. sekos

$$Y_0, Y_1, \ldots, Y_n, \ldots \pmod{d} \tag{29}$$

tolygumui tarp 0 ir $d - 1$ patikrinti.

Seka (28) gali būti pakeista (29) tipo seka, pavyzdžiui, naudojant formulę:

$$[dU_i] \text{ – sveikų skaičių iš intervalo } [0, d-1] \text{ seka.} \tag{30}$$

Aišku, jei (28) seka yra tolygiai pasiskirsčiusi, tai ir (30) seka yra tokia pati.

4.2.1 Tolygumo tikrinimas

Pirmasis svarbiausias reikalavimas (28) p.a.s. sekai yra jos narių tolygus pasiskirstymas intervale $[0, 1]$. Patikrinti tolygumą galima dviem būdais. Pirma, naudojant KS arba ω^2 kriterijų, palyginama empirinė pasiskirstymo funkcija su tolygaus dėsnio intervale $[0, 1]$ pasiskirstymo funkcija $F(x) = x$, $0 \leq x \leq 1$. Antra, naudojant χ^2 kriterijų. Šiuo atveju intervalą $[0, 1]$ reikia padalinti į r intervaliukų ir suskaičiuoti skaičius sekos narių, pakliūvančių į tuos intervaliukus. 6 ir 7 pavyzdžiuose tikrinamas atsitiktinės sekos tolygumas naudojant χ^2 ir KS kriterijus.

4.2.2 Serijų testas

Šiuo testu tikrinamas (29) sekos gretimų porų tolygumas ir nepriklausomumas. Paskaičiuojama kiek kartų yra sutinkama pora $(Y_{2k}, Y_{2k+1}) = (q, s)$, $0 \leq k < n$. Čia q ir s gali priimti bet kurias reikšmes nuo 0 iki $d-1$. Po to taikomas χ^2 kriterijus su $r = d^2$ (galimų skirtingų porų skaičius) ir lygiomis tikimybėmis $p_i = 1/d^2$. n ir d turėtų būti parinkti taip, kad χ^2 kriterijus duotų bent jau patenkinamus rezultatus. Tokie gaunami, kai $np_i > 5$. Kai $np_i > 20$, gaunami labai geri patikimi rezultatai.

Būtų klaidinga testą naudoti poroms $(Y_0, Y_1), (Y_1, Y_2), \ldots, (Y_{n-1}, Y_n)$. Jos yra stipriai priklausomos ir χ^2 kriterijus joms negali būti taikomas. Žinoma, vietoj porų (Y_{2k}, Y_{2k+1}) galima imti poras (Y_{2k+1}, Y_{2k+2}), $0 \leq k < n-1$. Galima tikrinti ir vienas ir kitas.

Šis testas gali būti apibendrintas trejetams, ketvertams ir t.t. Bet tuomet d reikšmės turėtų būti smarkiai sumažintos, kad būtų išlaikoma bent jau sąlyga $np_i > 5$. Todėl junginiams iš keturių ir daugiau elementų naudojami ne tokie tikslūs testai. Apie juos pakalbėsime kituose skyreliuose.

4.2.3 Intervalų testas

Šiame teste paimamas intervalas $[\alpha, \beta) \in [0, 1]$ ir (28) seka suskirstoma į dalis (intervalus) $U_j, U_{j+1}, \ldots, U_{j+k}$, kuriose tik paskutinis narys $U_{j+k} \in [\alpha, \beta)$. Sakoma, kad tokia $k+1$ skaičiaus seka apibrėžia k ilgio intervalą. Generavimo proceso metu reikia generuoti tiek dydžių U_i, kad turėtume iš viso n intervalų. (Teoriškai toks procesas gali būti begalinis, nes intervalo ilgis irgi gali būti kiek norima didelis. Dėl to kartais apsiribojama generuotos sekos ilgiu ir suskaičiuojamas intervalų skaičius.) Pasirenkamas skaičius t ir suskaičiuojama kiek yra 0 ilgio intervalų, kiek yra 1 ilgio intervalų, ir t.t., pagaliau, kiek yra t ilgio intervalų ir kiek yra intervalų, kurių ilgis didesnis už t. Taigi suskirstome atsitiktinius dydžius (intervalų ilgius) į $r = t + 2$ grupes. Pažymėkime $\beta - \alpha = p$. Nesunku suprasti, kad tikimybės p_i, jog intervalas turi ilgį i yra lygios:

$$p_i = P(\text{intervalo ilgis } = i) = p(1-p)^i,\ 0 \leq i \leq t,$$

$$p_{>i} = P(\text{intervalo ilgis } > t) = (1-p)^{t+1}.$$

Intervaliukų skaičių r turime, tikimybes p_i turime. Galima taikyti χ^2 kriterijų, tik reikia parinkti n ir t taip, kad $np_i > 5$.

Dažnai imama $\alpha = 0$ arba $\beta = 1$. Taip supaprastinami skaičiavimai. Daliniai atvejai, kai $[\alpha, \beta) = [0, 1/2)$ arba $[\alpha, \beta) = [1/2, 1)$ naudojami atsitiktinės sekos nukrypimui nuo vidurkio į vieną arba į kitą pusę patikrinti.

4.2.4 Kėlinių testas

Pradinė seka suskaidoma į n grupių po k elementų kiekvienoje iš jų: $(U_{jk}, U_{jk+1}, \ldots, U_{jk+k-1})$, $0 \le j < n$. Turint galvoje skaičių išsidėstymą grupėje (jų palyginimo prasme), iš viso gali būti $k!$ skirtingų variantų. Paskaičiuojama kiek kartų kiekvienas konkretus dėstinys sutinkamas tarp n grupių. Tada pritaikomas χ^2 kriterijus su $r = k!$ ir tikimybėmis $p_i = 1/k!$.

Pavyzdžiui, kai $k = 3$, iš viso yra 6 grupės:

$$\text{arba } U_{3j} < U_{3j+1} < U_{3j+2}, \text{ arba } U_{3j} < U_{3j+2} < U_{3j+1},$$

$$\text{arba } U_{3j+1} < U_{3j} < U_{3j+2}, \text{ arba } U_{3j+1} < U_{3j+2} < U_{3j},$$

$$\text{arba } U_{3j+2} < U_{3j} < U_{3j+1}, \text{ arba } U_{3j+2} < U_{3j+1} < U_{3j}.$$

Taikant šį testą tariama, kad tiksli lygybė tarp dviejų sekos narių negalima. Teoriškai tokia tikimybė lygi nuliui. Praktiškai generuodami skaičius, mes juos gauname su tam tikru tikslumu, ir lygybė pasidaro galima. Jei generuojamų sekų periodai ilgi, tai pervedant tuos skaičius į tolygiai pasiskirsčiusius intervale $[0, 1]$ ir apvalinant keleto ženklų po kablelio tikslumu, galima tikėtis, kad lygybė pasitaikys itin retai. Tokias retai pasitaikančias grupes galima išmesti, o nagrinėti tik likusias.

4.2.5 Atstumų testas

Šiame teste nagrinėjamos p.a.s. sekos $U_0, U_1, \ldots, U_n$ gretimos poros (U_0, U_1), $(U_2, U_3), \ldots, (U_{n-1}, U_n)$. Šios poros xOy koordinačių sistemoje reiškia taškus vienetiniame kvadrate $0 \le x \le 1,\ 0 \le y \le 1$. Tarp kiekvienos poros taškų paskaičiuojamas atstumo kvadratas d^2. Jei taškai (U_i, U_{i+1}) vienetiniame kvadrate yra pasiskirstę tikrai atsitiktinai, tai pasiskirstymo funkcija

$$F(u) = P(d^2 < u) = \begin{cases} \pi u - \dfrac{8u^{3/2}}{3} + \dfrac{u^2}{2}, & \text{kai } u \le 1, \\ \dfrac{1}{3} + (\pi - 2)u \\ +4(u-1)^{1/2} + \dfrac{8(u-1)^{3/2}}{3} \\ -\dfrac{u^2}{2} - 4u \operatorname{arcsec} \sqrt{u}, & \text{kai } 1 < u \le 2. \end{cases} \tag{31}$$

Paskaičiuojama empirinė atstumų d^2 pasiskirstymo funkcija. Naudojant KS arba ω^2 kriterijų gauta empirinė funkcija palyginama su tikrąja (31) pasiskirstymo funkcija.

4.2.6 Skaitmenų testai

Šie testai susiję su skaitmenų atsitiktinumu p.a.s. sekoje. Jeigu p.a.s. tikrai atsitiktiniai, tai skaitmenų $0, 1, \ldots, 9$ pasirodymas dešimtainėje p.a.s. sekoje turėtų būti apytiksliai vienodas. Kadangi čia eina kalba apie dešimtainių skaitmenų seką, tai tokią galime gauti iš p.a.s. sekos surašę, pavyzdžiui, juos eilės tvarka. Tik reikia pastebėti, jei generuojame penkiolikaženklius p.a.s. X_i, tai visus ir rašome kaip penkiolikaženklius. Tuo atveju, kai gaunamas X_i su mažiau negu 15 ženklų prirašome priekyje reikiamą skaičių nulių.

1. Pirmo testo esmė tokia. Tegul s vienas iš dešimties skaitmenų $0, 1, \ldots, 9$. Skaičiuokime intervalų tarp dviejų gretimų s ilgius. Jei tarp dviejų gretimų s yra k skaitmenų, nelygių s, tai sakome, kad intervalo ilgis yra lygus k. Jei stovi du s, vienas šalia kito, tai intervalo ilgis yra lygus nuliui. Jei seka tikrai atsitiktinė, tai tikimybė, kad tarp dviejų gretimų s pasitaikys k ilgio intervalas yra lygi

$$P(k) = (0.9)^k 0.1\,.$$

Dideliems k ši tikimybė yra maža, todėl reikėtų jungti didelius k į vieną grupę (pavyzdžiui, kai $k \geq 20$, $P(\geq 20) = 0.9^{20} \approx 0.1216$). Pritaikius χ^2 kriterijų, galima spręsti apie sekos atsitiktinumą.

2. Kitas skaitmenų dažnio tikrinimo testas dar paprastesnis. Reikia suskaičiuoti skaitmenų $0, 1, \ldots, 9$ pasirodymo sekoje empirinius dažnius ir, naudojant χ^2 kriterijų, palyginti juos su teorinėmis tikimybėmis

$$P(\text{skaitmens } s \text{ pasirodymas}) = \frac{1}{10}\,, \quad s = 0, 1, \ldots, 9\,.$$

Praktika rodo, kad pastarasis testas yra gana silpnas.

4.2.7 Maksimumo testas (didžiausias iš k)

Tegul

$$V_j = \max(U_{jk}, U_{jk+1}, \ldots, U_{jk+k-1}),\ 0 \leq j < n.$$

Jei dydžiai U_i nepriklausomi ir tolygiai pasiskirstę intervale $[0, 1]$, tai tikimybė

$$P(V_j < x) = P(U_{jk} < x) \cdot P(U_{jk+1} < x) \cdot \cdots \cdot P(U_{jk+k-1} < x) = x \cdot x \cdot \cdots \cdot x = x^k.$$

Taigi dydžiai $V_0, V_1, \ldots, V_{n-1}$ pasiskirstę pagal dėsnį $F(x) = x^k$, $0 \leq x \leq 1$. Testavimui reikia paskaičiuoti empirinę atsitiktinio dydžio V_j pasiskirstymo funkciją ir taikyti KS arba ω^2 kriterijų.

Vietoje dydžių V_j galima tikrinti dydžių $V_0^k, V_1^k, \ldots, V_{n-1}^k$ tolygumą. Nesunku parodyti, kad šie dydžiai turėtų būti pasiskirstę tolygiai intervale $[0, 1]$.

4.2.8 Kombinacijų tikrinimas (pokerio testas)

1. Klasikiniame pokerio teste sudaroma n grupių po 5 skaičius $(Y_{5j}, Y_{5j+1}, \ldots, Y_{5j+4})$, $0 \le j < n$. Išskiriami 7 kombinacijų tipai:

abcde	(visi skirtingi),
aabcd	(viena pora),
aabbc	(dvi poros),
aaabc	(trys vienos rūšies),
aaabb	(pilnas rinkinys),
aaaab	(keturi vienos rūšies),
aaaaa	(penki vienos rūšies).

Su χ^2 kriterijumi galima tikrinti ar kombinacijų dažniai atitinka teorines tikimybes.

Kad būtų lengvesni skaičiavimai ir paprastesnis algoritmas nustatant kombinacijos tipą, testas supaprastinamas. Paliekami 5 kombinacijų tipai:

5 skirtingi	= visi skirtingi,
4 skirtingi	= viena pora,
3 skirtingi	= dvi poros arba 3 vienos rūšies,
2 skirtingi	= pilnas rinkinys arba 4 vienos rūšies,
nėra skirtingų	= 5 vienos rūšies.

Testas nuo tokio jungimo beveik nepablogėja, nes jungiamos grupės su mažomis tikimybėmis, o skaičiavimai žymiai supaprastėja.

Dydžiai Y_i gali įgyti d skirtingų reikšmių, tai iš viso gali būti d^5 skirtingų rinkinių po 5. Paskaičiavę visų 5 tipų tikimybes, turėsime

$$P(\text{5 skirtingi}) = \frac{d(d-1)\ldots(d-4)}{d^5}\,,$$

$$P(\text{4 skirtingi}) = \frac{d(d-1)(d-2)(d-3)}{d^5}\cdot 10\,,$$

$$P(\text{3 skirtingi}) = \frac{d(d-1)(d-2)}{d^5}\cdot 25\,,$$

$$P(\text{2 skirtingi}) = \frac{d(d-1)}{d^5}\cdot 15\,,$$

$$P(\text{nėra skirtingų}) = \frac{d}{d^5}\,.$$

Tipų skaičių dar galime sumažinti (iš esmės nepabloginant testo) jungdami dvi paskutines grupes, nes jų tikimybės yra labai mažos.

2. Bendru atveju galime nagrinėti n grupių po k elementų kiekvienoje $(Y_{jk}, Y_{jk+1}, \ldots, Y_{jk+k-1})$, $0 \le j < n$. Paskaičiuojama grupių, kuriose yra i skirtingų skaičių, skaičius. Gauti statistiniai dažniai su χ^2 kriterijaus pagalba palyginami su teorinėmis tikimybėmis

$$p_i = \frac{d(d-1)\ldots(d-i+1)}{d^k}\,\sigma_k^{(i)},\ 1 \le i \le k.$$

Čia $\sigma_k^{(i)}$ yra antros rūšies Stirlingo skaičiai. Stirlingo skaičius $\sigma_k^{(i)}$ yra lygus k elementų aibės skirtingų suskaidymų į netuščius poaibius skaičiui. Jie gali būti paskaičiuoti naudojant formulę:

$$\sigma_k^{(i)} = \frac{1}{i!}\sum_{j=1}^{i}(-1)^{i-j}C_i^j j^k\,.$$

Kadangi tikimybės p_1 ir p_2 yra labai mažos, tai šios dvi grupės gali būti jungiamos į vieną.

4.2.9 Pilno rinkinio testas

Skaičiuojama (29) sekos segmentų $Y_{j+1}, Y_{j+2}, \ldots, Y_{j+k}$, kuriuose yra pilnas skaičių nuo 0 iki $d-1$ rinkinys, ilgiai. Praktiškai tai daroma tokiu būdu. (29) seka suskirstoma į segmentus

$$(Y_0, Y_1, \ldots, Y_{j_1}), (Y_{j_1+1}, Y_{j_1+2}, \ldots, Y_{j_2}), \ldots, (Y_{j_{n-1}+1}, Y_{j_{n-1}+2}, \ldots, Y_{j_n}),$$

kurių kiekviename yra skaičių $0, 1, \ldots, d-1$ rinkinys. Nustatomi skaičiai segmentų, turinčių ilgius $d, d+1, \ldots, k-1, \ge k$ $(k > d)$. Taigi atsitiktinis dydis – segmentų ilgis suskaidomas į $r = k-d+1$ grupių. Tada taikomas χ^2 kriterijus, lyginantis gautus dažnius su teorinėmis tikimybėmis

$$p_i = \frac{d!}{d^i}\,\sigma_{i-1}^{(d-1)},\ d \le i < k,$$

$$p_k = 1 - \frac{d!}{d^{k-1}}\,\sigma_{k-1}^{(d)}.$$

Šios tikimybės gaunamos remiantis gana sudėtingomis kombinatorikos formulėmis. Detalesnis paaiškinimas yra knygoje [9].

4.2.10 Monotoniškumo tikrinimas

Yra keletas monotoniškumo tikrinimo testų. Pora jų detaliai aprašyti knygoje [9]. Jie naudoja gana sudėtingą matematinį aparatą, todėl čia mes paaiškinsime tik šių testų esmę.

1. Tarkime, duota seka, sudaryta iš 10 skaitmenų 8 1 3 7 4 5 0 2 9 6. Suskirstome seką į didėjančius posekius:

$$|8|137|45|02|9|6|.$$

Gavome 6 posekius. Toliau reikėtų nagrinėti tų didėjančių posekių ilgius. Trijų posekių ilgiai lygūs 1, dviejų posekių ilgiai lygūs 2, ir vieno posekio ilgis lygus 3. Skirtingai negu ankstesniems testams šių posekių ilgiams negalima betarpiškai taikyti χ^2 kriterijaus, nes tie ilgiai nėra nepriklausomi atsitiktiniai dydžiai. Po ilgų posekių dažniau pasirodo trumpi, o po trumpų dažniau ilgi. Bet paskaičiavus sekų ilgių koreliaciją, galima sudaryti statistiką, kuri pasiskirsčiusi pagal χ^2 dėsnį.

Lygiai taip pat galima tirti ir sekos mažėjančius posekius.

2. Čia, neišvedinėdami matematinių formulių, pateiksime vieną monotoniškumo testą, atsižvelgiantį kartu ir į mažėjančius ir į didėjančius intervalus. Tegul $X_0, X_1, \ldots, X_n$ yra p.a.s. seka. Sudarykime n ilgio dvejetainę seką tokiu būdu. Kiekvienai gretimų narių porai (X_{i-1}, X_i) priskirkime 0, jei $X_{i-1} < X_i$, ir 1, jei $X_{i-1} > X_i$. Gautoje sekoje k nuliukų, einančių tarp dviejų vienetukų, sudaro k ilgio nuliukų posekį. Lygiai taip pat sekoje yra ir įvairių ilgių vienetukų posekiai. Prie posekių priskaičiuojami ir du dvejetainės sekos galuose esantys nuliukų ar vienetukų posekiai. Suskaičiuokime nuliukų ir vienetukų posekius, kurių ilgiai lygūs $1, 2, \ldots, n$. Jeigu p.a.s. seka $X_0, X_1, \ldots, X_n$ tikrai atsitiktinė, nuliukų ir vienetukų posekiai, kurių ilgiai lygūs $1, 2, \ldots, n$, turėtų vidutiniškai pasirodyti taip:

$$E(\text{skaičius posekių, kurių ilgis} = 1) = \frac{5n+6}{12},$$

$$E(\text{skaičius posekių, kurių ilgis} = 2) = \frac{11n-3}{60},$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$E(\text{skaičius posekių, kurių ilgis} = k \mid k < n) = \frac{2[(k^2+3k+1)n - (k^3+2k^2-4k-5)]}{(k+3)!},$$

$$E(\text{skaičius posekių, kurių ilgis} = n) = \frac{2}{(n+1)!}.$$

Turint šiuos vidurkius ir empirinius posekių ilgių pasirodymo p.a.s. sekoje skaičius, galima pritaikyti χ^2 kriterijų.

8 pavyzdys *Panaudodami monotoniškumo testą, patyrinėkime sekos* $X_0, X_1, \ldots, X_{20} =$

$$\begin{array}{cccccc} 0.61 & 0.18 & 0.93 & 0.65 & 0.36 & 0.24 \\ 0.97 & 0.30 & 0.74 & 0.83 & 0.02 & \\ 0.45 & 0.17 & 0.06 & 0.95 & 0.49 & \\ 0.54 & 0.71 & 0.80 & 0.37 & 0.12 & \end{array}$$

atsitiktinumą.

Atitinkanti monotoniškumą dvejetainė seka yra

$$1\,0\,1\,1\,1\,0\,1\,0\,0\,1\,0\,1\,1\,0\,1\,0\,0\,0\,1\,1\,.$$

Teoriniai vidurkiai ir empiriniai dvejetainės sekos intervalų ilgių skaičiai yra tokie:

Posekių ilgis	Teorinis vidurkis	Empirinis skaičius
1	≈ 8.83	8
2	≈ 3.62	3
3	≈ 0.98	2
> 3	≈ 0.24	0

Čia teorinį vidurkį

$$E(\text{skaičius posekių, kurių ilgis} > 3)$$

paskaičiavome remdamiesi (32) formule ir

$$E(\text{ilgiai } > 3) = E(\text{visi ilgiai}) - E(\text{ilgiai } = 1) - E(\text{ilgiai } = 2) - E(\text{ilgiai } = 3).$$

Dabar galima paskaičiuoti statistiką

$$\chi^2_{20} = \frac{(8.83-8)^2}{8.83} + \frac{(3.62-3)^2}{3.62} + \frac{(0.98-2)^2}{0.98} + \frac{(0.24-0)^2}{0.24} \approx 1.49.$$

Iš χ^2 su 3 laisvės laipsniais pasiskirstymo lentelės randame, kad

$$\chi^2(3, 0.05) \approx 7.8.$$

Taigi hipotezės, kad seka $X_0, X_1, \ldots, X_{20}$ yra atsitiktinė (su reikšmingumo lygmeniu 0.05) atmesti negalime. ◄

3. Sekos atsitiktinumo testas taip pat gali būti pagrįstas bendro sekos monotoniškumo intervalų skaičiaus tyrimu. Teorinis tokio skaičiaus vidurkis

$$E(\text{visų monotoniškumo intervalų skaičius}) = \frac{2n+1}{3}\,. \tag{32}$$

Sekos atsitiktinumo hipotezė atmetama, kai visų monotoniškumo intervalų skaičius yra mažas. Tegul I yra visų monotoniškumo intervalų skaičius. Jei seka tikrai atsitiktinė, tai atsitiktinis dydis

$$Z := \frac{I - \dfrac{2n+1}{3}}{\left(\dfrac{16n-13}{90}\right)^{1/2}}$$

yra pasiskirstęs pagal standartinį normalųjį dėsnį. Taigi gali būti sukonstruotas dar vienas monotoniškumo tikrinimo testas.

4. Skyrelio pabaigoje pateiksime testą, kurį reikėtų vadinti ne monotoniškumo testu, bet testu, kuris tiria kaip dažnai p.a.s. seka šokinėja apie vidurkį.

Testo esmė tokia. P.a.s. sekai $U_0, U_1, \ldots, U_n$ apibrėžkime dvejetainę seką $b_0, b_1, \ldots, b_n$ tokiu būdu:

$$b_i := \begin{cases} 0, & \text{jei } b_i < 1/2\,, \\ 1, & \text{jei } b_i > 1/2\,, \end{cases}$$

$i = 0, 1, \ldots, n$. Toliau kaip ir ankstesniame monotoniškumo teste skaičiuojami dvejetainės sekos intervalų tarp gretimų nuliukų ir gretimų vienetukų ilgiai. Jei seka $U_0, U_1, \ldots, U_n$ tikrai atsitiktinė, tai k ilgio intervalų skaičiaus vidurkis

$$E(\text{skaičius intervalų, kurių ilgis } = k) = \frac{n-k+4}{2^{k+1}}\,, \quad 1 \le k \le n+1.$$

Visų intervalų skaičiaus vidurkis

$$E(\text{visų intervalų skaičius}) = \frac{n+2}{2}\,.$$

Paskaičiuojama dvejetainės sekos $b_0, b_1, \ldots, b_n$ empiriniai k ilgio intervalų skaičiai. Gauti rezultatai su χ^2 kriterijaus pagalba lyginami su teoriniais vidurkiais. Labai didelis arba labai mažas intervalų skaičius reiškia, kad p.a.s. seka nepasižymi atsitiktinės sekos savybėmis.

4.2.11 Gretimų narių koreliacija

Statistikoje dažnai naudojamas koreliacijos koeficientas. Tarkime, turime du skirtingų dydžių rinkinius $U_0, U_1, \ldots, U_{n-1}$ ir $V_0, V_1, \ldots, V_{n-1}$. Jų koreliacijos

koeficientas apibrėžiamas taip:

$$C := \frac{n\sum_{i=0}^{n-1}(U_iV_i) - \left(\sum_{i=0}^{n-1}U_i\right)\left(\sum_{i=0}^{n-1}V_i\right)}{\sqrt{\left(n\sum_{i=0}^{n-1}U_i^2 - \left(\sum_{i=0}^{n-1}U_i\right)^2\right)\left(n\sum_{i=0}^{n-1}V_i^2 - \left(\sum_{i=0}^{n-1}V_i\right)^2\right)}}\,. \tag{33}$$

Koreliacijos koeficientas visuomet yra tarp -1 ir $+1$. Kai jis lygus nuliui ar labai mažas, tai yra nuoroda į dydžių U_i ir V_i nepriklausomumą. Kai $C = \pm 1$, tai dydžiai U_i ir V_i surišti tiesine priklausomybe: $V_i = a + bU_i$, visiems i. Jeigu tikimasi, kad dydžiai U_i ir V_i nepriklausomi, tai laukiama, kad jų koreliacijos koeficientas būtų nedidelis.

1. Statistika

$$C_n := \frac{n(U_0U_1 + \cdots + U_{n-2}U_{n-1} + U_{n-1}U_0) - (U_0 + U_1 + \cdots + U_{n-1})^2}{n(U_0^2 + U_1^2 + \cdots + U_{n-1}^2) - (U_0 + U_1 + \cdots + U_{n-1})^2} \tag{34}$$

yra gretimų sekos $U_0, U_1, \ldots, U_{n-1}$ narių koreliacijos matas. (34) formulė yra (33) formulės atskiras atvejis, vietoje sekos $V_0, V_1, \ldots, V_{n-1}$ paėmus seką $U_1, U_2, \ldots, U_{n-1}, U_0$. Gerai, kai (34) formule apibrėžtas koreliacijos koeficientas yra mažas. Praktiškai jis laikomas geru, kai

$$C_n \in [\mu_n - 2\sigma_n, \mu_n + 2\sigma_n].$$

Čia

$$\mu_n = -\frac{1}{n-1}\,, \quad \sigma_n = \frac{1}{n-1}\sqrt{\frac{n(n-3)}{n+1}}\,, \quad n > 2\,.$$

Faktiškai C_n reikšmė nurodytam intervalui turi priklausyti 95% visų atvejų.

Pageidautina, kad p.a.s. sekoje būtų maža koreliacija ir tarp narių U_i, U_{i+t}, kai $t = 1, 2, \ldots$ Tokią koreliaciją visiškai pakanka nagrinėti, pavyzdžiui, kai $t \le 10$. Išdėstytas metodas tinka ir koreliacijai tarp U_i ir U_{i+t} nagrinėti. Kuo t didesnis, tuo patenkinančios praktines reikmes koreliacijos koeficiento $C_n = C_n(t)$ reikšmės gali būti didesnės.

2. Šiame skyrelyje pateiksime dar vieną koreliacijos tarp narių U_i, U_{i+t} matą. Visa p.a.s. seka $U_0, U_1, U_2, \ldots$ suskirstoma į m grupių po k elementų kiekvienoje grupėje: $(U_{kj}, U_{kj+1}, \ldots, U_{kj+k-1})$, $0 \le j < m-1$. Kiekvienai grupei paskaičiuojama statistika

$$C_{tj} := \frac{1}{k-t}\sum_{i=0}^{k-t-1} U_{kj+i}U_{kj+i+t}\,.$$

Jeigu tarp U_{kj+i} ir U_{kj+i+t} nėra koreliacijos, tai dydis C_{jt} apytiksliai pasiskirstęs pagal normalųjį dėsnį

$$N\left(\frac{1}{4}, \frac{\sqrt{13k-19t}}{12(k-t)}\right).$$

Normališkumui patikrinti reikėtų sudaryti empirinę pasiskirstymo funkciją ir naudoti KS arba ω^2 kriterijų. Gali būti naudojamas ir χ^2 kriterijus, bet tada reikėtų, naudojantis normaliuoju dėsniu, paskaičiuoti teorines C_{tj} pakliuvimo į sudarytus intervalus tikimybes.

4.2.12 Posekių tikrinimas

Neretai modeliuose reikia ne vieno bet keleto p.a.s. rinkinių. Pavyzdžiui, jeigu modelyje yra trys atsitiktiniai skaičiai X, Y, Z, tai, norint juos apibrėžti, reikia tris kartus generuoti p.a.s. Tokiuose taikymuose svarbu, kad geromis atsitiktinėmis sekomis būtų sekos posekiai, minėtu atveju imant kas trečią generuotos sekos narį. Tokiu atveju pakanka generuoti vieną p.a.s. seką. Jeigu modelyje kiekvieną kartą reikia q atsitiktinių skaičių, tai su testų pagalba reikia tikrinti ne pradinę seką $U_0, U_1, U_2, \ldots$, o atskirai jos posekius

$$U_0, U_q, U_{2q}, \ldots;\ U_1, U_{q+1}, U_{2q+1}, \ldots;\ \ldots;\ U_{q-1}, U_{2q-1}, U_{3q-1}, \ldots$$

Patirtis rodo, kad, naudojant tiesinę kongruentinę seką, tokių posekių charakteristikos praktiškai niekada nebūna blogesnės negu pradinės sekos, išskyrus atvejus, kai skaičius q ir tiesinės kongruentinės sekos modulis m turi didelį bendrą daliklį.

Iš visų aprašytų empirinių testų tolygumo tikrinimo ir gretimų narių koreliacijos tikrinimo testai patys silpniausi. Silpniausi ta prasme, kad bandymuose šie testai beveik visada duoda patenkinamus rezultatus. Palyginus stiprus testas yra monotoniškumo testas. Maksimumo testas irgi nėra silpnas testas.

4.3 Teoriniai testai

Nors kiekvieną p.a.s. seką galima tikrinti su empiriniais testais, bet dar geriau žinoti iš anksto kokie bus to tyrimo rezultatai. Tokio tipo teoriniai rezultatai leidžia geriau įvertinti p.a.s. sekas negu empiriniai tyrinėjimai.

Šiame skyrelyje nagrinėsime tik tiesines kongruentines sekas. Rezultatai, liečiantys šias sekas, pagrindinai gaunami atliekant viso sekos periodo statistinę analizę. Pavyzdžiui, tikrinant tokios sekos tolygumą, ji visada bus labai tolygi. Bet patikrinti visą periodą tokiais empiriniais testais kaip serijų, intervalų, kėlinių, gretimų narių koreliacijos ir kitais tikslinga.

4.3.1 Kėlinių testas

Šis testas priklauso teoriniams testams ir skiriasi nuo tokį pat pavadinimą turinčio empirinio testo. Testo esmė tokia. Jei p.a.s. seka gera, tai maždaug pusė porų (X_n, X_{n+1}) turi tenkinti nelygybę $X_{n+1} < X_n$.

9 teorema *Tegul (X_0, a, c, m) yra tiesinė kongruentinė seka su maksimaliu periodu. Beto tegul $d = (a-1, m)$. Tuomet tikimybė*

$$P(X_{n+1} < X_n) = \frac{1}{2} + r.$$

Čia

$$r = \frac{2(c \bmod d) - d}{2m}.$$

Taigi

$$|r| \le \frac{d}{2m}.$$

Iš šios teoremos išplaukia, kad nelygybė visame $X_{n+1} < X_n$ periode pasitaikys su laukiamu tikslumu kaip bepasirinktume a ir c, išskyrus tuos atvejus, kai d yra didelis.

4.3.2 Gretimų narių koreliacija

Gretimų narių koreliacijos koeficientas rodo jų priklausomumo laipsnį. Jis jau buvo apibrėžtas (34) formulėje.

10 teorema *Tiesinės kongruentinės sekos (X_0, a, c, m) su maksimaliu periodu gretimų narių koreliacijos koeficientas*

$$C \approx \frac{1}{a}\left(1 - 6\frac{c}{m} + 6\left(\frac{c}{m}\right)^2\right). \tag{35}$$

Šios apytikslės formulės klaida yra nedidesnė už $(a+6)/m$.

Iš (35) formulės galima padaryti keletą svarbių išvadų. Pirma, reikia vengti mažų a reikšmių. Antra, didelės a reikšmės negarantuoja, kad bus maža gretimų narių koreliacija, nes (35) formulės paklaida gali išaugti iki a/m. Trečia, jei $a \approx \sqrt{m}$, tai koreliacijos koeficiento modulis yra aprėžtas dydžiu $2/\sqrt{m}$.

Iki šiol mes nieko nekalbėjome apie c parinkimą, išskyrus tai (žr. 1 teoremą), kad c ir m turi būti tarpusavyje pirminiai. (35) lygybė padeda

pasirinkti gerus c. Lygties $1 - 6x + 6x^2 = 0$ šaknys yra lygios $\frac{1}{2} \pm \frac{1}{6}\sqrt{3}$. Todėl pasirinkę c taip, kad

$$\frac{c}{m} \approx \frac{1}{2} \pm \frac{1}{6}\sqrt{3}\,, \tag{36}$$

gausime mažą koreliacijos koeficientą.

Pageidautina, kad p.a.s. sekoje būtų maža koreliacija ir tarp X_n ir X_{n+2}. Apskritai neblogai, jei koreliacija nedidelė tarp X_n ir X_{n+t}, kai, sakykime, $1 \le t \le 10$. Anksčiau buvo parodyta (žr. (5) formulę), kad

$$X_{n+t} \equiv a_t X_n + c_t \mod m\,.$$

Čia

$$a_t \equiv a^t \mod m\,,$$
$$c_t \equiv \frac{(a^t - 1)c}{a - 1} \mod m\,.$$

Su šių formulių pagalba galima paskaičiuoti koreliaciją tarp X_n ir X_{n+t}, jeigu vietoje a ir c paimsime a_t ir c_t. Žinoma, c_t parinkimui jau negalėsime naudoti (36) formulės.

4.3.3 Spektrinis testas

Tai svarbus testas 1965 m. pasiūlytas Koveju (R.R.Coveyou) ir Makfersono (R.D.MacPherson). Šis testas geras tuo, kad visos žinomos blogos p.a.s. sekos buvo jo atmestos, o geros praėjo išbandymus sėkmingai. Tai, ko gero, stipriausias ir tobuliausias iš visų žinomų testų.

Teorinis testo pagrindimas naudojasi Furje spektrine analize ir kitu sudėtingu matematiniu aparatu, todėl mes jo detaliai nenagrinėsime, tik išdėstysime kai kuriuos esminius momentus.

Testas naudojamas tiesinės kongruentinės sekos (X_0, a, c, m) daugikliui a parinkti.

11 teorema *Tarkime, kad $s_1, s_2, \ldots, s_l$ yra sveikieji skaičiai, $|s_k| \le m/2$, $1 \le k \le l$,*

$$s(a) := s_1 + s_2 a + s_3 a^2 + \cdots + s_l a^{l-1},$$

$$\nu_l := \min{}^* \sqrt{s_1^2 + s_2^2 + \cdots + s_l^2}. \tag{37}$$

*Čia * reiškia, kad minimumas imamas pagal visus skaičių rinkinius $(s_1, s_2, \ldots, s_l) \ne (0, 0, \ldots, 0)$, tenkinančius sąlygą $s(a) \equiv 0 \pmod m$. Tuomet iš tiesinės kongruentinės sekos (X_0, a, c, m) su maksimaliu periodu m gautos sekos*

$$\frac{X_0}{m}\,, \frac{X_1}{m}\,, \frac{X_2}{m}\,, \ldots \tag{38}$$

tolygaus pasiskirstymo intervale $[0,1]$ *laipsnis toks. Jei imame* l *gretimų* (38) *sekos narių, tai tokie rinkiniai yra nepriklausomi* $1/\nu_l$ *"tikslumu"* (*vidurkinant pagal visą periodą*).

Surasti svarbų tiesinės kongruentinės sekos skaičių ν_l nėra paprasta. Iš 11 teoremos išplaukia, kad ν_l nepriklauso nei nuo tiesinės kongruentinės sekos pradinės reikšmės X_0, nei nuo prieauglio c. Yra algoritmai naudojantys gana sudėtingus algebros ir spektrinės analizės dalykus, su kurių pagalba galima surasti ν_l (žr. [9]). Taip pat įrodyta, kad

$$\nu_l \le \gamma_l m^{1/l}. \tag{39}$$

Čia γ_k, kai $k = 1, 2, \dots, 8$, įgyja tokias reikšmes:

$$1,\ (4/3)^{1/4},\ 2^{1/6},\ 2^{1/4},\ 2^{3/10},\ (64/3)^{1/12},\ 2^{3/7},\ 2^{1/2}.$$

Kad geriau suprastume situaciją, panagrinėkime pavyzdį.

9 pavyzdys *Tegul* $a = 3141592621$, *o* $m = 10^{10}$. *Panagrinėkime sekos*

$$(U_0, U_1, U_2, \dots) = \left(\frac{X_0}{m}, \frac{X_1}{m}, \frac{X_2}{m}, \dots\right)$$

gretimų narių nepriklausomumą, naudodami spektrinį testą (11 *teorema*).

Naudojant sudėtingus algoritmus, galima surasti, kad

$$\nu_2 \approx 67654,$$

$$\nu_3 \approx 1017,$$

$$\nu_4 \approx 250,$$

$$\nu_5 \approx 42,$$

$$\nu_6 \approx 23.$$

Tai reiškia, kad viena po kitos einančios poros (išnaudojant visą periodą) $(U_0, U_1), (U_2, U_3), \dots$ arba $(U_1, U_2), (U_3, U_4), \dots$ yra nepriklausomos, jei apsiribosime tikslumu $1/67654$. T.y. jei užrašas dešimtainis, tai 4-5 ženklai po kablelio nepriklausomi, jei dvejetainis, tai būtų 16 ženklų po kablelio. Vienas po kito einantys trejetai (U_k, U_{k+1}, U_{k+2}) bus nepriklausomi, jei apsiribosime tikslumu $1/1017$. Vieną po kito einančius ketvertus $(U_k, U_{k+1}, U_{k+2}, U_{k+3})$ galima laikyti nepriklausomais, jei apsiribosime tikslumu $1/250$. Penketai nepriklausomi, apsiribojant tikslumu $1/42$, o šešetai, jei apsiribosime tikslumu $1/23$. ◀

Net teoriškai yra sunku pasakyti, kaip parinkti pačius geriausius daugiklius a, t.y. kada ν_l yra maksimaliai galimi. Labai gerai, kai ν_l yra artimi nurodytai (39) formulėje ribai.

Spektrinis testas nėra grynai teorinis. Jis turi daug empirinio testo bruožų. Teorinis testo pagrindas yra 11 teorema. Bruožai, kurie bendri empiriniams testams ir spektriniam testui yra tai, kad, naudojant algoritmus, reikia atlikti kai kuriuos empirinius skaičiavimus. Praktiškai, ν_l reikia skaičiuoti keliems parinktiems a, o paskui parinkti tą a, kurio rodikliai geresni. Aišku, parenkant a svarbiau dvejetų, trejetų, tik po to ketvertų, penketų, šešetų nepriklausomumas. Septynetų, aštuonetų ar didesnių rinkinių nepriklausomumo laipsnis labai retai kada skaičiuojamas.

Patirtis rodo, kad daugiklis a yra neblogas, jei

$$C_l := \frac{\pi^{l/2}\nu_l^l}{(l/2)!\, m} \geq 0.1\,,$$

kai $l = 2, 3, 4$. Čia

$$\left(\frac{n}{2}\right)! = \left(\frac{n}{2}\right)\left(\frac{n}{2}-1\right)\ldots\left(\frac{1}{2}\right)\sqrt{\pi}\,, \text{ kai } n \text{ nelyginis.}$$

Jei $C_2, C_3, C_4 \geq 1$, tai spektrinio testo rezultatai labai geri (a yra geras).

4.3.4 Gardelinis testas

Testas duoda labai panašius rezultatus kaip ir spektrinis testas bei turi ir teorinių ir empirinių testų bruožų. Testo pagrindines idėjas išdėstysime trumpai. Jis taikomas tiesinės kongruentinės sekos daugikliams palyginti ir parinkti tą, kurio testo duomenys geresni. Aptarkime svarbiausius testo momentus.

Tegul tiesinė kongruentinė seka (X_0, a, c, m) yra maksimalaus periodo. Tada seka $X_0, X_1, \ldots, X_{m-1}$ perbėga visus skaičius nuo 0 iki $m-1$. Nagrinėkime koordinačių sistemą xOy ir kvadratą $0 \leq x < m$, $0 \leq y < m$ joje. Šiame kvadrate yra m^2 taškų, kurių koordinatės sveiki skaičiai. Nagrinėkime poras (X_i, X_{i+1}), $i = 0, 1, \ldots, m-1$. Šias poras galima pavaizduoti minėto kvadrato taškais. Kvadrate yra m^2 sveikų taškų, o porų iš viso turime m. Aišku, kad seka $X_0, X_1, \ldots, X_{m-1}$ tuo geresnė, kuo labiau išsibarstę taškai (X_i, X_{i+1}).

Kadangi visi taškai (X_i, X_{i+1}) yra ant lygiagrečių tiesių, tai plokštumoje xOy galima gauti daug lygiagretainių, imant jų viršūnėmis šiuos taškus. Iš geometrinių samprotavimų aišku, kad mažiausio tokio lygiagretainio plotas yra m. Idealu, kai tas mažiausias lygiagretainis yra kvadratas. Tada seka labiausiai panaši į atsitiktinę.

Jeigu nagrinėtume visų vektorių, jungiančių taškus (X_i, X_{i+1}), aibę, tai ši aibė, aišku, yra dvimatė. Jei parinksime du trumpiausius vektorius, kad jie sudarytų bazę, tai tie vektoriai ir bus minėto mažiausio lygiagretainio kraštinėmis. Kuo lygiagretainis artimesnis kvadratui, tuo seka $X_0, X_1, \ldots, X_{m-1}$ tolygiau išsibarsčiusi ir labiau atsitiktinė. Tegul $\vec{\alpha}_1$ ir $\vec{\alpha}_2$ – trumpiausi baziniai vektoriai, $|\vec{\alpha}_2| \geq |\vec{\alpha}_1|$. Atsitiktinumo matu laikomas dydis

$$R(a) = \frac{|\vec{\alpha}_2|}{|\vec{\alpha}_1|}\,.$$

Šis dydis nepriklauso nuo c, o priklauso tik nuo a. Jis gali būti panaudotas ir tiesinėms kongruentinėms sekoms ne su maksimaliu periodu, tame tarpe ir multiplikatyviosioms kongruentinėms sekoms. Naudojant algebros aparatą, galima tuos vektorius surasti. Mes neįsime į teorinius išvedžiojimus, bet pabaigai pateiksime pavyzdį, kad būtų lengviau suprantama tai kas išdėstyta.

10 pavyzdys *Ištirkime tiesinę kongruentinę seką* $(X_0, a, 1, 16)$. *Kuris iš daugiklių* $a = 1$, $a = 5$, $a = 9$ *ar* $a = 10$ *yra geriausias* (*gardelinio testo prasme*)?

a) $a = 1$ b) $a = 5$

c) $a = 9$ d) $a = 13$

9 pav. Gardelinio testo pavyzdys

Iš 1 teoremos išplaukia, kad tik sąlygoje minėtiems a turėsime tiesinę kongruentinę seką $(X_0, a, 1, 16)$ su maksimaliu periodu lygiu 16. Kiekvienam iš minėtų a pavaizduokime xOy plokštumoje taškus (X_i, X_{i+1}), $i = 0, 1, \ldots, 15$. Žiūrėkite *9 pav.* Naudojantis *9 pav.* nesunkiai galima surasti trumpiausius bazinius vektorius $\vec{\alpha}_1$ ir $\vec{\alpha}_2$. Jie pavaizduoti paveikslėliuose. Paskaičiavę turime, kad

$$R(1) = \frac{\sqrt{128}}{\sqrt{2}} = 8,$$

$$R(5) = \frac{\sqrt{26}}{\sqrt{10}} \approx 1.61,$$

$$R(9) = \frac{\sqrt{34}}{\sqrt{8}} \approx 2.06,$$

$$R(13) = \frac{\sqrt{26}}{\sqrt{10}} \approx 1.61.$$

Taigi vienodai geras a pasirinkimas būtų 5 arba 13, blogiau 9, o blogiausias atvejis 1.

Aišku, kai modulis m nedidelis, nesunku nustatyti, kuris daugiklis bus geresnis (kai $m = 16$, geresnis daugiklis jau matosi iš *9 pav.*). Kai moduliai m dideli, nustatyti blogus daugiklius a nėra lengvas uždavinys. Dideliems m surasti $R(a)$ ir palyginti skirtingus a padeda nors ir sudėtingos bet pritaikomos algebrinės procedūros.

5 ĮVAIRIŲ ATSITIKTINIŲ DYDŽIŲ MODELIAVIMAS

Mes jau mokame gauti tolygiai pasiskirsčiusius intervale $[0,1)$ atsitiktinius dydžius $U_0, U_1, \ldots$ Tarkime, kad jie elgiasi taip tarsi būtų atsitiktiniai ir nepriklausomai parinkti.

Praktiniams taikymams dažnai reikia kitaip pasiskirsčiusių atsitiktinių dydžių rinkinių. Mes parodysime kaip tai galima gauti. Yra daug gavimo būdų. Neieškosime geriausių, nes įvertinti, kuris būdas geresnis, būtų kitas – sunkesnis uždavinys.

5.1

Sakykime, reikia gauti atsitiktinį skaičių X tarp 0 ir $k-1$(arba tarp 1 ir k), įgyjantį visas šitas reikšmes su vienoda tikimybe, lygia $\frac{1}{k}$. Pakanka paimti

$$X = [kU] \quad (\text{arba } X = [kU] + 1). \tag{40}$$

5.2

Tarkime, reikia gauti atsitiktinį dydį X tokį, kad

$$\begin{aligned} &X = x_1 \text{ su tikimybe } p_1, \\ &X = x_2 \text{ su tikimybe } p_2, \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ &X = x_k \text{ su tikimybe } p_k, \end{aligned}$$

$p_1 + \cdots + p_k = 1$. Tokiu bus dydis

$$X = \begin{cases} x_1, & \text{kai } 0 \le U < p_1, \\ x_2, & \text{kai } p_1 \le U < p_1 + p_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ x_k, & \text{kai } p_1 + \cdots + p_{k-1} \le U < 1. \end{cases} \tag{41}$$

5.3 Bendri metodai tolydžiai pasiskirsčiusiems dydžiams gauti

Sakykime, reikia gauti dydį X, pasiskirsčiusį pagal pasiskirstymo dėsnį $F(x)$, o $F(x)$ – didėjanti, tolydi pasiskirstymo funkcija. Tuomet egzistuoja atvirkštinė funkcija F^{-1}. Šiuo atveju dydį X galima gauti taip:

$$X = F^{-1}(U). \tag{42}$$

Iš tikrųjų,

$$P(X < x) = P(F^{-1}(U) < x) = P(U < F(x)) = F(x).$$

Pateiksime keletą žinių iš tikimybių teorijos, kuriomis naudojantis galima patobulinti kai kurių atsitiktinių dydžių gavimo procesą.

Jeigu X_1 ir X_2 – du nepriklausomi atsitiktiniai dydžiai su pasiskirstymo funkcijomis $F_1(x)$ ir $F_2(x)$, tai

$$\begin{aligned} &\max(X_1, X_2) \text{ turi pasiskirstymo funkciją } && F_1(x) \cdot F_2(x), \\ &\min(X_1, X_2) \text{ turi pasiskirstymo funkciją } && F_1(x) + F_2(x) - F_1(x) \cdot F_2(x). \end{aligned}$$

Taigi dydžiai

$$X = \sqrt{U} \text{ ir } Y = \max(U_1, U_2)$$

pasiskirstę vienodai. Jų pasiskirstymo funkcija

$$F(x) = \begin{cases} 0, & \text{kai } x < 0, \\ x^2, & \text{kai } 0 \le x < 1, \\ 1, & \text{kai } x \ge 1. \end{cases}$$

Turbūt paprasčiau, tokių dydžių gavimui naudoti antrąjį atvejį. Nereikia traukti šaknies.

Yra ir daugiau panašių dalykų, bet reikia gerai žinoti tikimybių teoriją ir tuo naudotis. Panagrinėkime taip vadinamą sumaišymo metodą. Tegul

$$F(x) = pF_1(x) + (1-p)F_2(x), \quad 0 < p < 1. \tag{43}$$

Galime sumodeliuoti atsitiktinį dydį X su pasiskirstymo funkcija $F(x)$ taip:

$$\begin{aligned} &X \text{ pasiskirstęs pagal } F_1(x), \text{ jei } U < p, \text{ ir} \\ &X \text{ pasiskirstęs pagal } F_2(x), \text{ jei } p \le U < 1. \end{aligned}$$

Toliau konstruojame X, pasiskirsčiusį jau pagal F_1 ar F_2, imdami kitą tolygiai pasiskirsčiusį dydį U_1. Tai padaryti lengviau, negu iš vieno dydžio U modeliuoti X, ieškant (43) funkcijos atvirkštinės funkcijos arba kokiu nors kitu keliu, jei funkcijos F_1 ir F_2 yra paprastesnės už F.

5.4 Įvairių a.d.gavimas

Šio skyriaus tikslas yra išsiaiškinti kaip gaunami įvairūs a.d. iš tolygiai pasiskirsčiusio intervale $[0, 1)$ a.d. U. Egzistuoja keturi pagrindiniai bendrieji

metodai: atvirkštinių transformacijų (inverse transform) metodas, kompozicijų (composition) metodas, priėmimo-atmetimo (acceptance-rejection) metodas ir tolygiųjų dydžių santykio (ratio-of-unforms) metodas. Plačiai naudojami ir įvairūs šių metodų mišiniai. Yra ir kitokių metodų, bet jie ne tiek daug taikomi kaip šie keturi metodai.

Dažnai konkrečiam a.d. gauti galima naudoti ne vieną metodą, bet parenkant metodus galima efektyviau sugeneruoti norimus a.d., t.y. per trumpesnį laiką, arba gauti geresnį tikslumą, jei dydžiai generuojami apytiksliai, arba sutaupyti kompiuterio atmintį. Dar platesniam susipažinimui su įvairių a.d. gavimu žiūrėkite Devroye [4] ir Fishman [6].

5.5 Bendrieji a.d. gavimo metodai

5.6 Atvirkštinių transformacijų metodas

Atvirkštinių transformacijų būdas a.d. gauti yra gana bendras, pats trumpiausias ir tiesioginis.

12 teorema *Tegul $F(z), a \le z \le b$, yra pasiskirstymo funkcija, o*

$$F^{-1}(u) = \inf\{z \in [a,b] : F(z) \ge u, 0 \le u \le 1\} \tag{44}$$

jos atvirkštinė funkcija[9]*. Tegul U yra tolygiai pasiskirstęs a.d. intervale $[0,1)$. Tada a.d.*

$$Z = F^{-1}(U) \tag{45}$$

skirstinys yra F.

Įrodymas.

$$\mathbb{P}(Z \le z) = \mathbb{P}(F^{-1} \le z) = \mathbb{P}(U \le F(z)) = F(z).$$

□

5.6.1 Tolydieji skirstiniai

Šiame skyrelyje pateiksime pagrindinių tolydžiųjų skirstinių atvirkštinių funkcijų lentelę. Tolydiesiems skirstiniams (44) formulė gali būti užrašyta taip:

$$F^{-1}(u) = \min\left\{ z : \int_{-\infty}^{z} f(y)dy \ge u, 0 \le u \le 1 \right\}.$$

[9] Ši funkcija nevisada yra atvirkštinė pagal atvirkštinės funkcijos apibrėžimą, bet visada egzistuoja, kadangi (44) formulėje imamas infimumas.

Čia f yra tolydžiojo skirstinio F tankio funkcija.

Skirstinys	Skirstinio tankis	Parametrų kitimo sritys	$F^{-1}(u)$
$\mathcal{U}(a,b)$ – tolygusis intervale $[a,b)$	$\dfrac{1}{b-a}$	$a \le z \le b$	$a+(b-a)u$
$\mathcal{B}e(\alpha,1)$ – beta	$\alpha z^{\alpha-1}$	$\alpha > 0$ $0 \le z \le 1$	$u^{1/\alpha}$
$\mathcal{B}e(1,\beta)$ – beta	$\beta(1-z)^{\beta-1}$	$\beta > 0$ $0 \le z \le 1$	$1-(1-u)^{1/\beta}$ $\left(\text{arba } 1-u^{1/\beta}\right)$
$\mathcal{E}(\beta)$ – eksponentinis	$\dfrac{1}{\beta}\, e^{-z/\beta}$	$\beta > 0$ $z \ge 0$	$-\beta \ln(1-u)$ $(\text{arba } -\beta \ln u)$
$\mathcal{L}(\alpha,\beta)$ – logistinis	$\dfrac{e^{-(z-\alpha)/\beta}}{\beta(1+e^{-(z-\alpha)/\beta})^2}$	$\beta > 0$ $-\infty < z < \infty$	$\alpha+\beta \ln \dfrac{u}{1-u}$
$\mathcal{C}(\alpha,\beta)$ – necentruotas Koši	$\dfrac{\beta}{\pi(\beta^2+(z-\alpha)^2)}$	$\beta > 0$ $-\infty < z < \infty$	$\alpha+\beta \tan \pi\left(u-\frac{1}{2}\right)$ $\left(\text{arba } \alpha+\dfrac{\beta}{\tan \pi u}\right)$
$\mathcal{P}a(\alpha,\beta)$ – Pareto	$\dfrac{\alpha\beta^\alpha}{z^{\alpha+1}}$	$\alpha > 0$ $z \ge \beta > 0$	$\beta/(1-u)^{1/\alpha}$ $\left(\text{arba } \beta/u^{1/\alpha}\right)$
$\mathcal{W}(\alpha,\beta)$ – Veibulo	$\dfrac{\alpha}{\beta^\alpha}\, z^{\alpha-1} e^{-(z/\beta)^\alpha}$	$\alpha,\beta > 0$ $z \ge 0$	$\beta(-\ln(1-u))^{1/\alpha}$ $\left(\text{arba } \beta(-\ln u)^{1/\alpha}\right)$
$\mathcal{N}(\mu,\sigma^2)$ – normalusis	$\dfrac{e^{-(z-\mu)^2/2\sigma^2}}{(2\pi\sigma^2)^{-1/2}}$	$\sigma^2 > 0$ $-\infty < z < \infty$	$\mu+\sigma H(u)$

Lentelės paskutinėje eilutėje, skaičiuojant normaliojo dėsnio atvirkštinę funkciją dydis $H(u)$ gaunamas apytiksliai:

$$H(u) = \left(\operatorname{sign}\left(u-\frac{1}{2}\right)\right)\left(t - \frac{c_0+c_1t+c_2t^2}{1+d_1t+d_2t^2+d_3t^3}\right),$$

$$t = \left(-2\ln \min(u,1-u)\right)^{1/2},$$

$$c_0 = 2.515517,\ c_1 = 0.802853,\ c_2 = 0.010328,$$

$$d_1 = 1.432788,\ d_2 = 0.189269,\ d_3 = 0.001308.$$

Apytikslio skaičiavimo absoliutinė paklaida $< 0.45 \times 10^{-3}$.

Tiems a.d. Z, kurių skirstinių atvirkštinės funkcijos gaunamos nesunkiai, šis metodas yra gana paprastas ir tiesioginis (ta prasme, kad nereikia nieko

daugiau). Jeigu atvirkštinę funkciją surasti gana sudėtinga ar iš viso negalima, kartai galima naudoti apytikslio skaičiavimo metodus.

Pavyzdžiui, sugeneravus $U \in \mathcal{U}[0,1)$, lygties

$$k(z) = F(z) - U = 0$$

z atžvilgiu sprendimui galima naudoti skaitinius metodus. Apytikslis jos sprendinys duos reikiamą a.d. $Z \approx z$. Jeigu z kitimo intervalas baigtinis, tinka intervalo dalinimo pusiau metodas. Jeigu z kitimo intervalas begalinis ir jame $f(z) > 0$, tai galima naudoti Niutono liestinių metodą.

Tarkime $F(z), a \leq z \leq b$, yra a.d. Z tolydus skirstinys. Paimkime modifikuoto (restricted) (kaip ir Z, tik sukoncentruoto mažesniame intervale) a.d. Z^* skirstinį

$$F^*(z) = \frac{F(z) - F(a')}{F(b') - F(a')} \quad a \leq a' \leq z \leq b' \leq b.$$

A.d. Z^* lengva generuoti, jei mokame gauti a.d. Z. Pakanka a.d. $U \in \mathcal{U}[0,1)$ pakeisti a.d.

$$U' = F(a') + (F(b') - F(a')U$$

ir imti

$$Z^* = F^{-1}(U').$$

5.6.2 Diskretieji skirstiniai

Apie diskrečiuosius skirstinius *** jau kalbėjome, o dabar pateiksime lentelę ir algorirmą kaip generuoti kitus dažnai pasitaikančius diskrečiuosius skirstinius.

Skirstinys	$p_k = \mathbb{P}(Z = k)$	Parametrų kitimo sritys	p_0	$\frac{p_{k+1}}{p_k} = c(k+1)$
$\mathcal{B}(r,p)$ – binominis	$\binom{r}{k} p^k (1-p)^{r-k}$	$k = 0, 1, \ldots, r$ $r \in \mathbb{N}$ $0 < p < 1$	$(1-p)^r$	$\frac{p(r-k)}{(1-p)(k+1)}$
$\mathcal{G}e(p)$ – geometrinis	$(1-p)p^k$	$k = 0, 1, \ldots$ $0 < p < 1$	$1-p$	p
$\mathcal{P}(\lambda)$ – Puasono	$\frac{\lambda^k e^{-\lambda}}{k!}$	$k = 0, 1, \ldots$ $\lambda > 0$	$e^{-\lambda}$	$\frac{\lambda}{k+1}$

Algoritmas 1

Tikslas: diskretaus a.d. Z, įgyjančio reikšmes $k = 0, 1, \ldots$ su tikimybėmis $p_0, p_1, \ldots$, generavimas.
Įvestis: p_0 ir $c(k+1)$ iš pateiktos lentelės.
Išvestis: Z.
Metodas:
$p \leftarrow p_0,\ q \leftarrow p_0,\ Z \leftarrow 0.$
Generuojame a.d. $U \in \mathcal{U}[0,1)$.
Kol $U > q:\ Z \leftarrow Z+1,\ p \leftarrow pc(Z),\ q \leftarrow q+p.$
Atsakymas: Z.

Naudojant šį algoritmą galima gauti ir daugiau diskrečiųjų a.d., tame tarpe hipergeometrinį ir neigiamą binominį, žr. Fishman [7].

Naudojant Algoritmą 1 galima gauti ir modifikuotus (restricted) diskrečiuosius a.d. Z^*, kurie įgyja reikšmes iš siauresnės aibės $k' = a', a'+1, \ldots, b'$ su tikimybėmis

$$\mathbb{P}(Z^* = k') = \frac{p_{k'}}{p_{a'} + \cdots + p_{b'}}.$$

Šiuo atveju Algoritme 1 pradinius duomenis reikėtų įvesti tokius: $p \leftarrow p_{a'},\ q \leftarrow (p_0 + \cdots + p_{a'}),\ Z \leftarrow a'$, ir generuotą a.d. $U \in \mathcal{U}[0,1)$ pakeisti a.d. $(p_0 + \cdots + p_{a'-1}) + U(p_{a'} + \cdots + p_{b'})$.

Pjūvio taško (cutpoint) metodas. Minėto Algoritmo 1 trūkumas, kad a.d. reikšmių mes ieškome iš eilės, didėjimo tvarka. Jei U yra didelis (artimas 1), tai algoritmas dirbs ilgai, paklaidos taip pat gausis didesnės, nes jos priklauso nuo žingsnių skaičiaus algoritme. Algoritmas 1 gali būti pagerintas ta prasme, kad galime pradėti procedūrą ne nuo a, bet nuo taškų, kurie yra arčiau sprendinio Z, jei skaičiuosime vidutiniškai (algoritme atliktų žingsnių vidurkio prasme).

Pjūvio taško metode a.d. reikšmių aibė $a, a+1, \ldots, b$ padalinama taškais

$$I_j = \begin{cases} \min(i : p_a + \cdots + p_i > \frac{j}{m}, a \le i \le b), & j = 0, \ldots, m-1, \\ b, & j = m, \end{cases} \tag{46}$$

į m intervalų. Intervalai parenkami taip, kad tikimybės patekti į kiekvieną iš intervalų $[I_j, I_{j+1})$ būtų kiek galima vienodesnės, $\approx \frac{1}{m}$, žr. Fishman [7]. Pjūvio taško metode a.d. generavimas susideda iš dviejų dalių. Pirma, generuojame $U \in \mathcal{U}[0,1)$ ir parenkame sveiką skaičių $L = [mU]$. Antra, parenkame a.d. $Z \ge I_L$.

Algoritmas 2

Tikslas: diskretaus a.d. Z, įgyjančio reikšmes $k = a, a+1, \ldots, b$ su tikimybėmis $p_a, p_{a+1}, \ldots, p_b$, generavimas.
Įvestis: $p_a, \ldots, p_b$ ir $I_0, \ldots, I_{m-1}$ iš pateiktos (46) formulės.
Išvestis: Z.
Metodas:
Generuojame a.d. $U \in \mathcal{U}[0,1)$.
$L \leftarrow [mU]$.
$Z \leftarrow I_L,\ q \leftarrow p_a + \cdots + p_Z$.
Kol $U > q$: $Z \leftarrow Z + 1$.
Atsakymas: Z.

5.7 Kompozicijų metodas

Tegul a.d. Z, su reikšmėmis intervale $[a,b]$, skirstinys F išsiskaido į sumą:

$$F(z) = \omega_1 F_1(z) + \cdots + \omega_r F_r(z), \tag{47}$$

$$0 < \omega_i < 1,\ i = 1, \ldots, r, \quad \omega_1 + \cdots + \omega_r = 1.$$

Čia $F_i(z), i = 1, \ldots, r$, yra kito a.d. Z_i, su reikšmėmis intervale $[a_i, b_i] \subset [a,b]$, skirstinys.

Šiuo atveju a.d. Z generuoti gali būti naudojamas kompozicijų metodas. Jo esmė tokia. Pirma, generuojamas a.d. Ω, įgyjantis reikšmes $1, \ldots, r$ su tikimybėmis $\omega_1, \ldots, \omega_r$. Tarkime, $\Omega = I$. Antra, generuojamas a.d. Z_I. A.d. Z paimamas lygus Z_I,t.y. $Z = Z_I$.

5.7.1 Išskyrimo į poras metodas

Šiame skyrelyje pateiksime kompozicijų metodo variantą, išskyrimo į poras (alias) metodą, tinkantį bet kokiam diskrečiam a.d. Z, įgyjančiam baigtinį skaičių reikšmių $z_1, z_2, \ldots, z_n$ su tikimybėmis $p_1, p_2, \ldots, p_n$. A.d. Z reikšmių aibė užrašoma dviejų aibių sąjunga:

$$\{z_1, z_2, \ldots, z_n\} = \{A_1, A_2, \ldots, A_n\} \cup \{B_1, B_2, \ldots, B_n\},$$

su priskirtomis tikimybėmis $\mathbb{P}(A_i) > 0$ ir $\mathbb{P}(B_i) = 1 - \mathbb{P}(A_i)$, tokiomis, kad

$$p_i = \frac{1}{n}\sum_{j=1}^{n} f_j(i), \quad f_j(i) = \begin{cases} \mathbb{P}(A_j), & \text{jei } z_i = A_j, \\ \mathbb{P}(B_j), & \text{jei } z_i = B_j, \\ 0 & \text{kitais atvejais.} \end{cases}$$

Taigi skirstinį F išskaidėme:

$$F(z) = \frac{1}{n} F_1(z) + \cdots + \frac{1}{n} F_n(z).$$

Čia F_i, a.d. Z_i, įgyjančio reikšmes A_i, B_i su tikimybėmis $\mathbb{P}(A_i), \mathbb{P}(B_i)$, skirstinys.

Teisinga tokia teorema.

13 teorema *Bet kokio diskretaus a.d., įgyjančio baigtinį skaičių n reikšmių, skirstinys F gali būti išskaidomas į n skirstinių sumą*

$$F(z) = \frac{1}{n} F_1(z) + \cdots + \frac{1}{n} F_n(z),$$

kiekvienas iš kurių sukoncentruotas ne daugiau kaip dviejuose taškuose.

Teoremos įrodymą galite rasti knygoje [7]. Pateiktas įrodymas duoda ir algoritmą, su kuriuo galima surasti aibes $\{A_1, \ldots, A_n\}$, $\{B_1, \ldots, B_n\}$ ir tikimybes $\mathbb{P}(A_1), \ldots, \mathbb{P}(A_n)$. Šis algoritmas taip pat pateiktas Fishman [7].

Dabar pateiksime algoritmą a.d. Z, įgyjančiam reikšmes $z_1, z_2, \ldots, z_n$ su tikimybėmis $p_1, p_2, \ldots, p_n$, generuoti. Naudosime išskyrimo į poras (alias) metodą ir tarsime, kad aibės $\{A_1, \ldots, A_n\}$, $\{B_1, \ldots, B_n\}$ ir tikimybės $\mathbb{P}(A_1), \ldots, \mathbb{P}(A_n)$ yra žinomos.

Algoritmas 3

Tikslas: diskretaus a.d. Z, įgyjančio reikšmes $z_1, z_2, \ldots, z_n$ su tikimybėmis $p_1, p_2, \ldots, p_n$, generavimas.
Įvestis: $\{A_1, \ldots, A_n\}$, $\{B_1, \ldots, B_n\}$, $\mathbb{P}(A_1), \ldots, \mathbb{P}(A_n)$.
Išvestis: Z.
Metodas:
Generuojame a.d. $U_1 \in \mathcal{U}[0, 1)$.
$I \leftarrow 1 + [nU_1]$.
Generuojame a.d. $U_2 \in \mathcal{U}[0, 1)$.
Jei $U_2 \leq \mathbb{P}(A_I)$, $Z \leftarrow A_I$; antraip $Z \leftarrow B_I$.
Atsakymas: Z.

Pabaigoje skyrelio pateiksime pora pavyzdžių kaip sudaryti aibes A_i, B_i ir tikimybes $\mathbb{P}(A_i)$ a.d., įgyjančiam 6 reikšmes.

z_i	p_i	A_i	B_i	$\mathbb{P}(A_i)$	$\mathbb{P}(B_i)$
1	0.1	$A_1 = 1$	$B_1 = 2$	0.6	0.4
2	0.2	$A_2 = 2$	$B_2 = 3$	0.8	0.2
3	0.3	$A_3 = 4$	$B_3 = 3$	0.6	0.4
4	0.1	$A_4 = 5$	$B_4 = 3$	0.6	0.4
5	0.1	$A_5 = 3$	$B_5 = 6$	0.8	0.2
6	0.2	$A_6 = 6$		1	

Arba kitas variantas.

z_i	p_i	A_i	B_i	$\mathbb{P}(A_i)$	$\mathbb{P}(B_i)$
1	0.1	$A_1 = 1$	$B_1 = 2$	0.6	0.4
2	0.2	$A_2 = 2$	$B_2 = 6$	0.8	0.2
3	0.3	$A_3 = 3$		1	
4	0.1	$A_4 = 4$	$B_4 = 3$	0.6	0.4
5	0.1	$A_5 = 5$	$B_5 = 3$	0.6	0.4
6	0.2	$A_6 = 6$		1	

5.8 Normaliai pasiskirsčiusio atsitiktinio dydžio gavimas

1. Tegul, kaip visada, U_1 ir U_2 – nepriklausomi ir tolygiai pasiskirstę intervale $[0, 1)$ atsitiktiniai dydžiai. Tuomet atsitiktiniai dydžiai $V_1 = 2U_1 - 1$ ir $V_2 = 2U_2 - 1$ yra nepriklausomi ir tolygiai pasiskirstę intervale $[-1, 1)$.

Pažymėkime

$$S = V_1^2 + V_2^2. \tag{48}$$

Jeigu $S \geq 1$, imkime naujus U_1 ir U_2 bei generuokime S iš naujo. Imkime tik $0 \leq S < 1$.

Parodysime, kad

$$X_1 = V_1\sqrt{\frac{-2\ln S}{S}}, \quad X_2 = V_2\sqrt{\frac{-2\ln S}{S}} \tag{49}$$

– nepriklausomi ir normaliai pasiskirstę atsitiktiniai dydžiai, t.y.

$$X_1, X_2 \in N(0,1), \text{ su pasiskirstymo funkcija } \Phi(x) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{x} \mathrm{e}^{-\frac{t^2}{2}} \mathrm{d}t.$$

Pereikime prie polinių koordinačių

$$V_1 = R\cos\theta,\ V_2 = R\sin\theta$$
$$\Rightarrow\ S = R^2$$
$$\Rightarrow\ X_1 = \sqrt{-2\ln S}\cos\theta,\ X_2 = \sqrt{-2\ln S}\sin\theta.$$

Pažymėkime

$$R' = \sqrt{-2\ln S}.$$

Dydžiai R ir θ yra nepriklausomi atsitiktiniai dydžiai, kintantys vienetinio skritulio viduje, nes V_1 ir V_2 nepriklausomi. Tuomet R' ir θ taip pat nepriklausomi. Tikimybė papulti taškui į vienodo kampo skritulio išpjovas vienoda (geometrinės tikimybės), todėl θ yra tolygiai pasiskirstęs intervale $[0, 2\pi)$. Tikimybė

$$P(R' < r) = P\left(\sqrt{-2\ln S} < r\right) = P\left(-2\ln S < r^2\right)$$
$$= P\left(\ln S > -\frac{r^2}{2}\right) = P\left(S > \mathrm{e}^{-\frac{r^2}{2}}\right) = P\left(R^2 > \mathrm{e}^{-\frac{r^2}{2}}\right)$$
$$= 1 - P\left(R^2 \le \mathrm{e}^{-\frac{r^2}{2}}\right) = 1 - \mathrm{e}^{-\frac{r^2}{2}},$$

nes atsitiktinis dydis R^2 yra tolygiai pasiskirstęs intervale $[0, 1)$. Tai lengva įsitikinti, paskaičiavus geometrinę tikimybę:

$$P\left(R^2 < t\right) = P\left(R < \sqrt{t}\right) = \frac{\text{skritulio, su spinduliu } \sqrt{t}\text{, plotas}}{\text{skritulio, su spinduliu 1, plotas}} = \frac{\pi t}{\pi} = t,$$
$$0 < t \le 1.$$

Taigi atsitiktinio dydžio R' pasiskirstymo funkcija

$$F_{R'}(r) := P(R' < r) = 1 - \mathrm{e}^{-\frac{r^2}{2}},\ r \ge 0.$$

Jau minėjome, kad dydis θ pasiskirstęs tolygiai intervale $[0, 2\pi)$, todėl jo pasiskirstymo funkcija

$$F_\theta(\varphi) := P(\theta < \varphi) = \frac{\varphi}{2\pi},\ 0 \le \varphi < 2\pi.$$

Dydžiai X_1 ir X_2 bus nepriklausomi ir pasiskirstę pagal $N(0, 1)$, jei

$$P(X_1 < x_1, X_2 < x_2) = P(X_1 < x_1)P(X_2 < x_2) = \Phi(x_1)\Phi(x_2).$$

Įsitikinsime, kad taip yra. Paskaičiuokime

$$\begin{aligned}
P(X_1 < x_1, X_2 < x_2) &= P(R'\cos\theta < x_1, R'\sin\theta < x_2) \\
&= \iint\limits_{\{(r,\varphi)|r\cos\varphi<x_1, r\sin\varphi<x_2\}} \mathrm{d}F_{R',\theta}(r,\varphi) \\
&\overset{R' \text{ ir } \theta \text{ nepriklausomi}}{=} \iint\limits_{\{(r,\varphi)|r\cos\varphi<x_1, r\sin\varphi<x_2\}} \mathrm{d}F_{R'}(r)\,\mathrm{d}F_{\theta}(\varphi) \\
&= \iint\limits_{\{(r,\varphi)|r\cos\varphi<x_1, r\sin\varphi<x_2\}} r\mathrm{e}^{-\frac{r^2}{2}}\frac{1}{2\pi}\,\mathrm{d}r\,\mathrm{d}\varphi \\
&= \left[\begin{matrix} x=r\cos\varphi \\ y=r\sin\varphi \end{matrix} \quad \left|\begin{matrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \varphi} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \varphi} \end{matrix}\right| = r\right] = \frac{1}{2\pi}\iint\limits_{\{(x,y)|x<x_1, y<x_2\}} \mathrm{e}^{-\frac{x^2+y^2}{2}}\,\mathrm{d}x\,\mathrm{d}y \\
&= \left(\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{x_1}\mathrm{e}^{-\frac{x^2}{2}}\,\mathrm{d}x\right)\left(\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{x_2}\mathrm{e}^{-\frac{y^2}{2}}\,\mathrm{d}y\right) = \Phi(x_1)\Phi(x_2).
\end{aligned}$$

2. Normalųjį dydį galima gauti ir kitu būdu. Sakykime,

$$F(x) = p_1F_1(x) + \cdots + p_nF_n(x), \tag{50}$$

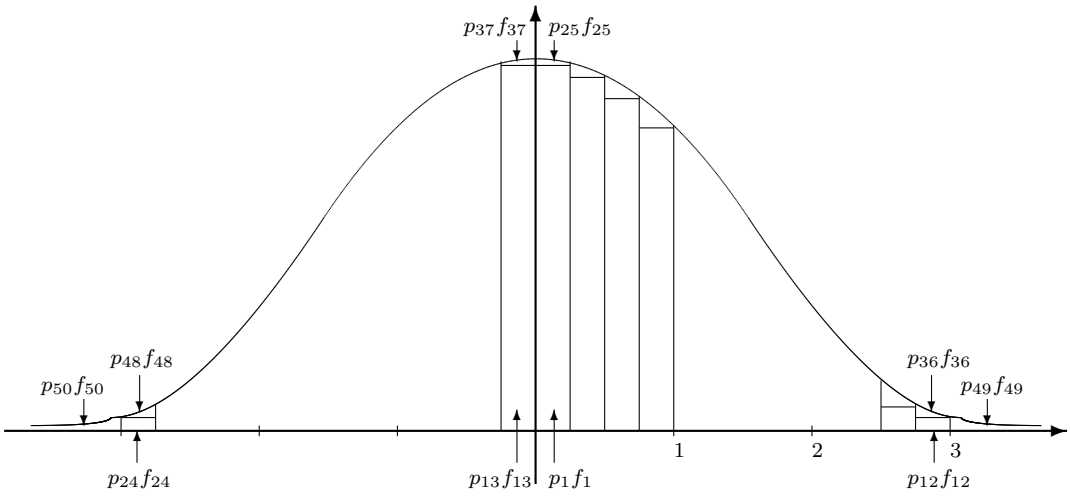
$F_1, \ldots, F_n$ – pasiskirstymo funkcijos, $p_1+\cdots+p_n = 1$. Taigi F – pasiskirstymo funkcija. Dydis X su tikimybe p_i parenkamas pagal pasiskirstymą F_i. Tai daroma taip:

$$X \text{ pasiskirstęs pagal } \begin{cases} F_1, & \text{jei } 0 \le U < p_1, \\ F_2, & \text{jei } p_1 \le U < p_1 + p_2, \\ \quad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ F_n, & \text{jei } p_1 + \cdots + p_{n-1} \le U < 1. \end{cases}$$

Perrašyta tankiams (50) lygybė atrodytų taip:

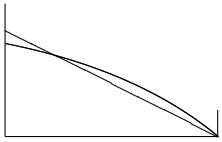
$$f(x) = p_1f_1(x) + \cdots + p_nf_n(x).$$

Tegul $f(x)$ normalusis tankis. Paimkime $n = 50$ ir sukonstruokime tankius $f_i(x)$ ir tikimybes p_i. Tegul tankiai (žr. 1 pav.) $f_1, \ldots, f_{24}$ yra tolygiai pasiskirsčiusių atsitiktinių dydžių tankiai ir užima didžiąją dalį ploto ($p_1 + \cdots + p_{24} > 0,9$). Todėl skaičiuoti X reikšmę pagal pasiskirstymus $F_{25}, \ldots, F_{50}$ retai kada reikės. Pastarieji pasiskirstymai ir daug sudėtingesni.

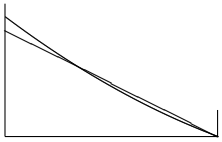


1 pav.

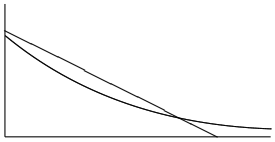
Jeigu pasitaiko tankiai $f_{25}, \ldots, f_{48}$, tai jie paprastai keičiami apytikriai trikampiais tankiais (žr. 2 ir 3 pav.). Kai tankiai trikampiai, jų pasiskirstymo funkcijos paprastesnės.



2 pav.



3 pav.



4 pav.

Su paskutiniais tankiais f_{49}, f_{50} elgiamasi analogiškai. Jie irgi keičiami trikampiais tankiais (žr. 4 pav.).

Užprogramuoti tokią procedūrą gana sudėtinga. Kadangi normalusis pasiskirstymas labai dažnai pasitaiko, tai kompiuteryje tokią programą verta turėti.

3. Normalusis atsitiktinis dydis su vidurkiu μ ir dispersija σ^2, o taip pat koreliuoti normalieji dydžiai. Dydis

$$Y = \mu + \sigma X, \ X \in N(0,1), \tag{51}$$

yra pasiskirstęs pagal $N(\mu, \sigma)$.

Tegul dydžiai $X_1, X_2 \in N(0,1)$ yra nepriklausomi. Tuomet

$$Y_1 = \mu_1 + \sigma_1 X_1, \ Y_2 = \mu_2 + \sigma_2 \left(\rho X_1 + \sqrt{1-\rho^2} X_2 \right), \tag{52}$$

yra pasiskirstę atitinkamai pagal $N(\mu_1, \sigma_1)$ ir $N(\mu_2, \sigma_2)$, o jų koreliacijos koeficientas lygus ρ (jie priklausomi).

Iš tikrųjų

$$DY_1 = \sigma_1^2,$$
$$DY_2 = \sigma_2^2 \left(\rho^2 DX_1 + \left(1 - \rho^2\right) DX_2\right) = \sigma_2^2,$$

$$M\left(\frac{Y_1 - \mu_1}{\sigma_1} \cdot \frac{Y_2 - \mu_2}{\sigma_2}\right)$$
$$= M\left(X_1\left(\rho X_1 + \sqrt{1-\rho^2} X_2\right)\right) = \rho M X_1^2 + \sqrt{1-\rho^2}\, M X_1\, M X_2$$
$$= \rho M X_1^2 = \rho\left(M X_1^2 - (M X_1)^2\right) = \rho D X_1 = \rho.$$

5.9 Eksponentinio atsitiktinio dydžio gavimas

Jei eksponentiškai pasiskirsčiusio atsitiktinio dydžio X vidurkis $MX = 1$, tai jo pasiskirstymo funkcija $F(x) = 1 - \mathrm{e}^{-x}$ ir

$$F^{-1}(y) = -\ln(1-y). \tag{53}$$

Taigi

$$X = -\ln(1-U).$$

Kadangi $1 - U$ taip pat tolygiai pasiskirstęs intervale $(0, 1]$, tai galima imti

$$X = -\ln U.$$

Tik reikia arba išmesti reikšmes, kai $U = 0$, arba jas pakeisti reikšmėmis $U = 1$.

Ir čia galima naudoti metodą, kurį naudojome normaliojo atsitiktinio dydžio atveju. Skirstinį galima išskaidyti į skirstinių sumą.

Taip pat pastebėkime, jei X turi eksponentinį pasiskirstymą ir $MX = 1$, tai

$$Y = \mu X$$

taip pat turės eksponentinį pasiskirstymą ir $MY = \mu$.

5.10 Atsitiktinių dydžių, turinčių χ^2 skirstinį su ν laisvės laipsnių (arba $\nu/2$ eilės gama skirstinį), gavimas

Atsitiktinio dydžio χ^2, su ν laisvės laipsnių, pasiskirstymo funkcija

$$F(x) = \frac{1}{2^{\frac{\nu}{2}}\Gamma\left(\frac{\nu}{2}\right)} \int_0^x t^{\frac{\nu}{2}-1} \mathrm{e}^{-\frac{t}{2}} \mathrm{d}t, \quad x \geq 0. \tag{54}$$

Tokį atsitiktinį dydį galima gauti, turint nepriklausomų eksponentiškai pasiskirsčiusių atsitiktinių dydžių rinkinį. Be to, jeigu ν yra nelyginis skaičius, reikalingas nepriklausomas nuo eksponentinių dydžių atsitiktinis dydis, turintis standartinį normalųjį skirstinį. Būtent dydis

$$X = \begin{cases} 2(Y_1 + \cdots + Y_k), & \text{kai } \nu = 2k, \\ 2(Y_1 + \cdots + Y_k) + Z^2, & \text{kai } \nu = 2k+1, \end{cases} \tag{55}$$

yra pasiskirstęs pagal χ^2 dėsnį su ν laisvės laipsnių. Čia $Y_1, \ldots, Y_k$ – nepriklausomi eksponentiškai pasiskirstę (su vidurkiu, lygiu 1) atsitiktiniai dydžiai, o $Z \in N(0,1)$ nepriklauso nuo Y_i .

Atsitiktinis dydis, turintis (54) pasiskirstymo funkciją, dar vadinamas atsitiktiniu dydžiu, turinčiu $\nu/2$ eilės gama pasiskirstymą.

5.11 Atsitiktinių dydžių, turinčių beta skirstinį su ν_1 ir ν_2 laisvės laipsnių, gavimas

Atsitiktinis dydis X turi beta pasiskirstymą su ν_1 ir ν_2 laisvės laipsnių, jei jo pasiskirstymo funkcija

$$F(x) = \frac{\Gamma\left(\frac{\nu_1+\nu_2}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)\Gamma\left(\frac{\nu_2}{2}\right)} \int_0^x t^{\frac{\nu_1}{2}-1}(1-t)^{\frac{\nu_2}{2}-1}\mathrm{d}t, \quad 0 \le x \le 1. \tag{56}$$

Taip pasiskirsčiusį atsitiktinį dydį X galima gauti imant

$$X = \frac{Y_1}{Y_1 + Y_2};$$

čia Y_1 ir Y_2 – nepriklausomi atsitiktiniai dydžiai, turintys χ^2 pasiskirstymą atitinkamai su ν_1 ir ν_2 laisvės laipsnių.

Kitas metodas yra teisingas visoms ν_1, ν_2 reikšmėms, nebūtinai natūraliosioms. Tegul

$$Y_1 = U_1^{\frac{2}{\nu_1}}, \quad Y_2 = U_2^{\frac{2}{\nu_2}};$$

čia U_1, U_2 – tolygiai pasiskirstę intervale $[0,1)$ ir nepriklausomi. Procesą reikia tęsti, pakol gauname, kad $Y_1 + Y_2 \le 1$. Tada dydis

$$X = \frac{Y_1}{Y_1 + Y_2}$$

turi beta pasiskirstymą su ν_1 ir ν_2 laisvės laipsnių.

5.12 Atsitiktinių dydžių, turinčių F skirstinį su ν_1 ir ν_2 laisvės laipsnių, gavimas

Atsitiktinis dydis X turi F pasiskirstymą su ν_1 ir ν_2 laisvės laipsnių, jei jo pasiskirstymo funkcija

$$F(x) = \frac{\nu_1^{\frac{\nu_1}{2}} \nu_2^{\frac{\nu_2}{2}} \Gamma\left(\frac{\nu_1+\nu_2}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)\Gamma\left(\frac{\nu_2}{2}\right)} \int_0^x t^{\frac{\nu_1}{2}-1}(\nu_2+\nu_1 t)^{-\frac{\nu_1}{2}-\frac{\nu_2}{2}} \mathrm{d}t, \quad x \geq 0. \tag{57}$$

Taip pasiskirsčiusį atsitiktinį dydį X galima gauti imant

$$X = \frac{\nu_2 Y_1}{\nu_1 Y_2};$$

čia Y_1 ir Y_2 – nepriklausomi atsitiktiniai dydžiai, turintys χ^2 pasiskirstymą atitinkamai su ν_1 ir ν_2 laisvės laipsnių.

F skirstinį taip pat turi ir dydis

$$X = \frac{\nu_2 Y}{\nu_1(1-Y)};$$

čia Y pasiskirstęs pagal beta skirstinį su ν_1 ir ν_2 laisvės laipsnių.

5.13 Atsitiktinių dydžių, turinčių t skirstinį su ν laisvės laipsnių, gavimas

Atsitiktinis dydis X turi t pasiskirstymą su ν laisvės laipsnių, jei jo pasiskirstymo funkcija

$$F(x) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi\nu}\,\Gamma\left(\frac{\nu}{2}\right)} \int_{-\infty}^x \left(1+\frac{t^2}{\nu}\right)^{-\frac{\nu+1}{2}} \mathrm{d}t. \tag{58}$$

Tegul $Y_1 \in N(0,1)$, o Y_2 nepriklauso nuo Y_1 ir pasiskirstęs pagal χ^2 skirstinį su ν laisvės laipsnių. Tada atsitiktinis dydis

$$X = Y_1 \sqrt{\frac{\nu}{Y_2}}$$

turi t pasiskirstymą su ν laisvės laipsnių.

5.14 n-matės sferos, su spinduliu 1, atsitiktinis taškas

Tegul $X_1, \ldots, X_n \in N(0,1)$ nepriklausomi atsitiktiniai dydžiai. Tuomet dydis

$$\left(\frac{X_1}{r}, \ldots, \frac{X_n}{r}\right), \quad r = \sqrt{X_1^2 + \cdots + X_n^2},$$

yra tolygiai pasiskirstęs n-matės sferos, su spinduliu lygiu 1, paviršiuje.

5.15 Geometrinį pasiskirstymą turinčių atsitiktinių dydžių gavimas

Tegul kažkoks atsitiktinis įvykis įvyksta su tikimybe p. Tikimybė, kad įvykis pirmąkart įvyks n-uoju bandymu, yra lygi $(1-p)^{n-1}p$. Su šiuo atsitiktiniu įvykiu susiekime atsitiktinį dydį N. Sakykime, kad atsitiktinis dydis N yra lygus n, $N = n$, jei nagrinėjamas atsitiktinis įvykis įvyko n-uoju bandymu. Tuomet

$$P(N = n) = (1-p)^{n-1}p, \quad n = 1, 2, 3, \ldots \tag{59}$$

Atsitiktinis dydis N, įgyjantis natūraliąsias reikšmes su (59) tikimybėmis, vadinamas atsitiktiniu dydžiu, pasiskirsčiusiu pagal geometrinį pasiskirstymą su parametru p, $0 < p < 1$.

Geometrinis atsitiktinis dydis gaunamas su formulės

$$N = \left\lceil \frac{\ln U}{\ln(1-p)} \right\rceil \tag{60}$$

pagalba.

Iš tikrųjų

$$P(N = n) = P\left(n - 1 < \frac{\ln U}{\ln(1-p)} \leq n\right)$$
$$= P\left((1-p)^n \leq U < (1-p)^{n-1}\right) = (1-p)^{n-1} - (1-p)^n = (1-p)^{n-1}p.$$

Taigi (60) dydis pasiskirstęs pagal geometrinį skirstinį su parametru p.

5.16 Binominių atsitiktinių dydžių gavimas

Tegul koks nors įvykis įvyksta su tikimybe p. Darome r nepriklausomų eksperimentų. Tuomet tikimybė, kad įvykis įvyks n ($N = n$) kartų, lygi

$$P(N = n) = C_r^n p^n (1-p)^{r-n}. \tag{61}$$

Atsitiktinis dydis N, galintis įgyti reikšmes $n \in \{0, 1, \ldots, r\}$ su (61) tikimybėmis, vadinamas binominiu atsitiktiniu dydžiu su parametrais r ir p.

Atsitiktiniam dydžiui, kuris turi šį pasiskirstymą, gauti nėra tiesioginio metodo, kaip (60) formulėje. Bet, kai bandymų skaičius r didelis, binominis skirstinys aproksimuojamas normaliuoju $N\left(rp, \sqrt{rp(1-p)}\right)$.

5.17 Puasono atsitiktinių dydžių gavimas

Nagrinėkime tokius vienodus nepriklausomus įvykius. Kiekvienas atskiras įvykis gali įvykti bet kuriuo laiko momentu. Sakykime, kad vidutiniškai per laiko vienetą įvyksta μ įvykių. Tegul N yra su šiais atsitiktiniais įvykiais surištas atsitiktinis dydis – per laiko vienetą įvykusių atsitiktinių įvykių skaičius. Tada tikimybė

$$P(N = n) = \frac{\mathrm{e}^{-\mu}\mu^n}{n!}, \quad n \geq 0, \tag{62}$$

vadinama Puasono vardu ir sakoma, kad atsitiktinis dydis N yra psiskirstęs pagal Puasono dėsnį su vidurkiu μ.

Atsitiktinis dydis N, pasiskirstęs pagal Puasono dėsnį su vidurkiu μ, gaunamas tokiu būdu. Tegul $\mathrm{e}^{-\mu} = p$. Generuojame tolygiai intervale $[0, 1)$ pasiskirsčiusį atsitiktinį dydį U_1.

$$\text{Jei } U_1 < p, \text{ tai } N = 0.$$

Jei $U_1 \geq p$, tai generuojame U_2.

$$\text{Jei } U_1 U_2 < p, \text{ tai } N = 1.$$

Ir t.t.
Jei $U_1 \geq p,\ U_1U_2 \geq p,\ \ldots,\ U_1U_2 \ldots U_n \geq p$, tai generuojame U_{n+1}.

$$\text{Jei } U_1 U_2 \ldots U_{n+1} < p, \text{ tai } N = n.$$

Ir t.t.

Iš tikrųjų

$$
\begin{aligned}
&P(N=n)\\
&\quad= P\left(U_1 \ge p,\ U_1U_2 \ge p,\ \ldots,\ U_1U_2\ldots U_n \ge p,\ U_1U_2\ldots U_nU_{n+1} < p\right)\\
&\quad= P\left(U_1 \ge p,\ U_2 \ge \frac{p}{U_1},\ \ldots,\ U_n \ge \frac{p}{U_1\ldots U_{n-1}},\ U_{n+1} < \frac{p}{U_1\ldots U_n}\right)\\
&\qquad= \int_p^1 \mathrm{d}u_1 \int_{\frac{p}{u_1}}^1 \mathrm{d}u_2\ \ldots \int_{\frac{p}{u_1\ldots u_{n-1}}}^1 \mathrm{d}u_n \int_0^{\frac{p}{u_1\ldots u_n}} \mathrm{d}u_{n+1}\\
&\qquad= p\int_p^1 \frac{1}{u_1}\,\mathrm{d}u_1 \int_{\frac{p}{u_1}}^1 \frac{1}{u_2}\,\mathrm{d}u_2\ \ldots \int_{\frac{p}{u_1\ldots u_{n-1}}}^1 \frac{1}{u_n}\,\mathrm{d}u_n\\
&\quad= p\int_p^1 \frac{1}{u_1}\,\mathrm{d}u_1 \int_{\frac{p}{u_1}}^1 \frac{1}{u_2}\,\mathrm{d}u_2\ \ldots \int_{\frac{p}{u_1\ldots u_{n-2}}}^1 \frac{1}{u_{n-1}}\,\ln\frac{u_1\ldots u_{n-1}}{p}\,\mathrm{d}u_{n-1}\\
&= p\int_p^1 \frac{\mathrm{d}u_1}{u_1} \int_{\frac{p}{u_1}}^1 \frac{\mathrm{d}u_2}{u_2}\ \ldots \int_{\frac{p}{u_1\ldots u_{n-3}}}^1 \frac{\mathrm{d}u_{n-2}}{u_{n-2}} \int_{\frac{p}{u_1\ldots u_{n-2}}}^1 \ln\frac{u_1\ldots u_{n-1}}{p}\,\mathrm{d}\ln\frac{u_1\ldots u_{n-1}}{p}\\
&\qquad= \frac{p}{2}\int_p^1 \frac{\mathrm{d}u_1}{u_1} \int_{\frac{p}{u_1}}^1 \frac{\mathrm{d}u_2}{u_2}\ \ldots \int_{\frac{p}{u_1\ldots u_{n-3}}}^1 \left(\ln\frac{u_1\ldots u_{n-2}}{p}\right)^2 \frac{\mathrm{d}u_{n-2}}{u_{n-2}}\\
&\qquad\qquad\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&\qquad\qquad= \frac{p}{(n-2)!}\int_p^1 \frac{\mathrm{d}u_1}{u_1} \int_{\frac{p}{u_1}}^1 \left(\ln\frac{u_1u_2}{p}\right)^{n-2} \frac{\mathrm{d}u_2}{u_2}\\
&\qquad\qquad\qquad= \frac{p}{(n-1)!}\int_p^1 \left(\ln\frac{u_1}{p}\right)^{n-1} \frac{\mathrm{d}u_1}{u_1}\\
&\qquad\qquad\qquad\qquad= \frac{p}{n!}\left(\ln\frac{1}{p}\right)^n = \frac{\mathrm{e}^{-\mu}}{n!}\,\mu^n.
\end{aligned}
$$

Taigi atsitiktinis dydis N turi Puasono skirstinį su vidurkiu μ.

6 Markovo grandinių Monte Karlo metodas

6.1 Įvadas į MGMK metodą

Markovo grandinės yra atsitiktinių procesų klasė, pasižyminti savybe: "memoryless property" – "procesai, kuriuose pamiršta praeitis" – viskas priklauso tik nuo esamos proceso būsenos. Tokius procesus nagrinėja Markovo teorija – viena iš šiuolaikinės tikimybių teorijos pagrindinių sričių. Šios srities negali nestudijuoti tie studentai, kurie užsiiminės atsitiktinių algoritmų dizainu ir taikymais. Markovo grandinės yra pagrindinė sudėtinė tokių algoritmų dalis. Faktiškai kiekvienas atsitiktinis algoritmas gali būti traktuojamas kaip Markovo grandinė (MG).

Dabar pateiksiu keletą uždavinių, kurių sprendimui gali būti sėkmingai taikomas Markovo grandinių Monte Karlo metodas.

> **11 pavyzdys Keliaujančio pirklio uždavinys.** Įsivaizduokime, kad pirklys gyvena viename mieste ir turi aplankyti daug kitų miestų, o po to grįžti namo. Kokia tvarka jis turi keliauti, kad bendras atstumas, kurį jis nukeliauja, būtų trumpiausias?

> **12 pavyzdys Grafo dalinimas pusiau.** Tarkime, turime grafą, kurio viršūnių aibę V sudaro $2k$ viršūnių. Grafo viršūnių aibę V reikia padalyti į dvi lygias dalis V_1 ir V_2, turinčias po k viršūnių, kad briaunų, jungiančių viršūnes, viena iš kurių yra aibėje V_1, o kita aibėje V_2, skaičius būtų minimalus. Šis uždavinys yra, pavyzdžiui, svarbus projektuojant internetines priemones. V galėtų būti aibė visų tinklo puslapių, kuriuose yra surastas pasirinktas konkretus žodis. Briaunos reikštų nuorodas iš vieno puslapio į kitą. Tuomet V_1 ir V_2 turėtų būti svarbus aibės V išskaidymas į dalis (šiuo atveju tikslinga atsisakyti reikalavimo, kad V_1 ir V_2 yra vienodo dydžio aibės). Pavyzdžiui, jeigu ieškomas žodis yra "football", tai galima tikėtis, kad aibėje V_1 bus puslapiai daugiausia apie amerikietišką futbolą, o aibėje V_2 bus puslapiai daugiausia apie mėgėjišką futbolą.

6.2 Kai kas iš tikimybių teorijos

Sakome, kad atsitiktiniai dydžiai $X_1, X_2, \ldots$ yra **nepriklausomi ir vienodai pasiskirstę** (**n.v.p.**), jei

- jie yra nepriklausomi:

$$\mathbf{P}(X_i < x, X_j < y) = \mathbf{P}(X_i < x)\,\mathbf{P}(X_j < y) \quad \forall i, j \text{ ir } \forall x, y;$$

- turi vienodus skirstinius:

$$\mathbf{P}(X_i < x) = \mathbf{P}(X_j < x) \quad \forall i, j \text{ ir } \forall x.$$

Tegul X_n – atsitiktinio dydžio X reikšmė diskrečiu laiko momentu $n = 1, 2, \dots$ Tuomet seka $(X_1, X_2, \dots)$ vadinama **atsitiktiniu procesu** arba **stochastiniu procesu**.

6.3 Markovo grandinės

Pradėkime nuo paprasto pavyzdžio. Sakykime, mieste A gatvės sudaro kvadratą, t.y. yra keturios gatvės ir keturios sankryžos–kampai (žr. 1 pav.). Pažymėkime tas sankryžas v_1, v_2, v_3, v_4. Tarkime, kad pradiniu 0-iu laiko momentu keleivis stovi gatvių kampe v_1. Jis meta monetą, ar eiti keliu pagal laikrodžio rodyklę (kai pasirodo herbas), ar eiti keliu prieš laikrodžio rodyklę (kai pasirodo skaičius). Per laiko vienetą jis pereina visą gatvę ir atsiduria arba viršūnėje v_2, arba v_4. Tada jis vėl meta monetą, ir priklausomai nuo to, pasirodo herbas ar skaičius, jis eina keliu, pasirinktu laikrodžio rodyklės kryptimi, arba keliu, pasirinktu prieš laikrodžio rodyklę. Ir t.t. Jis meta monetą laiko momentais $2, 3, \dots$ ir priklausomai nuo to pasirenka gatvę laikrodžio rodyklės kryptimi arba prieš laikrodžio rodyklę.

Tegul X_n žymi gatvių kampo, kuriame yra keleivis, indeksą n-uoju laiko momentu. Taigi $(X_0, X_1, \dots)$ yra atsitiktinis procesas, priimantis reikšmes iš aibės $\{1, 2, 3, 4\}$. Kadangi keleivis 0-iu laiko momentu stovi sankryžoje v_1, tai

$$\mathbf{P}(X_0 = 1) = 1.$$

Toliau su tikimybe $\frac{1}{2}$ jis nueis į sankryžą v_2 ir su tokia pat tikimybe į sankryžą v_4. Taigi

$$\mathbf{P}(X_1 = 2) = \frac{1}{2} \quad \text{ir} \quad \mathbf{P}(X_1 = 4) = \frac{1}{2}\,. \tag{63}$$

Apskaičiuoti X_n skirstinį, kai $n \geq 2$, reikia šiek tiek daugiau pastangų. Apskaičiuosime vėliau. O dabar panagrinėkime sąlygines tikimybes. Tarkime, kad n-uoju laiko momentu keleivis stovi sankryžoje v_2. Tuomet mes turime dvi sąlygines tikimybes:

$$\mathbf{P}(X_{n+1} = v_1 \,|\, X_n = v_2) = \frac{1}{2}$$

ir

$$\mathbf{P}(X_{n+1} = v_3 \,|\, X_n = v_2) = \frac{1}{2}\,,$$

nes jas apibrėžia monetos sankryžoje v_2 metimas. Faktiškai mes gauname tas pačias sąlygines tikimybes, kai išrašome visą proceso istoriją:

$$\mathbf{P}(X_{n+1} = v_1 \,|\, X_0 = i_0, X_1 = i_1, \ldots, X_{n-1} = i_{n-1}, X_n = v_2) = \frac{1}{2}$$

ir

$$\mathbf{P}(X_{n+1} = v_3 \,|\, X_0 = i_0, X_1 = i_1, \ldots, X_{n-1} = i_{n-1}, X_n = v_2) = \frac{1}{2}$$

bet kokiam $i_0, i_1, \ldots, i_{n-1}$ rinkiniui. Taip yra todėl, kad monetos metimas n-uoju laiko momentu nepriklauso nuo visų ankstesnių monetos metimų, taigi nepriklauso nuo $(X_0, \ldots, X_{n-1})$. Toks reiškinys vadinamas reiškiniu, **neturinčiu atminties** (pamirštama praeitis), o tokia savybė – **Markovo savybe**: dydis X_{n+1}, gautas pažingsniui kaip sekos $X_0, X_1, \ldots, X_n$ sekantis narys, priklauso tik nuo X_n. Kitais žodžiais sakant, prognozuojant kas atsitiks "rytoj" (laikas $n+1$), mums užtenka nagrinėti tik kas atsitinka "šiandien", o "praeitis" (laikai $0, 1, \ldots, n-1$) neduoda jokios papildomos naudingos informacijos.

Kitas įdomus šio atsitiktinio proceso bruožas yra, kad X_{n+1} sąlyginis skirstinys, kai $X_n = v_2$ (tarkime), yra toks pat visiems n (t.y. todėl, kad keleivio sprendimas kur eiti yra parenkamas taip pat bet kuriuo laiko momentu). Tokia savybė vadinama **laiko homogeniškumu** arba tiesiog **homogeniškumu**.

Pereikime prie bendro apibrėžimo.

1 apibrėžimas *Tegul* $\mathbb{P}$ *yra* $k \times k$ *matrica su elementais* $\{P_{i,j} : i, j = 1, \ldots, k\}$. *Atsitiktinis procesas* $(X_0, X_1, \ldots)$ *su baigtine būsenų aibe* $S = \{s_1, \ldots, s_k\}$ *vadinamas* ***(homogenine) Markovo grandine su perėjimo matrica*** $\mathbb{P}$, *jei* $\forall n$, $\forall\, i, j \in \{1, \ldots, k\}$ *ir* $\forall\, i_0, \ldots, i_{n-1} \in \{1, \ldots, k\}$ *teisinga lygybė*

$$\boldsymbol{P}(X_{n+1} = s_j \,|\, X_0 = s_{i_0}, X_1 = s_{i_1}, \ldots, X_{n-1} = s_{i_{n-1}}, X_n = s_i) = \boldsymbol{P}(X_{n+1} = s_j \,|\, X_n = s_i) = P_{i,j}.$$

Perėjimo matricos $\mathbb{P}$ elementai yra vadinami **perėjimo tikimybėmis**. Perėjimo tikimybė $P_{i,j}$ yra sąlyginė tikimybė: "rytoj" pereiti į būseną s_j, jei "šiandien" esama būsenoje s_i. Žodis homogeninė dažnai praleidžiamas, bet turimas galvoje, kai kalbama apie Markovo garndines (MG). Pavyzdžiui, keleivio klaidžiojimas gatvėmis (ankstesniame pavyzdyje) yra Markovo grandinė

su būsenų aibe $\{1, \ldots, 4\}$ ir perėjimo matrica

$$\mathbb{P} = \begin{pmatrix} 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \end{pmatrix}.$$

Kiekviena perėjimo matrica turi tenkinti dvi savybes:

$$P_{i,j} \geq 0 \quad \forall\, i, j \in \{1, \ldots, k\} \tag{64}$$

ir

$$\sum_{j=1}^{k} P_{i,j} = 1 \quad \forall i \in \{1, \ldots, k\}. \tag{65}$$

Savybė (64) reiškia tik, kad tikimybės yra visuomet neneigiamos, o (65) savybė, kad tikimybių suma yra lygi 1, t.y.

$$\mathbf{P}(X_{n+1} = s_1 \,|\, X_n = s_i) + \mathbf{P}(X_{n+1} = s_2 \,|\, X_n = s_i) + \ldots \\ + \mathbf{P}(X_{n+1} = s_k \,|\, X_n = s_i) = 1.$$

Dabar panagrinėkime kitą svarbią Markovo grandinių charakteristiką – **pradinį skirstinį**, kuris pasako kaip MG startuoja. Pradinis skirstinys yra pateikiamas kaip vektorius–eilutė:

$$\mu^{(0)} = (\mu_1^{(0)}, \mu_2^{(0)}, \ldots, \mu_k^{(0)}) = (\mathbf{P}(X_0 = s_1), \mathbf{P}(X_0 = s_2), \ldots, \mathbf{P}(X_0 = s_k)).$$

Kadangi $\mu^{(0)}$ yra tikimybinis skirstinys, tai

$$\sum_{i=1}^{n} \mu_i^{(0)} = 1.$$

Keleivio klaidžiojimo atveju turime

$$\mu^{(0)} = (1, 0, 0, 0),$$

nes $\mathbf{P}(X_0 = 1) = 1$.

Analogiškai vektoriais–eilutėmis $\mu^{(1)}, \mu^{(2)}, \ldots$pažymėkime MG skirstinius laiko momentais $1, 2, \ldots$ Tuomet turėsime, kad

$$\mu^{(n)} = (\mu_1^{(n)}, \mu_2^{(n)}, \ldots, \mu_k^{(n)}) = (\mathbf{P}(X_n = s_1), \mathbf{P}(X_n = s_2), \ldots, \mathbf{P}(X_n = s_k)).$$

Keleivio pavyzdyje iš (63) lygybių išplaukia, kad

$$\mu^{(1)} = \left(0, \frac{1}{2}, 0, \frac{1}{2}\right).$$

Iš tikrųjų, jei pradinis skirstinys $\mu^{(0)}$ ir perėjimo matrica $\mathbb{P}$ yra žinomi, tai galima surasti skirstinius $\mu^{(1)}, \mu^{(2)}, \dots$ tiesiog dauginant matricas.

14 teorema *Tegul*

$$\begin{aligned} (X_0, X_1, \dots) &- \textit{ Markovo grandinė,} \\ \{s_1, \dots, s_k\} &- \textit{ MG būsenų aibė,} \\ \mu^{(0)} &- \textit{ pradinis skirstinys,} \\ \mathbb{P} &- \textit{ perėjimo matrica.} \end{aligned}$$

Tuomet Markovo grandinės skirstinys n-uoju laiko momentu

$$\mu^{(n)} = \mu^{(0)}\mathbb{P}^n. \tag{66}$$

Įrodymas Tegul $n = 1$. Tuomet

$$\mu_j^{(1)} = \mathbf{P}(X_1 = s_j) = \sum_{i=1}^{k} \mathbf{P}(X_0 = s_i, X_1 = s_j)$$

$$= \sum_{i=1}^{k} \mathbf{P}(X_0 = s_i)\mathbf{P}(X_1 = s_j \,|\, X_0 = s_i) = \sum_{i=1}^{k} \mu_i^{(0)} P_{i,j} = (\mu^{(0)}\mathbb{P})_j, \quad j = 1, \dots, k.$$

Čia $(\mu^{(0)}\mathbb{P})_j$ žymi vektoriaus–eilutės $\mu^{(0)}\mathbb{P}$ j-ąjį elementą. Taigi $\mu^{(1)} = \mu^{(0)}\mathbb{P}$.

Lygybės (66) įrodymui bendruoju ayveju naudosime indukcijos metodą. Fiksuokime m ir tarkime, kad (66) lygybė teisinga, kai $n = m$. Kai $n = m+1$, turime

$$\mu_j^{(m+1)} = \mathbf{P}(X_{m+1} = s_j) = \sum_{i=1}^{k} \mathbf{P}(X_m = s_i, X_{m+1} = s_j)$$

$$= \sum_{i=1}^{k} \mathbf{P}(X_m = s_i)\mathbf{P}(X_{m+1} = s_j \,|\, X_m = s_i) = \sum_{i=1}^{k} \mu_i^{(m)} P_{i,j} = (\mu^{(m)}\mathbb{P})_j.$$

Taigi $\mu^{(m+1)} = \mu^{(m)}\mathbb{P}$. Bet pagal indukcinę prielaidą $\mu^{(m)} = \mu^{(0)}\mathbb{P}^m$. Todėl

$$\mu^{(m+1)} = \mu^{(m)}\mathbb{P} = \mu^{(0)}\mathbb{P}^m\mathbb{P} = \mu^{(0)}\mathbb{P}^{m+1}.$$

Teorema įrodyta. □

Dabar panagrinėkime kelis pavyzdžius.

13 pavyzdys V (vidutinio pagal giedras ir apsiniaukusias dienas) **miesto oras.** Kartais tvirtinama, kad geriausia spėti rytojaus orą yra teigti, kad jis bus toks pat kaip šiandien (geriau negu pasitikėti sinoptikų prognoze?). Jeigu mes tarsime, kad toks pasakymas korektiškas, tai galėsime pateikti oro modelį kaip MG. Paprastumui sakykime, kad oras yra tik dviejų tipų: arba lyja, arba giedra. Jeigu mūsų tvirtinimas yra teisingas 75% atvejų (nepaisant koks šiandien oras, ar lyja, ar giedra), tai turime orą atitinkančią MG su būsenų aibe $S = \{s_1, s_2\}$ ($s_1 =$ "lyja", $s_2 =$ "giedra") ir perėjimo matrica

$$\mathbb{P} = \begin{pmatrix} 0,75 & 0,25 \\ 0,25 & 0,75 \end{pmatrix}.$$

Pastebėkime, kad 13 pavyzdyje yra visiška simetrija tarp "lyja" ir "giedra", ta prasme, kad tikimybė, jog oras liks toks pat ar pasikeis, nepriklauso nuo pačio oro. Aišku, kad tokiame mieste turėtų maždaug vienodai būti ir lietingų, ir giedrų dienų.

14 pavyzdys G (giedro) **miesto oras.** Miestui G palikime tą pačią būsenų aibę $S = \{s_1, s_2\}$ kaip ir miestui V, o perėjimo matricą pakeiskime tokia:

$$\mathbb{P} = \begin{pmatrix} 0,5 & 0,5 \\ 0,1 & 0,9 \end{pmatrix},$$

kadangi ji labiau tinkama miestui, kuriame daug giedrų dienų.

1 uždavinys Sumodeliuokite MG **L** (lietingo) **miesto orui.**

15 pavyzdys Internetas – Markovo grandinė. Tarkime, kad mes naršome po internetą ir kiekvieną kartą, kai atsiverčiame naują puslapį, paspaudžiame nuorodą prieš tai atverstame puslapyje. Nuorodas spaudžiame atsitiktinai (ir tolygiai, t.y. visoms puslapyje esančioms nuorodoms paspausti tikimybė vienoda). Jei X_n reiškia vietą internete po n paspaudimų, tai $(X_0, X_1, \dots)$ gali būti aprašytas kaip MG su būsenų aibe S, lygia visų tinklo puslapių internete aibei, ir perėjimo matrica $\mathbb{P}$ su tikimybėmis

$$P_{i,j} = \begin{cases} \frac{1}{d_i}, & \text{jei puslapis } s_i \text{ turi nuorodą į puslapį } s_j, \\ 0 & \text{kitais atvejais;} \end{cases}$$

čia d_i – nuorodų skaičius puslapyje s_i. Kad ši MG būtų visiškai apibrėžta, numatomas ir atvejis, kai puslapis s_i visai neturi nuorodų. Šiuo atveju $P_{i,i} = 1$, o $P_{i,j} = 0$, kai $i \neq j$. Toks atvejis reiškia, kad

atsidūrę tokiame puslapyje toliau negalėsime niekur eiti. Ši MG yra labai sudėtinga (ypač palyginus su 13 ir 14 pavyzdžiais). Bet įvedus eilę supaprastinimų galima gauti įvairių, įdomių ir naudingų išvadų.

Šis modelis yra nagrinėtas literatūroje. Į modelį galima įtraukti ir klavišo "atgal" panaudojimą. Šiuo atveju procesas $(X_0, X_1, \dots)$ jau nebebus MG, kadangi tai kas atsitinka paspaudus klavišą "atgal" laiko momentu n jau priklauso ir nuo praeities, t.y. nuo $X_0, \dots, X_{n-1}$. Bet ir toks modelis gali būti nagrinėjamas naudojant MG teoriją.

Patogus būdas pavaizduoti MG yra **perėjimo grafas**. Perėjimo grafą sudaro mazgai, atitinkantys MG būsenas, ir rodyklės tarp mazgų, atitinkančios perėjimo tikimybes. Paprasčiausia tai paaiškinti pavyzdžiais (žr. x, y ir z pav.).

Visuose pateiktuose pavyzdžiuose, o taip pat ir 1 apibrėžime taisyklė, gaunant X_{n+1} iš X_n, nesikeičia keičiantis laikui. Kai kuriose situacijose vaizdas bus realesnis, jei ši taisyklė keisis keičiantis laikui. Taip mes pereitume prie **nehomogeninės Markovo grandinės**.

2 apibrėžimas *Tegul $\mathbb{P}^{(1)}, \mathbb{P}^{(2)}, \dots$ – $k \times k$ matricų, tenkinančių (64) ir (65) sąlygas, seka. Atsitiktinis procesas $(X_0, X_1, \dots)$ su baigtine būsenų aibe $S = \{s_1, \dots, s_k\}$ vadinamas **nehomogenine Markovo grandine su perėjimo matricomis** $\mathbb{P}^{(1)}, \mathbb{P}^{(2)}, \dots$, jei $\forall n$, $\forall i, j \in \{1, \dots, k\}$ ir $\forall i_0, \dots, i_{n-1} \in \{1, \dots, k\}$ turime, kad*

$$\boldsymbol{P}(X_{n+1} = s_j \,|\, X_0 = s_{i_0}, X_1 = s_{i_1}, \dots, X_{n-1} = s_{i_{n-1}}, X_n = s_i) = \boldsymbol{P}(X_{n+1} = s_j \,|\, X_n = s_i) = P_{i,j}^{(n+1)}.$$

16 pavyzdys Modifikuotas V miesto oro modelis. Yra daug kelių 13 pavyzdyje pateiktą modelį padaryti realistiškesnį. Vienas kelias būtų atsižvelgti į metų laikus: kai prognozuojamas rytojaus oras netaip pat, jei šiandien "sausis" ar "liepa". Todėl praplečiame būsenų aibę $S = \{s_1, s_2, s_3\}$; čia s_1 ="lietus", s_2 ="giedra", s_3 ="sniegas". Tegul

$$\mathbb{P}_{vasara} = \begin{pmatrix} 0,75 & 0,25 & 0 \\ 0,25 & 0,75 & 0 \\ 0,5 & 0,5 & 0 \end{pmatrix} \quad \text{ir} \quad \mathbb{P}_{žiema} = \begin{pmatrix} 0,75 & 0,25 & 0 \\ 0,25 & 0,75 & 0 \\ 0,5 & 0,5 & 0 \end{pmatrix}.$$

Taip pat tarkime, kad oras keičiasi gegužę–rugsėjį pagal $\mathbb{P}_{vasara}$ matricą, o spalį–balandį pagal $\mathbb{P}_{žiema}$ matricą. Turėsime nehomogeninę MG. Pastebėkime, kad gegužę–rugsėjį modelis lieka toks pat kaip 13 pavyzdyje, išskyrus gegužės 1 d., nes šią dieną dar gali pasnigti.

Dabar pateiksime 14 teoremos apibendrinimą nehomogeninėms MG. Remdamiesi šia teorema galėsime paskaičiuoti nehomogeninės MG skirstinius $\mu^{(1)}, \mu^{(2)}, \ldots$ laiko momentais $1, 2, \ldots$, kai duotas pradinis skirstinys $\mu^{(0)}$ ir perėjimo matricos $\mathbb{P}^{(1)}, \mathbb{P}^{(2)}, \ldots$

15 teorema *Tegul*

$$\begin{aligned} (X_0, X_1, \ldots) &- \textit{ nehomogeninė Markovo grandinė}, \\ \{s_1, \ldots, s_k\} &- \textit{ MG būsenų aibė}, \\ \mu^{(0)} &- \textit{ pradinis skirstinys}, \\ \mathbb{P}^{(1)}, \mathbb{P}^{(2)}, \ldots &- \textit{ perėjimo matricos}. \end{aligned}$$

Tuomet nehomogeninės Markovo grandinės skirstinys n-uoju laiko momentu

$$\mu^{(n)} = \mu^{(0)} \mathbb{P}^{(1)} \mathbb{P}^{(2)} \ldots \mathbb{P}^{(n)}.$$

6.4 Markovo grandinės modeliavimas

Pradėkime nuo melagingo teiginio:

> Dauguma programavimo kalbų turi generatorius, kurie sugeneruoja nepriklausomų ir tolygiai pasiskirsčiusių intervale $[0, 1]$ atsitiktinių dydžių seką $U_0, U_1, \ldots$

Tai yra melas dėl dviejų priežasčių.

1. Seka $U_0, U_1, \ldots$, gauta generatoriaus nėra tolygiai pasiskirsčiusi intervale $[0, 1]$. Paprastai šie skaičiai turi baigtinį dvejetainį (ar dešimtainį) skleidinį, ir dėl to yra racionalūs. Priešingai, galima įrodyti, kad atsitiktinis dydis, kuris tikrai tolygiai pasiskirstęs intervale $[0, 1]$ (faktiškai net bet kuris tolydus atsitiktinis dydis) yra iracionalus su tikimybe 1.

2. Seka $U_0, U_1, \ldots$ nėra atsitiktinė, kadangi ji gauta determinuotos procedūros pagalba. Dėl šios priežasties atsitiktinių skaičių generatoriai kartais vadinami (ir tai tiksliau) pseudoatsitiktinių skaičių generatoriais.

Pirmoji priežastis nėra jau tokia svarbi problema, nes skaičių dvejetainės ar dešimtainės išraiškos yra gana ilgos (sakykime, 32 bitų). Per dešimtmečius daug pastangų įdėta į (pseudo)atsitiktinių skaičių generatorių sukūrimą, kurie generuoja skaičius kiek galima panašesnius į tikruosius nepriklausomus ir tolygiai pasiskirsčiusius intervale $[0, 1]$. Šiandien yra sukurti generatoriai, kurie šią priežastį padaro nereikšmingą. Dėl to mes darysime nekorektišką

prielaidą, kad generuoti dydžiai $U_0, U_1, \dots$ yra nepriklausomi ir tolygiai pasiskirstę intervale $[0,1]$ atsitiktiniai dydžiai. Bet taip pat turėsime galvoje, kad šie dydžiai gali būti potencialus klaidų šaltinis sukurtuose kompiuteriniuose modeliuose.

Bet grįžkime prie skyrelio temos. Mums reikia sugeneruoti Markovo grandinę $(X_0, X_1, \dots)$, su duota būsenų aibe $S = \{s_1, \dots, s_k\}$, pradiniu skirstiniu $\mu^{(0)}$ ir perėjimo matrica $\mathbb{P}$. Pagrindinis ingredientas, kaip skaitytojas jau gali nuspėti, yra atsitiktiniai dydžiai $U_0, U_1, \dots$ Taip pat kiti pagrindiniai ingredientai yra dvi funkcijos, vadinamos **inicijuojančia** (initiation) **funkcija** ir **duomenų atnaujinimo** (update) **funkcija**.

Inicijuojanti funkcija $\psi : [0,1] \to S$ yra naudojama pradinio MG dydžio X_0 generavimui. Ji turi tenkinti du reikalavimus:

(A) ψ yra laiptuota funkcija, t.y. intervalas $[0,1]$ gali būti išskaidytas į baigtinį skaičių intervaliukų, kuriuose ψ yra konstanta;

(B) $\forall s \in S$ bendras intervaliukų, kuriuose $\psi(x) = s$, ilgis yra lygus $\mu^{(0)}(s)$.

Savybė (B) gali būti užrašyta ir šiek tiek formaliau:

$$\int_0^1 \mathbf{I}_{\{\psi(x)=s\}} dx = \mu^{(0)}(s) \quad \forall s \in S. \tag{67}$$

Čia $\mathbf{I}_{\{\psi(x)=s\}} : [0,1] \to \{0,1\}$ yra vadinamas aibės $\{\psi(x) = s\}$ **indikatoriumi** ir apibrėžiamas lygybe:

$$\mathbf{I}_{\{\psi(x)=s\}} = \begin{cases} 1, & \text{jei } \psi(x) = s, \\ 0 & \text{kitais atvejais.} \end{cases}$$

Kai jau turime funkciją ψ ir skaičių U_0, galime generuoti X_0. Tiesiog pakanka paimti $X_0 = \psi(U_0)$. Iš tikrųjų, šitaip apibrėžtas X_0 turi reikalingą skirstinį, nes

$$\begin{aligned} \mathbf{P}(X_0 = s) = \mathbf{P}(\psi(U_0) = s) &= (\text{intervalų, kuriuose } \psi(x) = s, \text{ bendras ilgis}) \\ &= \int_0^1 \mathbf{I}_{\{\psi(x)=s\}} dx = \mu^{(0)}(s) \quad \forall s \in S. \end{aligned}$$

Inicijuojančią funkciją nesunku parinkti. Pavyzdžiui, galima paimti, kad ir tokią:

$$\psi(x) = \begin{cases} s_1, & \text{kai } x \in \left[0, \mu^{(0)}(s_1)\right), \\ s_2, & \text{kai } x \in \left[\mu^{(0)}(s_1), \mu^{(0)}(s_1) + \mu^{(0)}(s_2)\right), \\ \dots & \\ s_i, & \text{kai } x \in \left[\sum_{j=1}^{i-1} \mu^{(0)}(s_j), \sum_{j=1}^{i} \mu^{(0)}(s_j)\right), \\ \dots & \\ s_k, & \text{kai } x \in \left[\sum_{j=1}^{k-1} \mu^{(0)}(s_j), 1\right]. \end{cases}$$

Savybė (A) čia yra akivaizdi, o (B) savybė lengvai patikrinama.

Dabar žinome kaip generuoti dydį X_0. Jei žinotume kaip generuoti X_{n+1} kiekvienam n (turint X_n), galėtume šią procedūrą iteraciniu būdu pritaikyti visai grandinei $(X_0, X_1, \dots)$. Kad gautume X_{n+1} turėdami X_n, panaudosime atsitiktinį dydį U_{n+1} ir duomenų atnaujinimo funkciją $\phi : S \times [0,1] \to S$. Panašiai kaip ir inicijuojančiai funkcijai ψ reikia, kad ϕ patenkintų du reikalavimus. Būtent,

(C) kiekvienam fiksuotam s_i funkcija $\phi(s_i, x)$ yra laiptuota funkcija x atžvilgiu;

(D) kiekvienai fiksuotai porai $s_i, s_j \in S$ bendras intervaliukų, kuriuose $\phi(s_i, x) = s_j$, ilgis yra lygus $P_{i,j}$, arba, naudojant indikatorių ir integralą, tai ekvivalentu užrašymui:

$$\int_0^1 \mathbf{I}_{\{\phi(s_i,x)=s_j\}} dx = P_{i,j} \quad \forall s_i, s_j \in S.$$

Jei duomenų atnaujinimo funkcija ϕ tenkina (C) ir (D) reikalavimus, tai pakanka paimti

$$X_{n+1} = \phi(X_n, U_{n+1}).$$

Šitaip apibrėžtas X_{n+1} pasiskirstęs taip kaip reikia, nes

$$\begin{aligned} &\mathbf{P}(X_{n+1} = s_j \,|\, X_n = s_i) = \mathbf{P}(\phi(X_n, U_{n+1}) = s_j \,|\, X_n = s_i) \\ &= \mathbf{P}(\phi(s_i, U_{n+1}) = s_j) = (\text{intervalų, kuriuose } \phi(s_i, x) = s_j, \text{ bendras ilgis}) \\ &\qquad = \int_0^1 \mathbf{I}_{\{\phi(s_i,x)=s_j\}} dx = P_{i,j} \quad \forall s_i, s_j \in S. \end{aligned}$$

Belieka sukonstruoti konkrečią duomenų atnaujinimo funkciją. Tai nėra sunku. Tokia funkcija, pavyzdžiui, galėtų būti funkcija, $\forall s_i \in S$ apibrėžta taip:

$$\phi(s_i, x) = \begin{cases} s_1, & \text{kai } x \in [0, P_{i,1}), \\ s_2, & \text{kai } x \in [P_{i,1}, P_{i,1} + P_{i,2}), \\ \dots & \\ s_j, & \text{kai } x \in \left[\sum_{l=1}^{j-1} P_{i,l}, \sum_{l=1}^{j} P_{i,l}\right), \\ \dots & \\ s_k, & \text{kai } x \in \left[\sum_{l=1}^{k-1} P_{i,l}, 1\right]. \end{cases} \tag{68}$$

Savybė (C) yra akivaizdi, o (D) savybė lengvai patikrinama.

Taigi mes turime viską ko reikia MG generuoti. Sukonstravę inicijuojančią ir duomenis atnaujinančią funkcijas, iš sugeneruotos nepriklausomų ir tolygiai pasiskirsčiusių intervale $[0, 1]$ dydžių sekos $U_0, U_1, \dots$ galime gauti MG:

$$\begin{aligned} X_0 &= \psi(U_0), \\ X_1 &= \phi(X_0, U_1), \\ X_2 &= \phi(X_1, U_2), \\ X_3 &= \phi(X_2, U_3) \end{aligned}$$

ir t.t.

Pažiūrėkime kaip visa tai veikia paprastame pavyzdyje.

17 pavyzdys V miesto oro generavimas. Nagrinėkime 13 pavyzdžio MG, kurios būsėnų aibė $S = \{s_1, s_2\}$, $s_1 =$ "lietus", $s_2 =$ "giedra", o perėjimo matrica

$$\mathbb{P} = \begin{pmatrix} 0,75 & 0,25 \\ 0,25 & 0,75 \end{pmatrix}.$$

Tarkime, MG prasideda lietinga diena, t.y. $\mu^{(0)} = (1, 0)$. Kad sugeneruotume šią MG, naudodami aukščiau aprašytą schemą, inicijuojančią funkciją imame tokią:

$$\psi(x) = s_1 \quad \forall x,$$

o duomenis atnaujinančią funkciją apibrėžtą taip:

$$\phi(s_1, x) = \begin{cases} s_1, & \text{kai } x \in [0, 0.75), \\ s_2, & \text{kai } x \in [0.75, 1], \end{cases} \qquad \phi(s_2, x) = \begin{cases} s_1, & \text{kai } x \in [0, 0.25), \\ s_2, & \text{kai } x \in [0.25, 1]. \end{cases}$$

Šio skyrelio pabaigoje parodysime, kaip išdėstytas metodas gali būti pritaikytas nehomogeninei MG generuoti. Tegul $(X_0, X_1, \dots)$ yra nehomogeninė MG su būsenų aibe $S = \{s_1, \dots, s_k\}$, pradiniu skirstiniu $\mu^{(0)}$ ir perėjimo matricomis $\mathbb{P}^{(1)}, \mathbb{P}^{(2)}, \dots$ Inicijuojančią funkciją ψ ir pradinį dydį X_0 galima gauti lygiai taip pat kaip ir homogeniniame atvejyje. Duomenų atnaujinimas gaunamas panašiai kaip ir homogeniniu atveju. Skirtumas toks, kad MG yra nehomogeninė, todėl reikia visos sekos skirtingų duomenų atnaujinimo funkcijų $\phi^{(1)}, \phi^{(2)}, \dots$, ir visos jos turi tenkinti (C) ir (D) reikalavimus. Reikalavimas (D), suprantamai, truputį keičiasi į tokį:

$$\int_0^1 \mathbf{I}_{\{\phi^{(n)}(s_i,x)=s_j\}}(x)dx = P_{i,j}^{(n)} \quad \forall n \text{ ir } \forall s_i, s_j \in S.$$

Tokios funkcijos gali būti gautos, apibendrinus (68) lygybes:

$$\phi^{(n)}(s_i, x) = \begin{cases} s_1, & \text{kai } x \in \left[0, P_{i,1}^{(n)}\right), \\ s_2, & \text{kai } x \in \left[P_{i,1}^{(n)}, P_{i,1}^{(n)} + P_{i,2}^{(n)}\right), \\ \dots & \\ s_j, & \text{kai } x \in \left[\sum_{l=1}^{j-1} P_{i,l}^{(n)}, \sum_{l=1}^{j} P_{i,l}^{(n)}\right), \\ \dots & \\ s_k, & \text{kai } x \in \left[\sum_{l=1}^{k-1} P_{i,l}^{(n)}, 1\right]. \end{cases}$$

Tokiu būdu nehomogeninė MG generuojama imant

$$\begin{aligned} X_0 &= \psi(U_0), \\ X_1 &= \phi^{(1)}(X_0, U_1), \\ X_2 &= \phi^{(2)}(X_1, U_2), \\ X_3 &= \phi^{(3)}(X_2, U_3) \end{aligned}$$

ir t.t.

6.5 Markovo garandinės II (neredukuojamumas, neperiodiškumas, stacionarumas, apgręžiamumas)

Kaip ir kitose matematikos šakose (ir ne tik matematikos) iš pradžių turime surasti sąlygas. Tokias, kad jos būtų gana griežtos ta prasme, kad galėtume jomis pasirėmę gauti įdomias ir naudingas išvadas. Iš kitos pusės jos turi būti ne per griežtos, kad galėtume jas lengvai patikrinti daugeliui įdomių

pavyzdžių. Šiame skyrelyje mes kalbėsime apie tokias sąlygas MG. Ir šiame skyrelyje ir vėliau nagrinėsime tik homogenines MG, nors galima būtų nagrinėti ir bendrą atvejį (nehomogenines). Rezultatai būtų panašūs, esmė ta pati, bet formulės, apibrėžimai, samprotavimai ir pan. komplikuotesni.

Pradėsime nuo neredukuojamumo sąvokos. Ši savybė žodžiais gali būti pasakyta taip: "iš bet kurios MG būsenos galima pereiti į bet kurią kitą". O dabar pereikime prie tikslesnių apibrėžimų. Nagrinėkime MG $(X_0, X_1, \dots)$ su būsenų aibe $S = \{s_1, \dots, s_k\}$ ir perėjimo matrica $\mathbb{P}$. Sakysime, kad būsena s_i **komunikuoja** su būsena s_j, rašysime $s_i \to s_j$, jei tikimybė kada nors pasiekti MG būseną s_j, kai startuojama būsenoje s_i, yra teigiama. Kitaip sakant, s_i komunikuoja su s_j, jei egzistuoja toks n, kad

$$\mathbf{P}(X_{m+n} = s_j \,|\, X_m = s_i) > 0.$$

Iš MG homogeniškumo išplaukia, kad ši tikimybė nepriklauso nuo m ir yra lygi $(P^n)_{i,j}$.

Jei $s_i \to s_j$ ir $s_j \to s_i$, tai sakysime, kad s_i ir s_j **abipuskomunikuoja**, ir ršysime $s_i \leftrightarrow s_j$. Dabar galime apibrėžti neredukuojamumą.

3 apibrėžimas *Markovo grandinė $(X_0, X_1, \dots)$ su būsenų aibe $S = \{s_1, \dots, s_k\}$ ir perėjimo matrica $\mathbb{P}$ vadinama* ***neredukuojama****, jei $\forall s_i, s_j \in S$ turime, kad $s_i \leftrightarrow s_j$. Kitu atveju MG vadinama* ***redukuojama****.*

Apibrėžimą galėtume perfrazuoti ir taip: MG yra neredukuojama, jei $\forall s_i, s_j \in S$ ezistuoja n toks, kad $(P^n)_{i,j} > 0$.

Lengvas kelias patikrinti ar MG yra neredukuojama yra toks: reikia pasižiūrėti į MG grafą ir patikrinti, ar iš kiekvienos būsenos yra rodyklių grandinė vedanti į kiekvieną kitą būseną. Žvilgsnis į xx, yy ir zz paveikslėlius leidžia įsitikinti, kad xx ,yy pavyzdžiuose, o taip pat klaidžiojančio keleivio pavyzdyje MG yra neredukuojamos. Dabar pateiksime MG pavyzdį, kuri nėra neredukuojama.

18 pavyzdys Nagrinėkime Markovo grandinę $(X_0, X_1, \dots)$ su būsenų aibe $S = \{1, 2, 3, 4\}$ ir perėjimo matrica

$$\mathbb{P} = \begin{pmatrix} 0,5 & 0,5 & 0 & 0 \\ 0,3 & 0,7 & 0 & 0 \\ 0 & 0 & 0,2 & 0,8 \\ 0 & 0 & 0,8 & 0,2 \end{pmatrix}.$$

Pažiūrėjus į grandinės perėjimo grafą (žr. xx pav.), iš karto aišku, kad grandinei startavus 1 arba 2 būsenoje, ji tose būsenose visam laikui

ir pasilieka. Panašiai, jei grandinė startuoja 3 arba 4 būsenoje, tai ji niekada nepalieka būsenų poaibio $\{3,4\}$. Taigi MG yra redukuojama.

Pastebėkime, jei grandinė startuoja 1 ar 2 būsenoje, tai ji elgiasi lygiai taip pat lyg ji būtų MG su būsenų aibe $\{1,2\}$ ir perėjimo matrica

$$\begin{pmatrix} 0,5 & 0,5 \\ 0,3 & 0,7 \end{pmatrix}.$$

Jei ji startuoja 3 ar 4 būsenoje, ji elgiasi taip tarsi būtų MG su būsenų aibe $\{3,4\}$ ir perėjimo matrica

$$\begin{pmatrix} 0,2 & 0,8 \\ 0,8 & 0,2 \end{pmatrix}.$$

Tai paaiškina redukuojamos MG pavadinimą "redukuojama": jei MG yra redukuojama, tai jos elgesio tyrimas gali būti išskaidytas į dviejų arba daugiau MG su mažesnėmis būsenų aibėmis tyrimą.

Pereikime prie neperiodiškumo sąvokos. Būsenos $s_i \in S$ **periodas** $d(s_i)$ apibrėžiamas lygybe:

$$d(s_i) := DBD\{n \geq 1 : (P^n)_{i,i} > 0\};$$

čia DBD reiškia aibės, nurodytos riestiniuose skliaustuose, didžiausią bendrąjį daliklį. Sakant žodžiais, s_i periodas yra laikų, per kuriuos grandinė gali sugrįžti į būseną s_i, aibės didžiausias bendrasis daliklis. Jei $d(s_i) = 1$, būsena s_i vadinama **neperiodine**.

4 apibrėžimas *Markovo grandinė vadinama* ***neperiodine****, jei visos jos būsenos yra neperiodinės. Kitais atvejais grandinė avdinama* ***periodine****.*

Prisiminkime 13 pavyzdį (V miesto oras). Lengva patikrinti, kad, kokį orą mes šiandien beturėtume, lietingą ar giedrą, tikimybė išlikti tokiam pat orui dar n dienų yra teigiama: $(P^n)i, i > 0\ \forall n$ ir $\forall s_i$. Taigi MG 13 pavyzdyje yra neperiodinė. Taip pat argumentuodami įsitikintume, kad ir 14 pavyzdyžio (G miesto oras) MG yra neperiodinė.

Panagrinėkime klaidžiojančio keleivio modelį (žr. xx pav.). Tegul keleivis 0-iu laiko momentu stovi sankryžoje v_1. Aišku, jis turės padaryti lyginį skaičių ėjimų, kad vėl grįžtų į v_1. Tai reiškia, kad $(P^n)_{1,1} > 0$ tik lyginiams n. Vadinasi,

$$DBD\{n \geq 1 : (P^n)_{i,i} > 0\} = DBD\{2,4,6,\dots\} = 2.$$

Taigi ši MG yra periodinė.

Svarbi ir naudinga yra tokia neperiodiškumo savybė.

16 teorema *Tarkime, kad* $(X_0, X_1, \dots)$ *yra neperiodinė Markovo grandinė su būsenų aibe* $S = \{s_1, \dots, s_k\}$ *ir perėjimo matrica* $\mathbb{P}$. *Tuomet* $\exists N$ *toks, kad*

$$(P^n)i, i > 0 \quad \forall i \in \{1, \dots, k\} \text{ ir } \forall n \geq N.$$

Kitas svarbus rezultatas tesingas neperiodinėms ir neredukuojamoms MG.

17 teorema *Tarkime, kad* $(X_0, X_1, \dots)$ *yra neredukuojama ir neperiodinė Markovo grandinė su būsenų aibe* $S = \{s_1, \dots, s_k\}$ *ir perėjimo matrica* $\mathbb{P}$. *Tuomet* $\exists M$ *toks, kad*

$$(P^n)i, j > 0 \quad \forall i, j \in \{1, \dots, k\} \text{ ir } \forall n \geq M.$$

Panagrinėkime kaip elgiasi MG per ilgą laiko tarpą, arba apie MG asimptotinį elgesį. Pavyzdžiui, jei $P_{i,i} = 1$, MG tampa visiškai apibrėžta, visada lieka būsenoje s_i. Ar gali X_n skirstinys turėti ribinį skirstinį (prie kurio artėja laikui bėgant)?

Grįžkime prie 14 pavyzdžio (G miesto oras). Jeigu vietoje pradinio skirstinio $\mu^{(0)}$paimtume $\mu^{(0)} = \left(\frac{1}{6}, \frac{5}{6}\right)$, tai X_n skirstinys išliktų visą laiką nepasikeitęs, t.y. $\mu^{(n)} = \mu^{(0)}$ $\forall n$. Joks kitas skirstinys neturi tokios savybės. Skirstinys $\left(\frac{1}{6}, \frac{5}{6}\right)$ šiai MG yra ypatingas ir vadinamas **stacionariuoju skirstiniu**. Dar kartais vadinama **invariantiniu** arba **pusiausvyros skirstiniu**.

5 apibrėžimas *Tegul* $(X_0, X_1, \dots)$ *yra Markovo grandinė su būsenų aibe* $S = \{s_1, \dots, s_k\}$ *ir perėjimo matrica* $\mathbb{P}$. *Vektorius–eilutė* $\pi = (\pi_1, \dots, \pi_k)$ *yra vadinamas MG* ***stacionariuoju skirstiniu****, jei*

(i) $\pi_i \geq 0 \ \ \forall i = 1, \dots, k$ *ir* $\sum_{i=1}^{k} \pi_i = 1$,

(ii) $\pi P = \pi$, *t.y.* $\sum_{i=1}^{k} \pi_i P_{i,j} = \pi_j \ \ \forall j = 1, \dots, k$.

Sabybė (i) paprasčiausiai reiškia, kad π yra tikimybinis skirstinys aibėje $\{s_1, \dots, s_k\}$. Iš (ii) savybės išplaukia, kad, paėmus pradinį skirstinį $\mu^{(0)}$ lygų π, ir skirstinys $\mu^{(1)}$ liks nepasikeitęs:

$$\mu^{(1)} = \mu^{(0)}\mathbb{P} = \pi\mathbb{P} = \pi.$$

Ir taip kiekvienam n: $\mu^{(n)} = \pi$.

Kadangi stacionaraus skirstinio apibrėžimas priklauso tik nuo perėjimo matricos $\mathbb{P}$, tai mes dažnai sakysime, kad skirstinys π, tenkinantis (i) ir (ii) sąlygas, yra **stacionarus matricai** $\mathbb{P}$ (o ne MG).

Mums prireiks dar keleto naujų sąvokų. Tegul MG $(X_0, X_1, \dots)$ su būsenų aibe $\{s_1, \dots, s_k\}$ ir perėjimo matrica $\mathbb{P}$ startuoja būsenoje s_i. Tuomet dydis

$$T_{i,j} = \min\{n \geq 1 : X_n = s_j\}$$

vadinamas s_j būsenos **aplankymo laiku** (hitting time), kai startuojama būsenoje s_i. Dydis

$$\tau_{i,j} = \mathbf{E}[T_{i,j}]$$

vadinamas būsenos s_j **vidutiniu aplankymo laiku**, kai startuojama būsenoje s_i. Tai reiškia, kad $\tau_{i,j}$ yra labiausiai tikėtinas laikas, per kurį patenkama į būseną s_j, kai startuojama būsenoje s_i. Tuo atveju, kai $i = j$, dydis $\tau_{i,i}$ vadinamas **vidutiniu grįžimo** į būseną s_i **laiku**. Svarbesnes dydžių $T_{i,j}$ ir $\tau_{i,j}$ savybes pateiksime teoremoje.

18 teorema *Tegul* $(X_0, X_1, \dots)$ *yra neredukuojama neperiodinė Markovo grandinė su būsenų aibe* $S = \{s_1, \dots, s_k\}$ *ir perėjimo matrica* $\mathbb{P}$. *Tuomet* $\forall s_i, s_j \in S$ *turime*

$$\mathbf{P}(T_{i,j} < \infty) = 1, \quad \mathbf{E}[T_{i,j}] < \infty.$$

Prieš nagrinėjant asimptotinį skirstinio $\mu^{(n)}$ elgesį, mes turime apibrėžti, ką reiškia tikimybinių skirstinių sekos $\nu^{(1)}, \nu^{(2)}, \dots$ konvergavimas į tikimybinį skirstinį ν. Taigi apibrėšime atstumą tarp tikimybinių skirstinių. Atstumai gali būti įvairiai apibrėžiami. Mes pasirinksime vieną, taip vadinamą **pilnosios variacijos atstumą** (total variation distance).

6 apibrėžimas *Jei* $\nu^{(1)} = \left(\nu_1^{(1)}, \dots, \nu_k^{(1)}\right)$ *ir* $\nu^{(2)} = \left(\nu_1^{(2)}, \dots, \nu_k^{(2)}\right)$ *yra aibės* $S = \{s_1, \dots, s_k\}$ *tikimybiniai skirstiniai, tai* ***pilnasis variacijos atstumas*** *tarp* $\nu^{(1)}$ *ir* $\nu^{(2)}$ *apibrėžiamas lygybe:*

$$d_{TV}\left(\nu^{(1)}, \nu^{(2)}\right) = \frac{1}{2} \sum_{i=1}^{k} \left|\nu_i^{(1)} - \nu_i^{(2)}\right|. \tag{69}$$

Jei $\nu^{(1)}, \nu^{(2)}, \dots$ *ir* ν *yra aibės* $S = \{s_1, \dots, s_k\}$ *tikimybiniai skirstiniai, tai sakysime, kad* $\nu^{(n)}$ ***konverguoja į*** ν ***pagal pilnąją variaciją****, kai* $n \to \infty$, *rašysime* $\nu^{(n)} \xrightarrow{TV} \nu$, *jei*

$$\lim_{n\to\infty} d_{TV}\left(\nu^{(n)}, \nu\right) = 0.$$

Konstanta $\frac{1}{2}$ (69) formulėje sunormuoja pilnosios variacijos atstumą, t.y. padaro, kad jis priimtų reišmes tarp 0 ir 1. Jei $d_{TV}(\nu^{(1)}, \nu^{(2)}) = 0$, tai $\nu^{(1)} = \nu^{(2)}$.

Dabar galime suformuluoti pagrindinį rezultatą, MG ribinę teoremą.

19 teorema *Tegul* $(X_0, X_1, \dots)$ *yra neperiodinė ir neredukuojama Markovo grandinė su būsenų aibe* $S = \{s_1, \dots, s_k\}$, *perėjimo matrica* $\mathbb{P}$ *ir bet kokiu*

pradiniu skirstiniu $\mu^{(0)}$. *Tuomet egzistuoja vienintelis MG stacionarusis skirstinys* π *ir*

$$\mu^{(n)} \xrightarrow{TV} \pi.$$

Be to, šis ribinis skirstinys

$$\pi = \left(\frac{1}{\tau_{1,1}}, \ldots, \frac{1}{\tau_{k,k}} \right).$$

Teorema sako, kad nepaisant koks buvo pradinis skirstinys, MG skirstinys n-uoju laiko momentu yra artimas stacionariam skirstiniui π, kai n yra pakankamai didelis. Tai dažnai pakeičiama pasakymu, kad MG priartėja prie **pusiausvyros** (equilibrium), kai $n \to \infty$.

Dabar pakalbėkime apie dar vieną MG rūšį, **apgręžiamas** (reversible) Markovo grandines. Jos vadinamos taip todėl, kad į jas galima žiūrėti kaip į MG ne tik kai laikas eina pirmyn, bet ir laikui einant priešinga kryptimi (atgal).

7 apibrėžimas *Tegul* $(X_0, X_1, \ldots)$ *yra Markovo grandinė su būsenų aibe* $S = \{s_1, \ldots, s_k\}$ *ir perėjimo matrica* $\mathbb{P}$. *Tikimybinis skirstinys* π *aibėje* S *yra vadinamas* ***apgręžiamu****, jei* $\forall i, j \in \{1, \ldots, k\}$ *teisinga lygybė:*

$$\pi_i P_{i,j} = \pi_j P_{j,i}. \tag{70}$$

Markovo grandinė vadinama ***apgręžiama****, jei yra tikimybinis skirstinys, kuris yra apgręžiamas šiai MG.*

Formulę (70) būtų galbūt galima paaiškinti taip: tikimybinės masės srautas, einantis iš būsenos s_i į būseną s_j (kairioji (70) lygybės pusė), yra lygus tikimybinės masės srautui, einančiam iš būsenos s_j į būseną s_i (dešinioji pusė). Ir tai matyt reiškia tam tikrą pusiausvyrą. Kad taip yra matyti ir iš teoremos apie apgręžiamas MG.

20 teorema *Tegul* $(X_0, X_1, \ldots)$ *yra Markovo grandinė su būsenų aibe* $S = \{s_1, \ldots, s_k\}$ *ir perėjimo matrica* $\mathbb{P}$. *Jei* π *yra šios MG apgręžiamas skirstinys, tai jis yra ir stacionarus šios MG skirstinys.*

Įrodymas Stacionaraus skirstinio apibrėžimo (i) savybė yra akivaizdi. Tad reikia patikrinti (ii) savybę. Turime, kad $\forall j \in \{1, \ldots, k\}$

$$\pi_j = \pi_j \sum_{i=1}^{k} P_{j,i} = \sum_{i=1}^{k} \pi_j P_{j,i} \overset{\text{apgręžiamumas}}{=} \sum_{i=1}^{k} \pi_i P_{i,j}.$$

Taigi $\pi = \pi\mathbb{P}$, ir π yra stacionarusis skirstinys. □

19 pavyzdys Atsitiktinis klaidžiojimas grafu. Šis pavyzdys yra atsitiktinio klaidžiojimo miesto gatvėmis, pateikto **??** pavyzdyje, apibendrinimas. **Grafas** $G = (V, E)$ susideda iš **viršūnių aibės** $V = \{v_1, \ldots, v_k\}$ ir **briaunų aibės** $E = \{e_1, \ldots, e_l\}$. Kiekviena briauna jungia dvi viršūnes. Briauna, jungianti viršūnes v_i ir v_j, žymima $\langle v_i, v_j \rangle$. Vieną viršūnių porą gali jungti tik viena briauna. Dvi viršūnės vadinamos **kaimynėmis**, jei jos yra sujungtos briauna.

Pavyzdžiui grafas **??** paveiksle turi viršūnių aibę $V = \{v_1, \ldots, v_8\}$ ir briaunų aibę

$$E = \{\langle v_1, v_3\rangle, \langle v_1, v_4\rangle, \langle v_2, v_3\rangle, \langle v_2, v_5\rangle, \langle v_2, v_6\rangle, \langle v_3, v_4\rangle, \\ \langle v_3, v_7\rangle, \langle v_3, v_8\rangle, \langle v_4, v_8\rangle, \langle v_5, v_6\rangle, \langle v_6, v_7\rangle, \langle v_7, v_8\rangle\}.$$

Atsitiktinis klaidžiojimas grafu $G = (V, E)$ yra Markovo grandinė su būsenų aibe $V = \{v_1, \ldots, v_k\}$ ir tokiu perėjimo mechanizmu: Jei klaidžiotojas n-uoju laiko momentu stovi viršūnėje v_i, tai $n+1$-uoju laiko momentu jis atsiduria vienoje iš v_i kaimynių, pasirinktų atsitiktinai su vienodomis tikimybėmis. Taigi, jei d_i yra viršūnės v_i kaimynių skaičius, tai perėjimo matricos elementai

$$P_{i,j} = \begin{cases} \frac{1}{d_i}, & \text{jei } v_i \text{ ir } v_j \text{ yra kaimynės,} \\ 0 & \text{kitais atvejais.} \end{cases}$$

Atsitiktinis klaidžiojimas grafu yra apgręžiama MG su apgręžiamu skirstiniu

$$\pi = \left(\frac{d_1}{d}, \ldots, \frac{d_k}{d}\right), \quad d = \sum_{i=1}^{k} d_i. \tag{71}$$

Lengva patikrinti, kad toks π parinkimas tenkina (70) formulę:

$$\pi_i P_{i,j} = \begin{cases} \frac{d_i}{d}\frac{1}{d_i} = \frac{1}{d} = \frac{d_j}{d}\frac{1}{d_j} = \pi_j P_{j,i}, & \text{jei } v_i \text{ ir } v_j \text{ yra kaimynės,} \\ 0 = \pi_j P_{j,i} & \text{kitais atvejais.} \end{cases}$$

Grafui, pavaizduotam xx paveiksle, (71) formulė tampa tokia:

$$\pi = \left(\frac{2}{24}, \frac{3}{24}, \frac{5}{24}, \frac{3}{24}, \frac{2}{24}, \frac{3}{24}, \frac{3}{24}, \frac{3}{24}\right).$$

Taigi, nusistovėjus pusiausvyros padėčiai, v_3 yra labiausiai tikėtina viršūnė, kurioje bus keleivis, tuo tarpu viršūnės v_1 ir v_5 yra mažiausiai tikėtinos.

20 pavyzdys Gimimo-mirties procesas. Tegul $(X_0, X_1, \dots)$ yra Markovo grandinė su būsenų aibe $S = \{s_1, \dots, s_k\}$ ir perėjimo matrica $\mathbb{P}$. Be to, tegul matrica $\mathbb{P}$ tenkina savybes:

(i) $P_{i,j} > 0$, jei $|i-j| = 1$, ir

(ii) $P_{i,j} = 0$, jei $|i-j| \geq 2$.

Tokia Markovo grandinė dažnai vadinama **gimimo-mirties procesu**. Jos perėjimo grafas yra pavaizduotas xx paveiksle (kai kurių ar net visų $P_{i,i}$-"kilpų" gali nebūti). Tokios rūšies MG visada yra apgręžiama. Sukonstruosime apgręžiamą skirstinį π. Paimkime π_1^* lygų bet kokiam teigiamam skaičiui a. Iš (70) sąlygos, kai $i = 1$, o $j = 2$, išplaukia, kad

$$\pi_2^* = \frac{aP_{1,2}}{P_{2,1}}\,.$$

Taikydami (70) dar kartą su $i = 2$ ir $j = 3$, gausime

$$\pi_3^* = \frac{\pi_2^* P_{2,3}}{P_{3,2}} = \frac{aP_{1,2}P_{2,3}}{P_{2,1}P_{3,2}}\,.$$

Tęsdami taip toliau, turėsime

$$\pi_i^* = \frac{a\prod_{l=1}^{i-1} P_{l,l+1}}{\prod_{l=1}^{i-1} P_{l+1,l}} \quad \forall i.$$

Taigi $\pi^* = (\pi_1^*, \dots, \pi_k^*)$ tenkina apgręžiamo skirstinio reikalavimus, išskyrus, gal būt, reikalavimą, kad jis būtų tikimybiniu skirstiniu, t.y., kad tikimybių suma būtų lygi 1. Tai lengva padaryti, padalijant visus π_i^* iš jų sumos. Galutinai gauname, kad

$$\pi = (\pi_1, \dots, \pi_k) = \left(\frac{\pi_1^*}{\sum_{i=1}^k \pi_i^*}, \dots, \frac{\pi_k^*}{\sum_{i=1}^k \pi_i^*}\right)$$

yra apgręžiamas skirstinys.

21 pavyzdys Neapgręžiama Markovo grandinė. Nagrinėkime šiek tiek modifikuotą keleivio klaidžiojimo gatvėmis (xx paveikslas) versiją. Tarkime, kad kiekvienoje sankryžoje keleivis pasirenka kryptį eiti pagal laikrodžio rodyklę su tikimybe $\frac{1}{4}$ ir prieš laikrodžio rodyklę su tikimybe $\frac{3}{4}$. Turėsime MG su perėjimo grafu, pavaizduotu xx paveiksle. Šiai MG stacionarus skirstinys bus $\pi = \left(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}\right)$ (patikrinkite!). Kadangi ši MG yra neredukuojama ir neperiodinė, tai iš 19 teoremos išplaukia, kad šis skirstinys bus vienintelis stacionarus šiai MG. Bet ši MG nėra apgręžiama, nes

$$\pi_1 P_{1,2} = \frac{1}{4}\cdot\frac{1}{4} = \frac{1}{16} < \frac{3}{16} = \frac{1}{4}\cdot\frac{3}{4} = \pi_2 P_{2,1}.$$

Intuityviai suprantama, kad ši MG nėra apgręžiama todėl, kad keleivis dažniau juda prieš laikrodžio rodyklę negu atgal. Vyrauja judėjimo prieš laikrodžio rodyklę tendencija. O jei mes judėtume laiku atgal, vyrautų judėjimo pagal laikrodžio rodyklę tendencija. Taigi yra skirtumas kaip eina laikas, pirmyn ar atgal.

Baigdami kalbą apie apgręžiamas MG paminėkime vieną paprastą ir nuostabų ekvivalentumą tarp apgręžiamų MG ir rezistorių sistemos. Naudojantis šiuo panašumu elektriniai argumentai (tokie kaip nuoseklaus ir lygiagretaus jungimo dėsniai) naudingi analizuojant MG, ir priešingai, tikimybiniai argumentai galimi studijuojant elektros sistemas. Gaila, bet tai neįeina į kurso programą.

6.6 Markovo grandinių Monte Karlo metodas

Šiame skyrelyje užsiimsime tokia problema. Aibėje $S = \{s_1, \dots, s_k\}$ duotas tikimybinis skirstinys π. Kaip imituoti atsitiktinį objektą, pasiskirsčiusį pagal skirstinį π? Pradėsime nuo pavyzdžio.

22 pavyzdys Nulių–vienetų (hard-core) modelis. Tegul $G = (V, E)$ yra grafas su viršūnių aibe $V = \{v_1, \dots, v_k\}$ ir briaunų aibe $E = \{e_1, \dots, e_l\}$. Kiekvienai viršūnei atsitiktinai priskirkime 0 arba 1 tokiu būdu, kad jokioms dviems kaimyninėms (t.y. tarp kurių yra briauna) viršūnėms nebūtų priskirti abu 1-tai. 0-ių ir 1-tų priskyrimai viršūnėms vadinami **konfigūracijomis** ir gali būti suprantami kaip aibės $\{0,1\}^V$ elementai. Konfigūracijos, kuriose jokie du 1-tai neužima kaimyninių viršūnių vadinamos **galimomis**. Atsitiktinę konfigūraciją parinksime iš visos galimų konfigūracijų aibės, imdami kiekvieną su vienoda tikimybe. Apibrėžėme tikimybę μ_G aibėje $\{0,1\}^V$:

$$\mu_G(\xi) = \begin{cases} \frac{1}{Z_G}, & \text{jei } \xi \text{ yra galima konfigūracija,} \\ 0 & \text{kitais atvejais;} \end{cases} \tag{72}$$

čia Z_G yra visų galimų konfigūracijų skaičius grafe G. Atsitiktinai parinktos konfigūracijos pavyzdį, kai grafas G yra kvadratinė 8×8 gardelė, žr. xx paveiksle.

Šis modelis (kai grafas yra trimatė gardelė) buvo naudojamas statistinėje fizikoje dujų savybių tyrimui, taip pat telekomunikacijos situacijų modeliavime.

Iškyla gana natūralus klausimas (modeliuose svarbus): Koks yra labiausiai tikėtinas 1-tų skaičius atsitiktinai parinktoje konfigūracijoje?

Tegul $n(\xi)$ yra vienetų skaičius konfigūracijoje ξ, o X – atsitiktinai parinkta konfigūracija. Tuomet tikimybinis vidurkis

$$\mathbf{E}[n(X)] = \sum_{\xi\in\{0,1\}^V} n(\xi)\mu_G(\xi) = \frac{1}{Z_G}\sum_{\xi\in\{0,1\}^V} n(\xi)\mathbf{I}_{\{\xi \text{ yra galima}\}}. \tag{73}$$

Suskaičiuoti šitą sumą galima nebent tada, kai grafas labai mažas, kadangi konfigūracijų skaičius (tuo pačiu ir dėmenų skaičius sumoje) auga eksponentiškai, palyginus su grafo matmenimis (pavyzdžiui, skirtingų konfigūracijų xx paveikslo grafe turėsime $2^{64} \approx 1.8\cdot 10^{19}$; fizikiniuose taikymuose paprastai sutinkami daug didesni grafai). Daliai viršūnių, žinoma, priskirti 0-iai, bet 1-tų skaičius vis tiek auga eksponentiškai. Taip pat pastebėkime, kad suskaičiuoti dydį Z_G nėra paprasta.

Kadangi tiksliai suskaičiuoti (73) išraišką neįmanoma, reikia kitokių idėjų. Viena iš jų būtų perėjimas prie imitacijos. Jei mes žinome kaip imituoti atsitiktinę konfigūraciją X su skirstiniu μ_G, tai mes galime tai padaryti daug kartų ir įvertinti $\mathbf{E}[n(X)]$ vidutiniu 1-tų skaičiumi, gautu imitacijoje. Iš didžiųjų skaičių dėsnio išplaukia, kad gautas įvertis konverguos į $\mathbf{E}[n(X)]$, kai imitacijų skaičius augs į begalybę. Todėl, naudojantis statistinėmis procedūromis, galima sukonstruoti pasikliautinuosius intervalus.

Turėdami galvoje šį pavyzdį, pagalvokime kaip imituoti atsitiktinį dydį X, pasiskirsčiusį būsenų aibėje S pagal duotą tikimybinį skirstinį π. Iš tikrųjų tai labai paprasta: sunumeruokime aibės S elementus, $s_1, \dots, s_k$; tada tegul

$$X = \psi(U);$$

čia U yra tolygus intervale $[0,1]$ atsitiktinis dydis, o funkcija $\psi : [0,1] \to S$ yra apibrėžta lygybe:

$$\psi(x) = \begin{cases} s_1, & \text{kai } x \in [0, \pi(s_1))\,, \\ s_2, & \text{kai } x \in [\pi(s_1), \pi(s_1)+\pi(s_2))\,, \\ \dots & \\ s_i, & \text{kai } x \in \left[\sum_{j=1}^{i-1}\pi(s_j), \sum_{j=1}^{i}\pi(s_j)\right)\,, \\ \dots & \\ s_k, & \text{kai } x \in \left[\sum_{j=1}^{k-1}\pi(s_j), 1\right]\,. \end{cases}$$

Praktiškai tai neįmanoma pritaikyti, nebent aibė S būtų maža. Nagrinėtame nulių–vienetų modelyje, kai gardelės dydis 8×8 ar didesnis, funkcijos ψ apskaičiavimas pasidaro praktiškai negalimas.

Tokioje situacijoje padeda **Markovo grandinių Monte Karlo (MGMK) metodas**. Metodas pradėtas naudoti fizikoje apie 1950 metus. Vėliau jį pradėta plačiai naudoti ir kitose srityse, ypatingai vaizdų analizėje nuo 1980 metų ir statistikos srityje, žinomoje **Bajeso statistikos** pavadinimu, nuo 1990 metų.

Idėja yra tokia. Tarkime, galime sukonstruoti neredukuojamą ir neperiodinę MG $(X_0, X_1, \dots)$, kurios (vienintelis) stacionarus skirstinys yra π. Šiai MG, paėmę bet kurį pradinį skirstinį, iš konvergavimo 19 teoremos gausime, kad MG skirstinys n-uoju laiko momentu π_n konverguoja į π, kai $n \to \infty$. Taigi, jei skaičiuosime MG pakankamai ilgai (dideliems n), gausime X_n gana artimus skirstiniui π. Iš tikrųjų, tai tik aproksimacija, bet ji gali būti pakankamai gera, imant didelius n.

Gali iškilti natūralus klausimas. Kodėl sukonstruoti MG, su norimomis savybėmis, lengviau negu sukonstruoti tiesiogiai atsitiktinį dydį, pasiskirsčiusį pagal skirstinį π? Kad būtų aišku, pateiksime pavyzdį.

23 pavyzdys Markovo grandinių Monte Karlo algoritmas nulių–vienetų modeliui. Nagrinėkime nulių–vienetų modelį, jau nagrinėtą 22 pavyzdyje. Taigi $G = (V, E)$ yra grafas (vienas iš jo konkrečių variantų pavaizduotas xx paveiksle) su viršūnių aibe $V = \{v_1, \dots, v_k\}$. Kad sukonstruotume Markovo grandinių Monte Karlo (MGMK) algoritmą šiam modeliui, mums reikia sukonstruoti MG, kurios būsenų aibė S yra galimos grafo G konfigūracijos, t.y.

$$S = \{\xi \in \{0,1\}^V : \xi \text{ yra galima}\}.$$

Taip pat reikia, kad MG būtų neredukuojama ir neperiodinė, o taip pat turėtų stacionarų skirstinį μ_G, apibrėžtą (72) lygybe.

MG su reikiamomis savybėmis gali būti gauta naudojant tokį perėjimo mechanizmą. Laiko momentu $n+1$ elgiamasi taip:

1. Atsitiktinai (ir tolygiai) parenkame viršūn $v \in V$.
2. Metame monetą.
3. Jei iškrinta herbas ir visi viršūnės v kaimynai yra nuliai n-uoju laiko momentu ($X_n(v_{\text{kaimynas}}) = 0$), tai imame $X_{n+1}(v) = 1$; atraip imame $X_{n+1}(v) = 0$.
4. Visas kitas viršūnes w, nesutampančias su v, paliekame nepakeistas, t.y. $X_{n+1}(w) = X_n(w),\ \forall w \neq v$.

Nesunku parodyti, kad ši Markovo grandinė yra neredukuojama ir neperiodinė. Reikia įsitikinti dar, kad skirstinys μ_G yra stacionarus. Iš xx teoremos išplaukia, jog užtenka parodyti, kad μ_G yra apgręžiamas. Tegul $P_{\xi,\xi'}$ yra perėjimo tikimybė iš būsenos ξ į būseną ξ' (kai

naudojamas aukščiau aprašytas perėjimo mechanizmas). Mums reikia patikrinti, kad

$$\mu_G(\xi)P_{\xi,\xi'} = \mu_G(\xi')P_{\xi',\xi} \tag{74}$$

kiekvienai galimai konfigūracijų porai ξ, ξ'. Tarkime, $d = d(\xi, \xi')$ yra viršūnių, kuriose ξ ir ξ' nesutampa, skaičius. Išskirkime tris atvejus $d = 0$, $d = 1$ ir $d \geq 2$. Kai $d = 0$, turime $\xi = \xi'$. Šiuo atveju (74) lygybė yra triviali. Atvejis $d \geq 2$ taip pat trivialus, kadangi grandinė per laiko vienetą niekada nepasikeičia daugiau kaip vienoje viršūnėje ir šiuo atveju $P_{\xi,\xi'} = P_{\xi',\xi} = 0$. Belieka atvejis $d = 1$, kai ξ ir ξ' nesutampa lygiai vienoje grafo viršūnėje. Todėl visi viršūnės v kaimynai turi būti nuliai, antraip viena iš dviejų konfigūracijų būtų negalima (kaimynais pasidarytų vienetai). Taigi turime

$$\mu_G(\xi)P_{\xi,\xi'} = \frac{1}{Z_G}\frac{1}{2k} = \mu_G(\xi')P_{\xi',\xi};$$

čia k yra grafo viršūnių skaičius. (74) lygybė teisinga, vadinasi, μ_G yra apgręžiamas (tuo pačiu ir stacionarus) skirstinys.

Naudodami xx skyrelio metodus dabar galime imituoti MG. Duomenų atnaujinimo funkcijos ϕ parinkimui patogu išskaidyti intervalą $[0, 1]$ į lygaus $\frac{1}{2k}$ ilgio intervaliukus, atitinkančius pasirinkimus

$$(v_1, \text{herbas}), (v_1, \text{skaičius}), (v_2, \text{herbas}), \ldots, (v_k, \text{skaičius}),$$

naudojamus aukščiau aprašytame perėjimo mechanizme. Jeigu mes ilgą laiką (n didelis) imituosime MG, startuodami su bet kuria galima konfigūracija (pavyzdžiui, su visais nuliais), tai gausime, kad X_n skirstinys yra artimas μ_G.

23 pavyzdžio MGMK algoritmas yra tipinis daugeliu aspektų. Pirma, nors mums reikia, kad MG skirstinys būtų stacionarus, randame net apgręžiamą skirstinį. Tokia savybe pasižymi dauguma MGMK algoritmų. Priežastis tame, kad lengviau patikrinti apgręžiamumo (70) sąlygą, negu stacionarumą. O jei skirstinys apgręžiamas, tai ir stacionarus (žinoma, jei MG neredukuojama ir neperiodinė).

Antra, 23 pavyzdžio MGMK algoritmas yra pavyzdys bendrai naudojamos specialios MGMK algoritmų klasės, žinomos kaip **Gibso modeliai**. Gibso modeliuose daroma taip. Būsenų aibėje, kuri turi formą S^V (čia S ir V yra baigtinės aibės) imituojamas tikimybinis skirstinys π. Kitais žodžiais tariant, turime baigtinę viršūnių aibę V su baigtine galimų reikšmių S aibe, o π yra koks nors atsitiktinis reikšmių iš S viršūnių aibėje V skirstinys (nulių–vienetų 22 pavyzdyje turėjome $S = \{0, 1\}$). Gibso modelis yra MG, kurioje $n + 1$-uoju laiko momentu atliekama:

1. Atsitiktinai (ir tolygiai) parenkama viršūnė $v \in V$.
2. X_{n+1} parenkamas priklausomai nuo sąlyginio π skirstinio, priklausančio nuo v reikšmės, laikant, kad visos kitos viršūnės priima X_n reikšmes.
3. Imama, kad $X_{n+1}(w) = X_n(w)\ \forall w \in V,\ w \neq v$.

Nesunku parodyti, kad tokia MG yra neperiodinė, ir kad π yra apgręžiamas (taigi ir stacionarus) skirstinys. Jeigu MG yra dar ir neredukuojama, tai ši MG yra korektiškas MGMK algoritmas, kurį galima naudoti atsitiktinio dydžio su skirstiniu π imitavimui. Pateiksime dar vieną pavyzdį.

24 pavyzdys MGMK algoritmas grafo atsitiktiniam q-spalvinimui. Tegul $G = (V, E)$ yra grafas, o $q \in \mathbb{N}, q \geq 2$. Grafo q-spalvinimas yra grafo viršūnių spalvinimas skirtingomis spalvomis iš aibės $\{1, \ldots, q\}$ taip, kad dvi kaimyninės viršūnės negali turėti tos pačios spalvos. Grafo G spalvinimas vadinamas atsitiktiniu, kai q-spalvinimas parenkamas tolygiai iš visos galimų q-spalvinimų aibės. Atitinkamą tikimybinį skirstinį[10] aibėje S^V žymėsime $\rho_{G,q}$.

Paimkime viršūnę $v \in V$. Simboliu ξ pažymėkime visų kitų nuspalvintų viršūnių rinkinį. Sąlyginis višūnės v spalvos skirstinys $\rho_{G,q}$ yra tolygus aibėje visų spalvų, kurių rinkinyje ξ neturi nei vienas v kaimynas. Atsitiktinio q-spalvinimo Gibso modelis yra MG su reikšmėmis iš aibės S^V, kai kiekvienu laiko momentu $n+1$, pasikeitimai vyksta taip:

1. Atsitiktinai (ir tolygiai) pasirenkama viršūnė $v \in V$.
2. $X_{n+1}(v)$ parenkamas pagal tolygų visų spalvų, kurių neturi nei vienas v kaimynas, skirstinį.
3. Visų kitų viršūnių spalvos paliekamos nepakeistos, t.y. $X_{n+1}(w) = X_n(w)\ \forall w \in V,\ w \neq v$.

Ši grandinė yra neperiodinė, o jos skirstinys $\rho_{G,q}$ yra stacionarus. Ar ši grandinė yra neredukuojama, priklauso nuo G ir q. Yra netriviali problema tai nustatyti bendruoju atveju[11]. Kai galime parodyti, kad grandinė yra neredukuojama, Gibso modelis tampa naudingu ir efektyviu MGMK algoritmu.

[10] Čia darome prielaidą, kad nagrinėjamam grafui G egzistuoja bent vienas q-spalvinimas. Taip būna nevisuomet. Pavyzdžiui, kai $q = 2$, o grafą G sudaro trys viršūnės, sujungtos trikampiu, tai jokio q-spalvinimo negalima surasti. Bendru atveju nustatyti, ar egzistuoja q-spalvinimas pasirinktiems G ir q, yra sudėtingas kombinatorikos uždavinys. Įžymi **keturių spalvų teorema** tvirtina, jei grafas G yra grafas plokštumoje (t.y. G gali būti pavaizduotas plokštumoje tokiu būdu, kad jokios dvi briaunos nekerta viena kitos), tai pakanka paimti $q = 4$.

[11] Palyginkite su ankstesne pastaba. Vienas dalykas, kurį nesunku parodyti, yra tai, kad bet kuriam grafui G, MG yra neredukuojama, kai q yra pakankamai didelis.

Benrai naudojamas Gibso modelio variantas yra toks. Vietoje to, kad rinktumės ir keistume viršūnes atsitiktinai, galime ciklišk ai pereiti per viršūnių aibę. Pavyzdžiui, kai $V = \{v_1, \ldots, v_k\}$, galime nuspręsti (pakeisti) viršūnes taip:

$$(75) \qquad \begin{cases} v_1 & \text{laiko momentais } 1, k+1, 2k+1, \ldots \\ v_2 & \text{laiko momentais } 2, k+2, 2k+2, \ldots \\ \ldots & \\ v_i & \text{laiko momentais } i, k+i, 2k+i, \ldots \\ \ldots & \\ v_k & \text{laiko momentais } k, 2k, 3k, \ldots \end{cases}$$

Tokia MG yra nehomogeninė, nes naudojama k skirtingų atnaujinimo (spalvų keitimo) taisyklių, naudojamų skirtingu laiku. Ji yra neperiodinė ir turi reikalaujamą skirstinį, be to, dar ir apgręžiamą. Taip pat ši MG yra neredukuojama $\Leftrightarrow$ pradinis Gibso modelis (pradinis viršūnių spalvų rinkinys) yra neredukuojamas. Tai neįrodinėsime. Šis Gibso modelio variantas yra vadinamas **sisteminiu Gibso modeliu** (systematic sweep Gibbs sampler).

Kita svarbi bendra procedūra, kaip gauti apgręžiamas Markovo grandines MGMK algoritmui, yra taip vadinamos **Metropolio grandinės**[12] konstravimas. Aprašykime būdą (ne patį bendriausią) kaip sukonstruoti Metropolio grandinę, kuri imituoja duotą tikimybinį skirstinį $\pi = (\pi_1, \ldots, \pi_k)$ aibėje $S = \{s_1, \ldots, s_k\}$. Pirma, reikia sukonstruoti kažkokį grafą G su viršūnių aibe S. Šio grafo briaunų aibė (kaimynų struktūra) gali būti bet kokia, išskyrus, kad

- grafas turi būti jungus, kad būtų užtikrintas gautos grandinės neredukuojamumas,
- kiekviena viršūnė neturi turėti per daug kaimynų, kadangi tokiu atveju grandinė tampa per sudėtinga skaičiavimams atlikti ir imituoti.

Kaip visada, sakysime, kad s_i ir s_j yra kaimynai, jei grafas turi briauną $\langle s_i, s_j \rangle$, jungiančią šias viršūnes. Taip pat būsenos s_i kaimynų skaičių žymėsime d_i. Metropolio grandinė, atitinkanti pasirinktą grafą, turės tokias perėjimo

[12] Dar bendresnė (ir plačiau naudojama) MG klasė, naudojama MGMK imitacijoje, yra **Metropolio-Hastingso grandinės.**

tikimybes:

$$P_{i,j} = \begin{cases} \frac{1}{d_i} \min\{\frac{\pi_j d_i}{\pi_i d_j}, 1\}, & \text{jei } s_i \text{ ir } s_j \text{ yra kaimynai,} \\ 0, & \text{jei } s_i \neq s_j \text{ nėra kaimynai,} \\ 1 - \sum\limits_{\substack{l \\ s_l \sim s_i}} \frac{1}{d_i} \min\{\frac{\pi_l d_i}{\pi_i d_l}, 1\}, & \text{jei } i = j. \end{cases} \tag{76}$$

Čia simbolis $\sim$ naudojamas kaimynams žymėti. Perėjimo matrica atitinka tokį perėjimo mechanizmą. Tarkime, $X_n = s_i$. Pirma, naudodami tolygų skirstinį, iš visos s_i kaimynų aibės parinkime kaimyną s_j (kiekvienas kaimynas parenkamas su tikimybe $\frac{1}{d_i}$). Tada paimkime

$$X_{n+1} = \begin{cases} s_j & \text{su tikimybe } \min\{\frac{\pi_j d_i}{\pi_i d_j}, 1\}, \\ s_i & \text{su tikimybe } 1 - \min\{\frac{\pi_j d_i}{\pi_i d_j}, 1\}. \end{cases}$$

Tokios MG skirstinys bus stacionarus, jei jis tenkins apgręžiamumo sąlygą

$$\pi_i P_{i,j} = \pi_j P_{j,i} \tag{77}$$

visiems i ir j. Įrodydami pastarąją lygybę elgsimės kaip ir 23 pavyzdyje. Pirma, pastebėkime, kad (77) lygybė yra triviali, kai $i = j$. Kai $i \neq j$, o s_i ir s_j nėra kaimynai, (77) lygybė teisinga, nes abi jos pusės yra lygios 0. Atvejį, kai s_i ir s_j yra kaimynai, skirsime į du, kai $\frac{\pi_j d_i}{\pi_i d_j} \geq 1$ ir < 1. Jei $\frac{\pi_j d_i}{\pi_i d_j} \geq 1$, tai

$$\begin{cases} \pi_i P_{i,j} = \pi_i \frac{1}{d_i}, \\ \pi_j P_{j,i} = \pi_j \frac{1}{d_j} \frac{\pi_i d_j}{\pi_j d_i} = \frac{\pi_i}{d_i}, \end{cases}$$

ir (77) lygybė teisinga. Panašiai, jei $\frac{\pi_j d_i}{\pi_i d_j} < 1$, tai

$$\begin{cases} \pi_i P_{i,j} = \pi_i \frac{1}{d_i} \frac{\pi_j d_i}{\pi_i d_j} = \frac{\pi_j}{d_j}, \\ \pi_j P_{j,i} = \pi_j \frac{1}{d_j}, \end{cases}$$

ir (77) lygybė vėl teisinga. Taigi π yra apgręžiamas (todėl ir stacionarus) Metropolio grandinės skirstinys, kuris gali būti panaudotas skirstinio π MGMK imitacijoje.

6.7 Vėsinimo imitacija

Šiame skyrelyje nagrinėsime tokį uždavinį. Tarkime, turime aibę $S = \{s_1, \ldots, s_k\}$ ir funkciją $f : S \to \mathbb{R}$. Tikslas yra surasti $s_i \in S$, kuriame funkcija įgyja minimumą: $f_{\min} = f(s_i)$ (kartais maksimumą: $f_{\max} = f(s_i)$).

Kai k nėra didelis, tai uždavinys yra trivialus. Reikia tik skaičiuoti iš eilės $f(s_i),\ i = 1, \ldots, k$, ir laikyti $f(s_i)$ atmintyje tol, kol nesurandame mažesnės reikšmės. Mes turėsime galvoje tokius uždavinius, kai k yra labai didelis, ir kai minėtas metodas praktiškai negali būti pritaikomas. Pateiksime du pavyzdžius.

25 pavyzdys Optimalus pakavimas. Tegul G yra grafas su viršūnių aibe V ir briaunų aibe E. Tarkime, reikia patalpinti objektus šio grafo viršūnėse tokiu būdu:

- kiekvienoje viršūnėje galima patalpinti ne daugiau kaip vieną objektą ir
- jokie du objektai negali užimti kaimyninių viršūnių,

ir norime tų objektų sutalpinti kiek įmanoma daugiausiai. Pavaizduokime objektus 1-ais, o tuščias viršūnes 0-iais. Tuomet uždavinys tampa tokiu. Reikia surasti (naudojant xx pavyzdžio terminologiją) galimą[13] konfigūraciją(as) $\xi \in \{0,1\}^V$, kuri(ios) turi maksimalų skaičių 1-ų[14]. Kaip jau minėta xx pavyzdyje, galimų konfigūracijų skaičius auga labai greitai (eksponentiškai) lyginant su grafo dydžiu. Taigi aukščiau minėtas paprastas reikšmių $f(\xi)$ (šiuo atveju $f(\xi)$ yra 1-ų skaičius konfigūracijoje ξ) skaičiavimas visiems ξ praktiškai neįmanomas net vidutinio dydžio grafams.

26 pavyzdys Keliaujančio pirklio uždavinys. Tarkime, kad yra m miestų, o simetrinėje matricoje $D_{m\times m}$ surašyti atstumai tarp šių miestų. Įsivaizduokime pirklį, gyvenantį viename iš šių miestų, kuriam reikia aplankyti kitus $m-1$ miestą ir grįžti namo. Kokia tvarka jis turėtų aplankyti visus miestus, kad nukeliautų trumpiausią atstumą? Ekvivalentus uždavinys yra toks. Reikia surasti aibės $(1, \ldots, m)$ kėlinį $\xi = (\xi_1, \ldots, \xi_m)$, kuriame funkcija

$$f(\xi) = \sum_{i=1}^{m-1} D_{\xi_i,\xi_{i+1}} + D_{\xi_m,\xi_1} \tag{78}$$

[13] Prisiminkim,kad konfigūracija $\xi \in \{0,1\}^V$ yra vadinama galima, jei jokios dvi kaimyninės viršūnės nėra 1-ai.

[14] Imant 8×8 gardelę (xx brėž.) optimalaus pakavimo uždavinys yra trivialus. Įsivaizduokime viršūnes kaip šachmatų lentos langelius ir patalpinkime 1-us į 32 tamsius langelius. Nesunku suprasti, kad toks atvejis yra optimalus. Bet kitokioms grafų struktūroms gali būti ne taip lengva surasti optimalų pakavimą

įgyja minimumą. Paprastas reikšmių $f(\xi)$ skaičiavimas gali būti pritaikytas tik tuo atveju, kai m nedidelis, nes skirtingų kėlinių ξ skaičius yra $m!$. O $m!$ auga net greičiau negu eksponentiškai, lyginant su m.

Tokios rūšies optimizacijos uždavinių sprendimui naudota daug metodų. Šiame skyrelyje išsiaiškinsime vieną iš tokių metodų: **vėsinimo imitaciją**.

Vėsinimo imitacijos idėja yra tokia.Tarkime, MG su būsenų aibe S turi vienintelį stacionarų skirstinį, tokį, kad būsenos $s \in S$ tikimybė pati didžiausia, o reikšmė $f(s)$ maža. Jeigu mes imituosime tokią grandinę gana ilgą laiką, tai tikėtina, kad baigsime būsenoje s. Tarkime, kad po to vėl imituojame kitą MG. Ji turi vienintelį stacionarų skirstinį su dar didesne tikimybe patekti į būseną s, kurioje $f(s)$ įgyja minimumą. Taigi po kažkokio laiko su dar didesne tikimybe galime atsidurti funkciją minimizuojančioje būsenoje s. Tada imituojame dar kitą MG su dar didesnėmis galimybėmis atsidurti minimizuojančioje būsenoje s ir t.t. Jeigu tokią schemą sukonstruosime rūpestingai, tai tikimybė, kad n-uoju laiko momentu būsime f-minimizuojančioje būsenoje artės į 1, kai $n \to \infty$.

Jei pirmoji MG turi perėjimo matricą $\mathbb{P}'$ ir imituojama laiką N_1, antroji MG turi perėjimo matricą $\mathbb{P}''$ ir imituojama laiką N_2 ir t.t., tai visas algoritmas yra nehomogeninė MG su perėjimo matricomis

$$\mathbb{P}^{(n)} = \begin{cases} \mathbb{P}', & \text{kai } n = 1, \dots, N_1, \\ \mathbb{P}'', & \text{kai } n = N_1 + 1, \dots, N_1 + N_2, \\ \dots \end{cases}$$

Yra bendras kelias pasirinkti tikimybinį skirstinį, kuris aibėje S turi didžiausias tikimybes tų būsenų s, kuriose f įgyja mažas reikšmes, būtent, taip vadinamą **Bolcmano skirstinį**, apibrėžtą žemiau. MG su bolcmano skirstiniu, kaip vieninteliu stacionariu, gali būti sukonstruota naudojant MGMK idėjas, išdėstytas 6.6 skyrelyje.

8 apibrėžimas ***Bolcmano skirstiniu*** $\pi_{f,T}$ *baigtineje aibėje* S *su* ***energijos funkcija*** $f : S \to \mathbb{R}$ *ir* ***temperatūros parametru*** $T > 0$ *vadinamas tikimybinis skirstinys aibėje* S*, kuris kiekvienam elementui* $s \in S$ *priskiria tikimybę*

$$\pi_{f,T}(s) = \frac{1}{Z_{f,T}} \exp\left(\frac{-f(s)}{T}\right). \tag{79}$$

Čia

$$Z_{f,T} = \sum_{s \in S} \exp\left(\frac{-f(s)}{T}\right) \tag{80}$$

yra normuojanti konstanta, užtikrinanti lygybę:

$$\sum_{s\in S}\pi_{f,T}(s)=1$$

Jeigu tikslas yra rasti ne funkcijos f minimumą, bet maksimumą, tai tada vietoje Bolcmano skirstinio naudojamas **modifikuotas Bolcmano skirstinys**, kuris gaunamas (79) ir (80) formulėse eksponentes $\frac{-f(s)}{T}$ pakeičiant į $\frac{f(s)}{T}$.

Dabar pateiksime rezultatą, kuris tvirtina, kad Bolcmano skirstinys su mažomis temperatūros parametro T reikšmėmis turi minėtas savybes, t.y. tikimybės didžiausios tų elementų s, kurie minimizuoja $f(s)$.

21 teorema *Tegul S yra baigtinė aibė, o $f : S \to \mathbb{R}$ bet kokia funkcija. Tegul $\alpha(T)$, $T > 0$, yra tikimybė, kad atsitiktinis elementas Y, parinktas pagal Bolcmano skirstinį $\pi_{f,T}$ aibėje S, minimizuoja f:*

$$f(Y)=\min_{s\in S} f(s).$$

Tuomet

$$\lim_{T\to 0}\alpha(T)=1.$$

Įrodymas. Įrodysime tik tuo atveju, kai f turi vienintelį minimumą. Tegul $S=\{s_1,\ldots,s_k\}$, $\min_{s'\in S} f(s')=f(s)=a$, $\min_{s'\in S\setminus\{s\}} f(s')=b$. Tuomet

$$\begin{aligned}\pi_{f,T}(s)&=\frac{1}{Z_{f,T}}\exp\left(\frac{-a}{T}\right)=\frac{\exp\left(\frac{-a}{T}\right)}{\sum_{s'\in S}\exp\left(\frac{-f(s')}{T}\right)}\\&=\frac{\exp\left(\frac{-a}{T}\right)}{\exp\left(\frac{-a}{T}\right)+\sum_{s'\in S\setminus\{s\}}\exp\left(\frac{-f(s')}{T}\right)}\\&\geq\frac{\exp\left(\frac{-a}{T}\right)}{\exp\left(\frac{-a}{T}\right)+(k-1)\exp\left(\frac{-b}{T}\right)}\\&=\frac{1}{1+(k-1)\exp\left(\frac{a-b}{T}\right)}.\end{aligned}$$

Kadangi $a<b$, tai

$$\lim_{T\to 0}\pi_{f,T}(s)=1.$$

□

Dabar išdėstysime algoritmą, kaip gali būti surastas funkcijos f minimumas aibėje S. Naudosime Bolcmano skirstinį $\pi_{f,T}$ aibėje S. Tai atliksime konstruodami Metropolio grandinę (žr. 6.6 skyrelį). Pirma, fiksuojama temperatūrų seka $T_1 > T_2 > \ldots$, kad $T_i \overset{i\to\infty}{\to} 0$, (čia paaiškėja, kodėl naudojamas

vėsinimo terminas) ir natūraliųjų skaičių seka $N_1, N_2, \ldots$ Startuojama bet kurioje aibės S būsenoje. Imituojama MG su temperatūros parametru T_1 laiką N_1, po to su temperatūros parametru T_2 laiką N_2 ir t.t.

Dydžių $T_1, T_2, \ldots$ ir $N_1, N_2, \ldots$ parinkimas vadinamas vėsinimo tvarkaraščiu ir yra labai svarbus. Yra teoremos, tvirtinančios, kad jei temperatūrų seka artėja į 0 pakankamai lėtai (tai ekvivalentu, jei laikų seka $N_1, N_2, \ldots$ auga pakankamai greitai), tai tikimybė, kad n-uoju laiko momentu surasime f minimumą artėja į 1, kai $n \to \infty$[15]. Vėsinimo tvarkaraščiai, kurie garantuoja konvergavimą pagal šias teoremas, daugeliu atvejų yra nerealūs, nes reikia astronominio laiko, kol temperatūra pasidaro pakankamai maža ir, kai jau easme beveik garantuoti, kad surastas f minimumas. Bet vėsinimo procedūros praktiniuose pritaikymuose paprastai yra greitesnės. Ir vis tik išlieka pavojus, jei vėdinsime per greitai, galime atsidurti ne absoliutaus minimumo, bet lokaliojo minimumo taške. Dėl to, sprendžiant konkrečius uždavinius, reikia nemažai eksperimentuoti. Taigi tokie uždaviniai kartais tampa ne matematinio, bet inžinerinio pobūdžio.

27 pavyzdys Vėsinimo imitacija keliaujančio pirklio uždavinyje. Nagrinėkime keliaujančio pirklio uždavinį (26 pavyzdys). Ieškokime aibės $(1, \ldots, m)$ kėlinio $\xi = (\xi_1, \ldots, \xi_m)$, kuris minimizuoja atstumą $f(\xi)$, apibrėžtą (78) lygybe. Pirma, naudodami Bolcmano skirstinį $\pi_{f,T}$, sukonstruokime Metropolio grandinę aibės $(1, \ldots, m)$ kėlinių aibėje. Taigi reikia apibrėžti perėjimo mechanizmą tarp kėlinių, t.y. apibrėžti "kaimynus" tarp kėlinių. Vienas iš paprastesnių būdų galėtų būti toks. Du kėlinius ξ ir ξ' vadinkime kaimynais, jei $\exists\, i, j \in (1, \ldots, m)$, $i < j$, tokie, kad kėlinys ξ' gaunamas iš kėlinio ξ apgręžiant segmentą $(\xi_i, \ldots, \xi_j)$:

$$\begin{aligned}\xi' &= (\xi'_1, \ldots, \xi'_m) \\ &= (\xi_1, \xi_2 \ldots, \xi_{i-1}, \xi_j, \xi_{j-1}, \ldots, \xi_{i+1}, \xi_i, \xi_{j+1}, \xi_{j+2}, \ldots, \xi_m).\end{aligned}$$

Kelionės po miestus grafe tai atitiktų dviejų briaunų pakeitimą kitomis dviejomis briaunomis (tokiame keitime dalyvauja tik 4 viršūnės), kad gautųsi kitokia kelionė. Žr. xx paveikslą. Tokią Metropolio grandinę atitinkanti perėjimo matrica gaunama Bolcmano skirstinio tikimybes

[15] Viena iš tokių teoremų (Gemano): Jei temperatūra n-uoju laiko momentu $T^{(n)}$ artėja į 0 pakankamai lėtai, t.y.

$$T^{(n)} \geq \frac{k\left(\max_{s\in S} f(s) - \min_{s\in S} f(S)\right)}{\log n}$$

visiems pakankamai dideliems n, tai tikimybė, kad n-uoju laiko momentu surasime f minimumą artėja į 1, kai $n \to \infty$.

įstačius į (76) formulę:

$$P_{\xi,\xi'} = \begin{cases} \frac{2}{m(m-1)} \min\left\{\exp\left(\frac{f(\xi)-f(\xi')}{T}\right), 1\right\}, & \text{jei } \xi \sim \xi', \\ 0, & \text{jei } \xi \neq \xi' \text{ ir } \xi \nsim \xi', \\ 1 - \sum\limits_{\substack{\xi'' \\ \xi'' \sim \xi}} \frac{2}{m(m-1)} \min\left\{\exp\left(\frac{f(\xi)-f(\xi'')}{T}\right), 1\right\}, & \text{jei } \xi = \xi'. \end{cases}$$

Čia ženklas $\sim$ reiškia kaimynus. Taigi turime tokį perėjimo mechanizmą. Pirma, pasirenkame $i, j \in \{1, \ldots, m\}$, $i < j$. Renkamės tolygiai iš visos galimų pasirinkimų aibės. Tada su tikimybe $\min\left\{\exp\left(\frac{f(\xi)-f(\xi')}{T}\right), 1\right\}$ kėlinį ξ pakeičiame kėliniu ξ', o su tikimybe $1-\min\left\{\exp\left(\frac{f(\xi)-f(\xi')}{T}\right), 1\right\}$ paliekame tą patį kėlinį ξ kitam laikui (padidintam laiko vienetu). Tokia grandinė turi Bolcmano skirstinį $\pi_{f,T}$. Jis yra apgręžiamas (pagal bendrą Metropolio grandinių teoriją). Taip pat galima parodyti, kad grandinė yra ir neredukuojama.

Belieka nuspręsti kokį vėsinimo tvarkaraštį pasirinkti, t.y. pasirinkti dvi sekas $T_1, T_2, \ldots$ ir $N_1, N_2, \ldots$ Galima tai atlikti ekperimentuojant.

Pastebėkime vieną aplinkybę. Pateiktame pavyzdyje, įstatant Bolcmano skirstinio tikimybes į (76) formulę, normuojančios konstantos $Z_{f,T}$ susiprastina. Tai gana gerai, nes priešingu atveju reikėtų skaičiuoti $Z_{f,T}$. (Suskaičiuoti jų faktiškai neįmanoma.) Tai atsitinka visuomet, kai naudojami Bolcmano skirstiniai Metropolio grandinei. Taigi tokia schema yra patogi skaičiavimams atlikti.

Pateiksime dar vieną pavyzdį, parodantį, kad per greitas vėsinimas gali neduoti gero rezultato.

28 pavyzdys Per greito vėsinimo tvarkaraščio naudojimo rizika. Tegul $S = \{s_1, s_2, s_3, s_4\}$, o $f : S \to \mathbb{R}$ yra tokia:

$$\begin{cases} f(s_1) = 1, \\ f(s_2) = 2, \\ f(s_3) = 0, \\ f(s_4) = 2. \end{cases}$$

Tarkime, turime rasti f minimumą naudodami vėsinimo imitaciją.[16] Bolcmano skirstinio aibėje S (su temperatūra T) Metropolio grandinės

[16] Žinoma yra kvaila naudoti vėsinimo imitaciją tokiam uždaviniui kaip šis. Juk galime nustatyti, kad minimumas $f(s_3) = 0$, tiesiog palygindami visas f reikšmes. Šis pavyzdys yra parinktas tik tam, kad pailiustruotume kas gali atsitikti naudojant vėsinimo imitacijos algoritmą žymiai sudėtingesniuose ir įdomesniuose uždaviniuose.

sukonstravimui turime apibrėžti grafo struktūrą aibėje S. Tarkime, turime grafą pavaizduotą xx brėžinyje. Įstatę $\pi_{f,T}$ iš (79) formulės į (76) formulę gausime tokią perėjimo matricą:

$$\begin{pmatrix} 1-\mathrm{e}^{-1/T} & \frac{1}{2}\mathrm{e}^{-1/T} & 0 & \frac{1}{2}\mathrm{e}^{-1/T} \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2}\mathrm{e}^{-2/T} & 1-\mathrm{e}^{-2/T} & \frac{1}{2}\mathrm{e}^{-2/T} \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \end{pmatrix}.$$

Toliau tarkime, kad generuojame MG $X_0, X_1, \dots$ aibeje S pagal kažkokį vėsinimo tvarkaraštį, startuodami su $X_0 = s_1$. Tegul $T^{(n)}$ yra temperatūra n-uoju laiko momentu. Tegul A yra įvykis, kad grandinė pasiliks būsenoje s_1 visam laikui (taigi minimali reikšmė $f(s_3)$ niekada nebus surasta). Turėsime

$$\begin{aligned} P(A) &= P(X_1 = s_1, X_2 = s_1, \dots) \\ &= \lim_{n\to\infty} P(X_1 = s_1, X_2 = s_1, \dots, X_n = s_1) \\ &= \lim_{n\to\infty} P(X_1 = s_1 | X_0 = s_1) P(X_2 = s_1 | X_1 = s_1) \\ &\dots P(X_n = s_1 | X_{n-1} = s_1) \\ &= \lim_{n\to\infty} \prod_{i=1}^{n} \left(1 - \mathrm{e}^{-1/T^{(i)}}\right) = \prod_{i=1}^{\infty} \left(1 - \mathrm{e}^{-1/T^{(i)}}\right). \end{aligned}$$

Iš čia aišku, kad $P(A) = 0 \Leftrightarrow \sum_{i=1}^{\infty} \mathrm{e}^{-1/T^{(i)}} = \infty$. Taigi, jei $T^{(n)}$ konverguoja į 0 per greitai, t.y. taip, kad eilutė $\sum_{i=1}^{\infty} \mathrm{e}^{-1/T^{(i)}} < \infty$, tai $P(A) > 0$. Vadinasi su teigiama tikimybe grandinė gali pasilikti būsenoje s_1 visam laikui. Tai atsitinka, pavyzdžiui, jei pasirinksime $T^{(n)} = \frac{1}{n}$. Šiuo atveju imituojamas vėsinimo algoritmas gali nesurasti tikrojo f minimumo $f(s_3)$. Tokia klaida gauta dėl dviejų priežasčių:

- vėsinimo tvarkaraštis per greitas, ir
- būsena s_1 yra lokaliojo minimumo taškas (ta prasme, kad f įgyja didesnes reikšmes kaimyninėse s_1 viršūnėse).

Skyriaus pradžioje pateiktame 12 pavyzdyje apie grafo dalinimą pusiau vėsinimo imitacija galėtų būti sėkmingai taikoma.

Paskutiniaisiais metais keletas mokslininkų atsisakė vėsinimo tvarkaraščio idėjos. Vietoje to sukonstravo Metropolio grandinę fiksuotai temperatūrai, kuri buvo parinkta po kruopščių matematinių paskaičiavimų. Pavyzdžiui, sprendžiant grafo dalinimo pusiau uždavinį, buvo sudarytas algoritmas, kuris $\forall \epsilon > 0$, jei k yra didelis, o T yra eilės $n^{5/6+\epsilon}$, per laiką $Ck^{2+\epsilon}$ su didele tikimybe (artėjančia į 1, kai $k \to \infty$) suranda optimalų grafo padalinimą pusiau.

6.8 Propo-Vilsono algoritmas

Nors MGMK metodas naudojamas modeliavime, ir ten yra labai naudingas, paminėkime du tokių taikymų trūkumus.

(A) Teorinis MGMK metodo pagrindas yra xx teorema, kuri garantuoja, kad neredukuojamos ir neperiodinės MG skirstinys $\mu^{(n)}$ laiko momentu n, nesvarbu kokia pradinė būsena bebūtų, konverguoja į stacionarų skirstinį π, kai $n \to \infty$. Bet iš to neišplaukia, kad $\mu^{(n)}$ kada nors bus lygus π, bet tiktai, kad $\mu^{(n)}$ labai priartėja prie π. Iš tikrųjų, didžiojoje daugumoje pavyzdžių mes gauname, kad $\mu^{(n)} \neq \pi$ visiems n. Vadinasi kokį didelį n bepaimtume vis tiek bus kažkoks skirtumas tarp gauto skirstinio $\mu^{(n)}$ ir ieškomo skirstinio π.

(B) Norint padaryti paklaidą, apie kurią kalbama (A) pastaboje, maža, reikia žinoti, kokį n turėtume paimti, kad skirtumas (matuojamas pilnosios variacijos atstumu $d_{TV}(\mu^{(n)}, \pi)$) būtų mažesnis negu duotas $\epsilon > 0$. Daugelyje situacijų yra labai sudėtinga nustatyti kokius n reikia paimti, kad gautume minėtą skirtumą pakankamai mažą. Dažnai nustatytos ribos neturi jokios praktinės naudos.[17]

1996 m. šias problemas išsprendė Propas(Propp) ir Vilsonas(Wilson). Jie sugalvojo MGMK metodo patobulinimą, pateikdami algoritmą, kuris išsprendė abi (A) ir (B) problemas:

(A*) jų algoritmas duoda tiksliai skirstinį π ir

(B*) automatiškai suranda laiką, kada reikia sustoti, nereikalaudamas MG konvergavimo greičio skaičiavimų.

Šis algoritmas vadinamas **Propo-Vilsono algoritmu**. Apie jį ir kalbėsime šiame skyrelyje. Nuo kitų MGMK algoritmų jis skiriasi tuo, kad generuojama ne viena MG, bet kelios MG kopijos[18] su skirtingomis pradinėmis reikšmėmis. Kitas svarbus (greitai suprasime kodėl) bruožas yra tai, kad grandinė pradedama ne 0-iu laiku, bet nuo kažkurio laiko praeityje ir imituojama iki 0-io laiko.

[17] Bendrai, kruopščiai tiriant xx teoremos įrodymą, galime gauti viršutinę n ribą (priklausančią nuo ϵ ir nuo grandinės). Tačiau dažnai tos ribos yra astronomonio dydžio, pavyzdžiui, tokios kaip "$d_{TV}(\mu^{(n)}, \pi) < 0.01$, jei $n \geq 10^{100}$". Tai visiškai nenaudinga, nes MG 10^{100} žingsnių imitacija nesibaigtų mūsų gyvenimo laiku. Tokiose situacijose mes galime tik tikėtis, kad konvergavimas yra daug greitesnis (kad, gal būt, $n = 10^5$ užteks), bet įrodyti tai dažnai būna pernelyg sudėtinga.

[18] Dėl priežasčių, kurias išsiaiškinsime, Propas ir Vilsonas pavadino savo algoritmą "dubliavimas praeityje" (coupling from the past).

Turint omenyje (A*) savybę, Propo-Vilsono algoritmas dažnai vadinamas **tiksliu** ar **tobulu** imitavimo algoritmu.

Pereikime prie detalaus algoritmo aprašymo. Tarkime, reikia parinkti elementą iš aibės $S = \{s_1, \dots, s_k\}$ pagal duotą skirstinį π. Kaip ir kituose MGMK algoritmuose konstruojame apgręžiamą, neredukuojamą ir neperiodinę MG su būsenų aibe S ir stacionariu skirstiniu π. Tegul $\mathbb{P}$ yra perėjimo matrica, o $\phi : S \times [0,1] \to S$ – kažkokia duomenų atnaujinimo funkcija. Tegul $N_1, N_2, \dots$ yra didėjanti natūraliųjų skaičių seka. Pasirinkimas, kurį pagrįsime būtų $(N_1, N_2, \dots) = (1, 2, 4, 8, \dots)$. Neigiami skaičiai $-N_1, -N_2, \dots$ bus naudojami kaip MG starto laiko momentai. Pagaliau tarkime, kad $U_0, U_{-1}, U_{-2}, \dots$ yra nepriklausomų ir tolygiai pasiskirsčiusių intervale $[0,1]$ atsitiktinių dydžių seka. Tuomet algoritmas veikia taip.

1. Tegul $m = 1$.

2. Kiekvienam $s \in \{s_1, \dots, s_k\}$ imituojame MG, laiko momentu $-N_m$ startuodami būsenoje s, ir tęsdami imitaciją iki 0-inio laiko. Naudojame duomenų atnaujinimo funkciją ϕ ir atsitiktinius skaičius $U_{-N_m+1}, U_{-N_m+2}, \dots, U_{-1}, U_0$. Pastarieji atsitiktiniai skaičiai yra tie patys visoms k grandinių.

3. Jei visos k grandinių 2-ame žingsnyje 0-iniu laiku baigiasi toje pačioje būsenoje s', tai išėjimas yra s', ir procesą pabaigiame. Jei ne, pereiname prie 4-ojo žingsnio.

4. Padidiname m vienetu ir pereiname prie 2-ojo žingsnio.

Yra svarbu, kad m-uoju laiku, kai ateiname į 2-ąjį žingsnį, naudojame atsitiktinius skaičius $U_{-N_m+1}, U_{-N_m+2}, \dots, U_{-1}, U_0$, kurių dalis $U_{-N_{m-1}+1}, U_{-N_{m-1}+2}, \dots, U_{-1}, U_0$ jau naudota ankstesniame žingsnyje. Tai yra būtina, kad algoritmas dirbtų korektiškai (žr. xx pavyzdį). Bet tai daro uždavinį gremėzdišku, nes reikia išsaugoti, gal būt, ilgą atsitiktinių dydžių seką.[19]

XX brėžinyje nagrinėjamas pavyzdys, kai $(N_1, N_2, \dots) = (1, 2, 4, 8, \dots)$, o $S = \{s_1, s_2, s_3\}$. Kadangi $N_1 = 1$, startuojame su grandine prasidedančia laiku -1 ir besibaigiančia laiku 0. Tarkime,

$$\begin{cases} \phi(s_1, U_0) = s_1, \\ \phi(s_2, U_0) = s_2, \\ \phi(s_3, U_0) = s_1. \end{cases}$$

[19] Kaip apeiti šią problemą diskutuosime vėliau.

Vadinasi, 0-iniu laiku turėsime dvi skirtingas būsenas (s_1 arba s_2), priklausančias nuo būsenos -1-uoju laiku. Todėl turime startuoti dar kartą laiku $-N_2 = -2$. Turime

$$\begin{cases} \phi(\phi(s_1, U_{-1}), U_0) = \phi(s_2, U_0) = s_2, \\ \phi(\phi(s_2, U_{-1}), U_0) = \phi(s_3, U_0) = s_1, \\ \phi(\phi(s_3, U_{-1}), U_0) = \phi(s_2, U_0) = s_2. \end{cases}$$

Ir vėl 0-iniu laiku gauname dvi skirtingas būsenas. Taigi turime dar kartą startuoti laiku $-N_3 = -4$. Gauname

$$\begin{cases} \phi(\phi(\phi(\phi(s_1, U_{-3}), U_{-2}), U_{-1}), U_0) = \cdots = s_2, \\ \phi(\phi(\phi(\phi(s_2, U_{-3}), U_{-2}), U_{-1}), U_0) = \cdots = s_2, \\ \phi(\phi(\phi(\phi(s_3, U_{-3}), U_{-2}), U_{-1}), U_0) = \cdots = s_2. \end{cases}$$

Taigi startuodami laiku -4, nepriklausomai nuo būsenos tuo laiku, 0-iniu laiku visada turime būseną s_2. Algoritmas baigiasi, fiksuodamas išėjime s_2 būseną. Pastebėkime, kad, jei būtume toliau imitavę grandinę, pradėdami laiku $-8, -16$ ir t.t., būtume turėję tą pačią būseną s_2 (išėjime) 0-iniu laiku. Taigi galime galvoti, kad būseną s_2 būtume turėję, jei būtume pradėję imituoti grandinę laiku $-\infty$, ir kad grandinė įgijo pusiausvyrą. Tai Propo-Vilsono algoritmo intuicija.

Pastebėkime, kad, bendrai paėmus, Propo-Vilsono algoritmas gali niekada nesibaigti (neturime kolkas jokių garantijų, kad Propo-Vilsono algoritmas kada nors pasibaigs). Iš tikrųjų, jis gali niekad nepasibaigti, jei duomenų atnaujinimo funkcija parinkta blogai, žr. xx uždavinį. Iš kitos pusės dažnai galima parodyti, kad algoritmas turi pabaigą su tikimybe 1.[20] Šiuo atveju išėjimas (atsitiktinis dydis, gautas 0-iniu laiku) turi skirstinį π. Tai suformuluosime kaip teoremą.

22 teorema *Tegul $\mathbb{P}$ yra neredukuojamos ir neperiodinės Markovo grandinės, su būsenų aibe $S = \{s_1, \dots, s_k\}$ ir stacionariu skirstiniu $\pi = (\pi_1, \dots, \pi_k)$, perėjimo matrica. Tegul ϕ yra duomenų atnaujinimo funkcija. Tarkime, imitacijai naudojamas Propo-Vilsono algoritmas su $(N_1, N_2, \dots) = (1, 2, 4, 8, \dots)$. Taip pat tarkime, algoritmas baigiasi per baigtinį laiką su tikimybe 1. Tegul Y yra išėjimas 0-iniu laiku. Tuomet $\forall i \in \{1, \dots, k\}$ turime*

$$P(Y = s_i) = \pi_i.$$

[20] Kad tai parodytume galima naudotis taip vadinamu tikimybiniu nulio-vieneto dėsniu: Propo-Vilsono algoritmas baigiasi per baigtinį laiką tik su tikimybe 0 arba 1. Taigi, jei pasisektų parodyti, kad Propo-Vilsono algoritmas baigiasi per baigtinį laiką su teigiama tikimybe, tai galėtume daryti išvadą, kad jis baigiasi per baigtinį laiką su tikimybe 1.

Dabar dar negalime pasakyti, kad toks Propo-Vilsono algoritmas praktiškai naudingas, nebent aibė S labai nedidelė.[21] Kokia didelė aibė S gali būti, kad galėtume generuoti algoritmą pradėdami visose būsenose s_i? Tai užimtų per daug kompiuterio laiko net ne itin dideliems k.

Atsakymas glūdi kitur. Galima surasti įvairių idėjų, panaudojant MG monotoniškumo savybę ir neimituoti visų variantų. Apie tai diskutuosime vėlesniuose skyreliuose.

Skyrelį baigsime pateikdami pora pavyzdžių, kuriuose bandoma supaprastinti Propo-Vilsono algoritmą. Deja, šie pavyzdžiai nekorektiški.

29 pavyzdys Dubliavimas ateityje.(Coupling to the future.)

Šis pavyzdys atsakys į klausimą, kodėl Propo-Vilsono algoritme startuojama vis toliau praeityje, o ne 0-iniu laiku ir einama vis toliau į ateitį.

Nagrinėkime tokį pavyzdį. Tegul $X_0, X_1, \ldots$ yra MG su būsenų aibe $S = \{s_1, s_2\}$ ir perėjimo matrica

$$\mathbb{P} = \begin{pmatrix} 0.5 & 0.5 \\ 1 & 0 \end{pmatrix}.$$

Perėjimo grafas pavaizduotas xx brėžinyje. Grandinė yra apgręžiama su stacionariu skirstiniu

$$\pi = (\pi_1, \pi_2) = \left(\frac{2}{3}, \frac{1}{3}\right). \tag{81}$$

Tarkime, generuojame dvi grandinės kopijas, startuodami 0-iniu laiku, vieną kopiją pradėdami būsenoje s_1, kitą s_2. Jos susijungs (įgis tą pačią reikšmę) pirmą kartą kažkokiu atsitiktiniu laiku N. Nagrinėkime situaciją, kai laikas yra $N-1$. Šiuo laiku viena iš grandinių yra būsenoje s_2. Bet iš perėjimo matricos išplaukia, kad ši grandinė su tikimybe 1 bus būsenoje s_1 po vieno žingsnio, t.y. laiku N. Vadinasi grandinės su tikimybe 1 pirmą kartą sutaps būsenoje s_1. Taigi šio modifikuoto Propo-Vilsono algoritmo išėjimas yra s_1 su tikimybe 1. Tai nesutampa su (81) stacionariu skirstiniu. Gauname, kad algoritmas nėra korektiškas.

30 pavyzdys Čia pateiktas kitas Propo-Vilsono algoritmo supaprastinimo pavyzdys.

[21] Jeigu S nedidelė, Propo-Vilsono algoritmas irgi nereikalingas, nes tuomet galima modeliuoti reikalingą dydį pačiu elementariausiu būdu.

> Kai grandinė startuoja iš naujo laiku $-N_{m+1}$, anksčiau naudoti atsitiktiniai dydžiai $U_{-N_m+1}, U_{-N_m+2}, \dots, U_0$ panaudojami vėl. Kodėl paprasčiausiai juos negeneruoti iš naujo?

Vėl naudokime tą pačią MG kaip ir ankstesniame 29 pavyzdyje. Tarkime, kad šiai grandinei taikome Propo-Vilsono algoritmą su $(N_1, N_2, \dots) = (1, 2, 4, 8, \dots)$ ir duomenų atnaujinimo funkcija

$$\phi(s_1, x) = \begin{cases} s_1, & \text{kai } x \in [0, 0.5), \\ s_2, & \text{kai } x \in [0.5, 1], \end{cases}$$

$$\phi(s_2, x) = s_1 \quad \forall x \in [0, 1].$$

Grandinę atnaujinkime kiekvieną kartą, generuodami vis naujus U_i. Tegul Y yra šio modifikuoto algoritmo išėjimas. Apibrėžkime atsitiktinį dydį M kaip didžiausią iš m, kuriems algoritmas nusprendžia startuoti laiku $-N_m$ (t.y. laikais $-N_M$ grandinės startuoja paskutinį kartą). Tiesioginiai skaičiavimai duoda

$$\begin{aligned} P(Y = s_1) &= \sum_{m=1}^{\infty} P(M = m, Y = s_1) \\ &\geq P(M = 1, Y = s_1) + P(M = 2, Y = s_1) \\ &= P(M = 1)P(Y = s_1 | M = 1) \\ &\quad + P(M = 2)P(Y = s_1 | M = 2) \\ &= \frac{1}{2} \cdot 1 + \frac{3}{8} \cdot \frac{2}{3} = \frac{3}{4} > \frac{2}{3}. \end{aligned}$$

Išėjimo Y skirstinys nesutampa su (81) skirstiniu π. Taigi pateiktas modifikuotas algoritmas nėra korektiškas.

7 MONTE KARLO METODO TAIKYMAI

7.1 Radioaktyviojo skilimo generavimas

Tarkime, turime N_t radioaktyvių atomų laiko momentu t. Suskilusių atomų skaičius dN per laiko intervalą $(t, t+dt)$ yra tiesiogiai proporcingas radioaktyvių atomų skaičiui N_t ir laiko intervalo ilgiui dt

$$dN = -\lambda N_t dt. \tag{82}$$

Minuso ženklas reiškia, kad radioaktyvių atomų skaičius mažėja (dalis suskyla, jų nelieka). Konstanta λ priklauso tik nuo radioaktyviosios medžiagos. Jos dimensija yra $\frac{1}{s}$. λ parodo kuri dalis atomų suskyla vidutiniškai per 1 sek. (Tikimybė atomui suskilti per 1 sek.)

Jei laiko momentu $t = 0$ radioaktyvių atomų yra N_0, o laiko momentu t yra N_t, tai iš (82) gausime

$$\frac{dN}{N_t} = -\lambda dt, \quad \Big| \quad \int_0^t$$

$$\ln \frac{N_t}{N_0} = -\lambda t,$$

$$N_t = N_0 e^{-\lambda t} \tag{83}$$

– radioaktyvių atomų skaičiaus mažėjimo dėsnis.

Iš tikrųjų atomas gali gyventi nesuskilęs labai ilgai. Jo gyvavimo laikas $t \in (0, +\infty)$. Bet vidutinis gyvavimo laikas T baigtinis. Tegul $T_{1/2}$ – radioaktyvaus atomo gyvavimo pusamžis $T_{1/2} = \dfrac{T}{2}$. Tuomet po tiek laiko suskils pusė atomų:

$$\frac{N_0}{2} = N_0 e^{-\lambda T_{1/2}},$$

$$T_{1/2} = \frac{\ln 2}{\lambda} \approx \frac{0.7}{\lambda}.$$

Skaičius atomų, suskylančių laiko intervale $(t, t + dt)$, yra (iš (82) ir (83))

$$\lambda N_t dt = \lambda N_0 e^{-\lambda t} dt.$$

Šių suskilusių atomų gyvavimo laikas $\approx t$ (tarp t ir $t + dt$). Susumuojam visus gyvavimo laikus:

$$T_\Sigma = \int_0^{+\infty} t\lambda N_0 e^{-\lambda t} dt = \frac{N_0}{\lambda}.$$

$\frac{T_\Sigma}{N_0}$ yra vidutinis atomo gyvavimo laikas:

$$\tau = \frac{T_\Sigma}{N_0} = \frac{1}{\lambda} \approx \frac{T_{1/2}}{0.7} \approx 1.4 T_{1/2}.$$

Tikimybė, kad radioaktyvus atomas skils intervale $(t, t+dt)$ yra

$$\frac{\text{suskilusių intervale } (t, t+dt) \text{ atomų skaičius}}{\text{visų atomų skaičius}} =$$

$$= \frac{\lambda N_0 e^{-\lambda t} dt}{N_0} = \lambda e^{-\lambda t} dt =$$

$$= \frac{1}{\tau} e^{-\frac{t}{\tau}} dt = f(t) dt.$$

Čia $f(t)$ yra eksponentinis tankis.

GRAFIKAS

Atsitiktiniam laikui t, momentams kuriais įvyksta radioaktyviųjų atomų skilimai, generuoti galime naudoti procedūrą:

$$U = \frac{1}{\tau} \int e^{-\frac{s}{\tau}} ds = \left(-e^{-\frac{t}{\tau}} \right) \Big|_0^t = -e^{-\frac{t}{\tau}} + 1,$$

$$e^{-\frac{t}{\tau}} = 1 - U.$$

Galima pakeisti

$$e^{-\frac{t}{\tau}} = U.$$

Čia U yra tolygiai pasiskirstęs intervale $(0, 1)$ atsitiktinis dydis. Toks yra ir $1 - U$. Taigi atsitiktiniais laiko momentais

$$t = -\tau \ln U$$

įvyksta radioaktyvios medžiagos skilimai. Tai laiko momentų modelis.

7.2 Monte-Karlo metodo bendroji taikymo schema

Monte-Karlo (MK) metodo taikymuose labai svarbi yra centrinė ribinė teorema (CRT). Todėl ją ir suformuluokime. Nepateiksime jos bendriausiu atveju, o tik nesudėtingą, paprastą variantą.

Įveskime pažymėjimus. Tegul

$$\xi_1, \xi_2, \dots, \xi_N$$

yra N nepriklausomų vienodai pasiskirsčiusių atsitiktinių dydžių. Tarkime,

$$E\xi_1 = E\xi_2 = \cdots = E\xi_N = m,$$

ir

$$D\xi_1 = D\xi_2 = \cdots = D\xi_N = b^2$$

yra šių dydžių vidurkiai ir dispersijos. Tegul

$$\rho_N = \xi_1 + \xi_2 + \cdots + \xi_N. \tag{84}$$

Tuomet

$$E\rho_N = Nm := a, \quad D\rho_N = Nb^2 := \sigma^2. \tag{85}$$

23 teorema (CRT) *Tegul $\zeta \in N(a, \sigma^2)$ (normaliai pasiskirstęs atsitiktinis dydis su vidurkiu a ir dispersija σ^2). Tuomet*

$$P(\alpha < \rho_N < \beta) \approx \int_\alpha^\beta p_\zeta(x)\, dx,$$

kai N pakankamai dideli. Čia

$$p_\zeta(x) = \frac{1}{\sqrt{2\pi}\sigma}\, e^{-\frac{(x-a)^2}{2\sigma^2}}$$

yra atsitiktinio dydžio ζ tankio funkcija.

Prireiks dar vienos teoremos.

24 teorema (3σ taisyklė) *Tegul $\zeta \in N(a, \sigma^2)$. Tuomet*

$$\int_{a-3\sigma}^{a+3\sigma} p_\zeta(x)\, dx \approx 0.977.$$

Dabar galime pereiti prie bendrosios MK metodo taikymo schemos.

Sakykime, reikia apskaičiuoti nežinomą dudį m. Reikėtų sugalvoti tokį atsitiktinį dydį ξ, kad $E\xi = m$. Tegul $D\xi = b^2$.

Toliau nagrinėkime N nepriklausomų atsitiktinių dydžių $\xi_1, \xi_2, \ldots, \xi_N$, kurie visi pasiskirstę kaip ir dydis ξ. Laikykimės pažymėjimų (105) ir (106). Iš CRT ir 3σ taisyklės turėsime, kad

$$P\{a - 3\sigma < \rho_n < a + 3\sigma\} \approx 0.977,$$

kai N pakankamai dideli. Iš pastarosios formulės gauname, kad

$$P\left\{m - \frac{3b}{\sqrt{N}} < \frac{\rho_N}{N} < m + \frac{3b}{\sqrt{N}}\right\} \approx 0.977$$

arba

$$P\left\{\left|\frac{1}{N}\sum_{j=1}^{N}\xi_j - m\right| < \frac{3b}{\sqrt{N}}\right\} \approx 0.977.$$

Taigi teisinga tokia teorema.

25 teorema

(86) $$P\left\{\left|\frac{1}{N}\sum_{j=1}^{N}\xi_j - m\right| < \frac{3b}{\sqrt{N}}\right\} \approx 0.977.$$

Paskutinioji (107) formulė gana svarbi MK metode. Ji duoda algoritmą, kaip paskaičiuoti m ir kartu paklaidos įvertinimą. Iš (107) formulės, žinodamo kokią paklaidą galime daryti, surandame N, o po to ir m.

7.3 Apibrėžtinio integralo skaičiavimas

Sakykime, reikia suskaičiuoti integralą

$$I = \int_a^b g(x)\,dx.$$

Pasirinkime kokį nors atsitiktinį dydį ξ, turintį tankį $p_\xi(x)$ intervale (a, b):

$$\int_a^b p_x i(x)\,dx = 1.$$

Kartu su atsitiktiniu dydžiu ξ nagrinėkime kitą atsitiktinį dydį

$$\eta = \frac{g(\xi)}{p_\xi(\xi)}\,.$$

Tuomet atsitiktinio dydžio η vidurkis

$$E\eta = \int_a^b \left(\frac{g(x)}{p_\xi(x)}\right) p_\xi(x)\,dx = I.$$

Dabar nagrinėkime N vienodai pasiskirsčiusių atsitiktinių dydžių $\eta_1, \eta_2, \ldots, \eta_N$ ir jų sumai taikykime 3 teoremą. Gausime

$$P\left\{\left|\frac{1}{N}\sum_{j=1}^{N}\eta_j - I\right| < 3\sqrt{\frac{D\eta}{N}}\right\} \approx 0.977. \tag{87}$$

Iš čia išplaukia, kad jei mes paimame N reikšmių $\xi_1, \xi_2, \ldots, \xi_N$, tai

$$\frac{1}{N}\sum_{j=1}^{N}\frac{g(\xi_j)}{p_\xi(\xi_j)} \approx I. \tag{88}$$

Taip pat gauname, kad (88) apytikslės lygybės paklaida su labai didele tikimybe neviršija

$$3\sqrt{D\eta/N}. \tag{89}$$

7.3.1 Apibrėžtinio integralo skaičiavimo pavyzdys

31 pavyzdys *Naudodami MK metodą suskaičiuokime integralą*

$$I = \int_0^{\pi/2} \sin x\,dx.$$

1. Tegul

$$p_\xi(x) = \frac{2}{\pi}$$

yra tolygiai intervale $(0, \pi/2)$ pasiskirsčiusio atsitiktinio dydžio ξ tankis,

$$\int_0^{\pi/2} \frac{2}{\pi}\, dx = 1.$$

Jeigu atsitiktinis dydis U yra tolygiai pasiskirstęs intervale $(0, 1)$, tai reikėtų imti

$$\xi = \frac{\pi}{2}\, U.$$

Taigi, gausime

$$I \approx \frac{\pi}{2N} \sum_{j=1}^{N} \sin \xi_j. \tag{90}$$

◀

2. Pasirinkime kitą atsitiktinį dydį ξ, pasiskirsčiusį intervale $(0, \pi/2)$ pagal tiesinį tankį

$$p_\xi = \frac{8x}{\pi^2}, \qquad \int_0^{\pi/2} \frac{8x}{\pi^2}\, dx = 1.$$

Dydžio ξ generavimui naudokimės universaliu metodu, panaudokime dydžio ξ skirstinį F_ξ. Tarkime, U yra tolygusis atsitiktinis dydis intervale $(0, 1)$. Tuomet

$$F_\xi(\xi) = U \quad \text{arba} \quad \int_0^{\xi} \frac{8x}{\pi^2}\, dx = U.$$

Išsprendę pastarąją lygybę ξ atžvilgiu, gausime

$$\xi = \frac{\pi}{2}\sqrt{U}.$$

Taigi

$$I \approx \frac{\pi^2}{8N} \sum_{j=1}^{N} \frac{\sin \xi_j}{\xi_j}\,. \tag{91}$$

◀

Kuris iš parinktų tankių geresnis? Antrasis, nes jis geriau aproksimuoja (jo grafikas artimesnis) pointegrinę funkciją (žr. xx brėž.).

MK metodas nenaudojamas tokių paprastų integralų skaičiavimui. Jis nenaudojamas net ir sudėtingiems vienalypiams integralams skaičiuoti, nes yra geresni ir tikslesni metodai – kvadratūrinės formulės. Bet kai reikia suskaičiuoti daugialypius integralus, kvadratūrinės formulės pasidaro labai sudėtingos. Lieka faktiškai vienintelis kelias integravimas MK metodu.

7.4 Tiesinių lygčių sistemos sprendimas Monte Karlo metodu

7.4.1 Iteracijų metodas. Įvadas

Nagrinėkime tiesinių lygčių sistemą:

$$
\begin{aligned}
x_1 &= a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n + b_1,\\
x_2 &= a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n + b_2,\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
x_n &= a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n + b_n,.
\end{aligned}
\tag{92}
$$

kurioje koeficientai $a_{11}, a_{12}, \ldots, a_{nn}, b_1, b_2, \ldots, b_n$ yra duoti, o kintamuosius $x_1, x_2, \ldots, x_n$ reikia surasti. Naudodami matricas (92) lygčių sistemą galime užrašyti trumpiau:

$$\mathbf{x} = \mathbf{A}\mathbf{x} + \mathbf{b}\,; \tag{93}$$

čia

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \ldots \\ b_n \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \ldots \\ x_n \end{pmatrix}.$$

Tarkime, kad $\mathbf{x}^{(0)}$ yra pradinė nežinomųjų vektoriaus reikšmė, o vektorius $\mathbf{x}^{(k)}$ paskaičiuojamas naudojantis lygtimis:

$$\mathbf{x}^{(k)} = \mathbf{A}\mathbf{x}^{(k-1)} + \mathbf{b}, \quad k = 1, 2, \ldots \tag{94}$$

Tarkime taip pat, kad viena iš dviejų nelygybių

$$\sum_{i,j=1}^{n} a_{ij}^2 < 1 \quad \text{arba} \quad \max_{1\le i\le n} \sum_{j=1}^{n} |a_{ij}| < 1 \tag{95}$$

yra patenkinta. Yra žinoma, kad tuomet seka $\mathbf{x}^{(k)}$, gauta naudojant (94) iteracinę procedūrą, konverguoja į (93) lygčių sistemos sprendinį $\mathbf{x}^*$, t.y.

$$\mathbf{x}^* = \mathbf{A}\mathbf{x}^* + \mathbf{b}, \quad \lim_{k\to\infty} \mathbf{x}^{(k)} = \mathbf{x}^*.$$

Parinkę pradinį vektorių $\mathbf{x}^{(0)} = \mathbf{b}$, apytikslį (93) lygčių sistemos sprendinį $\mathbf{x}^{(k)}$ galime užrašyti taip:

$$\mathbf{x}^{(k)} = (\mathbf{E} + \mathbf{A} + \mathbf{A}^2 + \cdots + \mathbf{A}^k)\mathbf{b}, \tag{96}$$

čia $\mathbf{E}$ yra vienetinė matrica. Iš (96) formulės gauname, kad apytikslio sprendinio $\mathbf{x}^{(k)}$ i-oji komponentė

$$(97)\quad x_i^{(k)} = b_i + \sum_{j_1=1}^{n} a_{ij_1}b_{j_1} + \sum_{j_1=1}^{n}\sum_{j_2=1}^{n} a_{ij_1}a_{j_1j_2}b_{j_2} + \cdots + \\ + \sum_{j_1=1}^{n}\sum_{j_2=1}^{n}\cdots\sum_{j_k=1}^{n} a_{ij_1}a_{j_1j_2}\ldots a_{j_{k-1}j_k}b_{j_k}.$$

Net vidutiniams (neypač dideliems) k ir n tokių sumų skaičiavimas gana ilgas procesas. Sutrumpinti skaičiavimams naudojamas Monte Karlo metodas. Kituose skyreliuose sukonstruosime atsitiktinį dydį, kurio vidurkis yra lygus (97) išraiškai ir išsiaiškinsime kaip panaudoti Monte Karlo metodą.

7.4.2 Monte Karlo metodo taikymas sumoms skaičiuoti

Trumpai pakartosime Monte Karlo metodo esmę. Tarkime, mums reikia bent jau apytiksliai suskaičiuoti sumą

$$S = \sum_{i=1}^{m} c_i.$$

Tam tikslui parenkame tikimybes $p_1, p_2, \ldots, p_m$: $p_i > 0, i = 1, 2, \ldots, m$, $\sum_{i=1}^{m} p_i = 1$. Toliau apibrėžkime diskretųjį dydį ξ, įgyjantį reikšmes $d_i = c_i/p_i$ su tikimybėmis p_i, $i = 1, 2, \ldots, m$, t.y.

$$P(\xi = d_i) = p_i, \quad i = 1, 2, \ldots, m.$$

26 teorema *Atsitiktinio dydžio ξ vidurkis yra lygus S.*

Įrodymas. Įrodymas išplaukia iš atsitiktinio dydžio vidurkio apibrėžimo

$$E\xi = \sum_{i=1}^{m} p_i d_i = \sum_{i=1}^{m} c_i = S.$$

◀

Taigi sumai S paskaičiuoti gali būti naudojamas Monte Karlo metodas. Turime generuoti N atsitiktinio dydžio ξ realizacijų $\xi^{(i)}$ ir paimti jų aritmetinį vidurkį

$$S_N = \frac{1}{N}\sum_{i=1}^{N} \xi^{(i)}.$$

Tai ir bus sumos S apytikslė reikšmė. Žinoma, tokią procedūrą verta naudoti, jei reikiamam tikslumui pasiekti užtenka N realizacijų, kai $N \ll m$,

(pavyzdžiui, $m/N > 10$). Atsitiktinio dydžio ξ dispersija

$$D^2\xi = \sum_{i=1}^{m} p_i d_i^2 - E^2(\xi) = \sum_{i=1}^{n} \frac{c_i^2}{p_i} - E^2(\xi). \tag{98}$$

Reikėtų taip parinkti tikimybes, kad dispersija būtų kuo mažesnė. Iš (98) dispersijos išraiškos aišku, kad reikia minimizuoti dydžius c_i^2/p_i, $i = 1, 2, \ldots, m$. Iš tikimybių teorijos žinome, kad geresni apytiksliai rezultatai gaunami tada, kai atsitiktinio dydžio ξ dispersija yra maža. Dažniausiai dydžiai c_i^2 nėra žinomi. Bet kokia informacija apie dydžių c_i elgesį gali būti naudinga parenkant p_i taip, kad dispersiją $D^2(\xi)$ gautume kiek galima mažesnę. Pavyzdžiui, jei seka yra $c_1, c_2, \ldots, c_m$ monotoniškai mažėjanti, tai ir tikimybės $p_1, p_2, \ldots, p_m$ turėtų elgtis taip pat.

7.4.3 XXX

Ankstesniame skyrelyje aprašytas idėjas pritaikykime (97) sumų įvertinimui. Paimkime matricą

$$\mathbf{P} = \mathbf{P}_{n\times n} = (p_{ij}),$$

kurios elementai p_{ij} tenkina sąlygas:

$$\begin{aligned} &\sum_{j=1}^{n} p_{ij} = 1,\ i = 1, 2, \ldots, n, \\ &\qquad\qquad \text{– kiekvieno matricos stulpelio elementų suma lygi 1;} \\ &p_{ij} > 0, \text{ jei } a_{ij} \neq 0, \text{ ir } \ p_{ij} = 0, \text{ jei } a_{ij} = 0. \end{aligned} \tag{99}$$

Iliustruotame pavyzdyje mums užteko vieno atsitiktinio dydžio. Dabar reikės n atsitiktinių dydžių ir jų skirstinių. Skaičius d_{ij} parinkime taip:

$$d_{ij} = \begin{cases} \dfrac{a_{ij}}{p_{ij}}, & \text{jei } a_{ij} \neq 0, \\ 0, & \text{jei } a_{ij} = 0. \end{cases} \tag{100}$$

Panašiai kaip ir ankstesniame skyrelyje pateiktame pavyzdyje (97) sumos nariai generuojami su tam tikromis (specifinėmis) tikimybėmis. Dydis $d_{ij_1} d_{j_1 j_2} \ldots d_{j_{k-1} j_k} b_{j_k}$ parenkamas su tikimybe $p = p_{ij_1} p_{j_1 j_2} \ldots p_{j_{k-1} j_k}$. Galima tai suformuluoti ir kitaip – indeksų kalba. T.y. pakanka parinkti indeksų seką $I = (i, j_1, j_2, \ldots, j_k)$ su tikimybe p, nes šie indeksai vienareikšmiškai apibrėžia dydį $d_{ij_1} d_{j_1 j_2} \ldots d_{j_{k-1} j_k} b_{j_k}$. Taigi mes galime apsiriboti klausimu: Kokia procedūra tiks generuoti (parinkti) I indeksų sekai su tikimybe p?

Jei indeksas j_1 parenkamas su tikimybe p_{ij_1}, indeksas j_2 parenkamas su tikimybe $p_{j_1 j_2}$, $\ldots$, ir indeksas j_k parenkamas su tikimybe $p_{j_{k-1} j_k}$, tai

I indeksų aibė yra parenkama su tikimybe $p = p_{ij_1}p_{j_1j_2}\dots p_{j_{k-1}j_k}$. Norėdami pademonstruoti indeksų generavimo procesą, panaudosime Markovo grandinių terminologiją. Diskretieji atsitiktiniai dydžiai $\xi_1, \xi_2, \dots, \xi_k$ gali įgyti reikšmes. Jei $\xi_i = j$, tai sakoma, kad atsitiktinis taškas i-uoju momentu (i-ajame žingsnyje) yra j-ojoje būsenoje. Atsitiktiniai dydžiai $\xi_1, \xi_2, \dots, \xi_k$ aprašo atsitiktinį k kelią, aplankant kažkiek būsenų iš n galimų. Įvykis, kad atsitiktinis taškas pateks iš r būsenos į s būseną, tikimybė žymima p_{rs} ir vadinama perėjimo tikimybe. Šios tikimybės nepriklauso nuo žingsnių, kurie buvo padaryti iki patekimo į r būseną. T.y. kiekvienam i turime, kad

$$p_{rs} = P\{\xi_{i+1} = s | \xi_i = r\}, \quad r, s = 1, 2, \dots, n. \tag{101}$$

Indeksų aibė I gali būti suprantama kaip atsitiktinis kelias, susidedantis iš k žingsnių, kai atsitiktinis taškas startuoja iš būsenos i ir po k žingsnių atsiduria būsenoje j_k aplankydamas būsenas $j_1, j_2, \dots, j_{k-1}$. Iš kitos pusės, paėmus pradinę reikšmę $\xi_0 = i$, atsitiktinių dydžių $\xi_1, \xi_2, \dots, \xi_k$ realizacija (gauta pagal (101) lygybes, naudojant perėjimo tikimybių matricą $\mathbf{P}$) duoda indeksų seką I, ir ši indeksų seka I generuota tiksliai su tikimybe $p = p_{ij_1}p_{j_1j_2}\dots p_{j_{k-1}j_k}$.

Dabar reziumuokime tai apie ką kalbėjome šiuose skyreliuose. Lygties (93) apytikslio sprendinio komponentė (vieno nežinomojo reikšmė) gali būti paskaičiuota naudojant Monte Karlo metodą. Kad galėtume tai padaryti, turime

- paimti matricą P, tenkinančią (99) sąlygas,
- nustatyti indekso i reikšmę, t.y. to kintamojo, kurį skaičiuosime,
- nustatyti dydį k, kuris apibrėžia aproksimacijos tikslumą,
- nustatyti generuojamų atsitiktinių egzempliorių skaičių N, nuo kurio priklauso Monte Karlo metodo padarytos klaidos dydis (tikslumas),
- naudodami (100) formulę, paskaičiuoti dydžius d_{ij}.

Taigi turėsime parinkti N atsitiktinių kelių. r-asis atsitiktinis kelias apibrėžia indeksų seką $I = (i, j_1, j_2, \dots, j_k)$. Aišku, ji priklauso nuo r, t.y. $I = I_r$, bet 65 indeks1, kad nebūtų labai griozdiški užrašai, mes praleidžiame. Toliau gautai indeksų sekai I suskaičiuojame sumą

$$S^{(r)} = b_i + d_{ij_1}b_{j_1} + d_{ij_1}d_{j_1j_2}b_{j_2} + \dots + d_{ij_1}d_{j_1j_2}\dots d_{j_{k-1}j_k}b_{j_k}. \tag{102}$$

Šios sumos generavimo tikimybė yra lygi $p = p_{ij_1}p_{j_1j_2}\dots p_{j_{k-1}j_k}$.

Kai jau turime sugeneravę N atsitiktinių kelių, dydžio $x_i^{(k)}$ aproksimaciją gauname paimdami sumų $S^{(r)}$ aritmetinį vidurkį

$$x_i^{(k)} \approx \frac{1}{N}\sum_{r=1}^{N} S^{(r)}.$$

27 teorema *Suma*

$$\frac{1}{N}\sum_{r=1}^{N} S^{(r)} \tag{103}$$

yra nepasislinkęs dydžio $x_i^{(k)}$*, apibrėžto* (97) *formule, įvertinimas.*

Įrodymas. Sumos (103) (atsitiktinio dydžio) vidurkis yra lygus

$$\begin{aligned} E\left(\frac{1}{N}\sum_{r=1}^{N} S^{(r)}\right) &= \frac{1}{N}\sum_{r=1}^{N} E\left(S^{(r)}\right) = E\left(S^{(r)}\right) \\ &= \sum_{j_1=1}^{n}\sum_{j_2=1}^{n}\cdots\sum_{j_k=1}^{n} S^{(r)}\cdot p_{ij_1}p_{j_1j_2}\dots p_{j_{k-1}j_k} \\ &= \sum_{j_1=1}^{n}\sum_{j_2=1}^{n}\cdots\sum_{j_k=1}^{n} b_i\cdot p_{ij_1}p_{j_1j_2}\dots p_{j_{k-1}j_k} \\ &+ \sum_{j_1=1}^{n}\sum_{j_2=1}^{n}\cdots\sum_{j_k=1}^{n} d_{ij_1}b_1\cdot p_{ij_1}p_{j_1j_2}\dots p_{j_{k-1}j_k} \\ &+ \sum_{j_1=1}^{n}\sum_{j_2=1}^{n}\cdots\sum_{j_k=1}^{n} d_{ij_1}d_{j_1j_2}b_{j_2}\cdot p_{ij_1}p_{j_1j_2}\dots p_{j_{k-1}j_k} \\ &+ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ &+ \sum_{j_1=1}^{n}\sum_{j_2=1}^{n}\cdots\sum_{j_k=1}^{n} d_{ij_1}d_{j_1j_2}\dots d_{j_{k-1}j_k}b_{j_k}\cdot p_{ij_1}p_{j_1j_2}\dots p_{j_{k-1}j_k}. \end{aligned} \tag{104}$$

Įstatę į (104) dydžius d_{ij} iš (100) formulės, turėsime

$$\begin{aligned} E\left(\frac{1}{N}\sum_{r=1}^{N} S^{(r)}\right) = b_i + \sum_{j_1=1}^{n} a_{ij_1}b_{j_1} + \sum_{j_1=1}^{n}\sum_{j_2=1}^{n} a_{ij_1}a_{j_1j_2}b_{j_2} + \dots + \\ + \sum_{j_1=1}^{n}\sum_{j_2=1}^{n}\cdots\sum_{j_k=1}^{n} a_{ij_1}a_{j_1j_2}\dots a_{j_{k-1}j_k}b_{j_k}. \end{aligned}$$

Gauta suma sutampa su (97) formulės dešiniąja puse. ◄

Iš tikrųjų Monte Karlo metodas tiesinių lygčių sistemoms spręsti naudojamas tik išimtiniais atvejais. Pavyzdžiui, tada, kai lygčių sistemą sudaro gana daug lygčių ir reikia surasti tik keletą nežinomųjų. Kitais atvejais geriau naudoti skaitinių metodų deterministinę techniką.

8 ATSITIKTINUMO SAMPRATA

8.1 k-sekos

Imkime aibę iš N elementų $\Omega = \{\omega_1, \ldots, \omega_N\}$.

9 apibrėžimas *Sakysime,kad seka*

$$\omega_{j_1}, \omega_{j_2}, \ldots, \quad \omega_{j_i} \in \Omega, \tag{105}$$

yra tolygiai pasiskirsčiusi aibėje Ω*, jeigu*

$$\lim_{n\to\infty} \frac{n(\omega_i)}{n} = \frac{1}{N} \quad \forall i. \tag{106}$$

Čia $n(\omega_i)$ *yra elementų* ω_i *skaičius tarp* n *pirmųjų (105) sekos narių.*

10 apibrėžimas *Tolygiai pasiskirsčiusią aibėje* Ω *seką taip pat vadinsime 1-seka.*

11 apibrėžimas *(105) seką vadinsime 2-seka, jei*

$$\lim_{n\to\infty} \frac{n(\omega_i, \omega_j)}{n} = \frac{1}{N^2} \quad \forall i, j. \tag{107}$$

Čia $n(\omega_i, \omega_j)$ *rodo kiek yra porų* (ω_i, ω_j) *tarp* n *pirmųjų sekos* $(\omega_{j_1}, \omega_{j_2}), (\omega_{j_2}, \omega_{j_3}), \ldots$ *narių.*

12 apibrėžimas *(105) seką vadinsime k-seka, jei*

$$\lim_{n\to\infty} \frac{n(\omega_{i_1}, \ldots, \omega_{i_k})}{n} = \frac{1}{N^k} \quad \forall i_1, \ldots, i_k. \tag{108}$$

Čia $n(\omega_{i_1}, \ldots, \omega_{i_k})$ *rodo kiek yra rinkinių* $(\omega_{i_1}, \ldots, \omega_{i_k})$ *tarp* n *pirmųjų sekos* $(\omega_{j_1}, \ldots, \omega_{j_k}), (\omega_{j_2}, \ldots, \omega_{j_{k+1}}), \ldots$ *narių.*

Apibendrinkime šias sąvokas, kai Ω nėra diskreti. Tegul $\Omega = [0, 1)$.

13 apibrėžimas *Seką*

$$x_1, x_2, \ldots, x_n, \ldots, \quad x_n \in [0, 1), \tag{109}$$

vadinsime tolygiai pasiskirsčiusia intervale $[0, 1)$*, jei* $\forall u, v \in [0, 1], u < v$*,*

$$\lim_{n\to\infty} \frac{n([u, v))}{n} = v - u. \tag{110}$$

Čia $n([u, v))$ *yra skaičius sekos (109) narių iš pirmųjų* n*, kurie patenka į intervalą* $[u, v)$*.*

14 apibrėžimas *Tolygiai pasiskirsčiusi intervale* $[0,1)$ *(109) seka taip pat vadinama 1-seka.*

15 apibrėžimas *(109) seką vadinsime 2-seka, jei* $\forall u_1, v_1, u_2, v_2 \in [0,1], u_1 < v_1, u_2 < v_2,$

$$\lim_{n\to\infty} \frac{n([u_1,v_1),[u_2,v_2))}{n} = (v_1-u_1)(v_2-u_2). \tag{111}$$

Čia $n(A,B)$ *rodo kiek porų* $(x_1,x_2),(x_2,x_3),\ldots,(x_n,x_{n+1})$ *priklauso aibei* $A\times B$*.*

Panašiai yra apibrėžiamos k-sekos.

16 apibrėžimas *(109) seką vadinsime k-seka, jei* $\forall u_1, v_1, u_2, v_2, \ldots, u_k, v_k \in [0,1], u_1 < v_1, u_2 < v_2, \ldots, u_k < v_k,$

$$\lim_{n\to\infty} \frac{n([u_1,v_1),[u_2,v_2),\ldots,[u_k,v_k))}{n} = (v_1-u_1)(v_2-u_2)\ldots(v_k-u_k). \tag{112}$$

Čia $n(A_1,A_2,\ldots,A_k)$ *rodo kiek rinkinių* $(x_1,x_2,\ldots,x_k),(x_2,x_3,\ldots,x_{k+1}),$ $\ldots,(x_n,x_{n+1},\ldots,x_{n+k-1})$ *priklauso aibei* $A_1\times A_2\times\cdots\times A_k$*.*

28 teorema *k-seka yra ir (k-1)-seka.*

Įrodymas. Paprastumo dėlei imkime $k=2$.

Tolydžiuoju atveju įrodymas gaunamas iš (111) paėmus $u_2=0, v_2=1$.

Jeigu turime diskretųjį atvejį, tai galima įrodyti prieštaros būdu. Iš tikrųjų, jei ω_k pasikartoja dažniau už ω_l, tai kažkurios poros $(\omega_k,\cdot)$ pasikartos dažniau už poras $(\omega_l,\cdot)$. Teorema įrodyta. ◀

17 apibrėžimas *(105) arba (109) seka vadinama* ∞*-seka, jeigu ji yra k-seka* $\forall k \in \mathbb{N}$.

`k-sekų pavyzdžiai` . . . . . .

k-sekų apibrėžimuose vietoje ribų galima rašyti tikimybes. Galime užrašyti 4 apibrėžimo analogą naudodami tikimybės sąvoką.

18 apibrėžimas *(105) seką vadinsime k-seka, jei*

$$P((\omega_{i_1},\ldots,\omega_{i_k}) = \frac{1}{N^k} \quad \forall i_1,\ldots,i_k. \tag{113}$$

Čia $P((\omega_{i_1},\ldots,\omega_{i_k}))$ *yra rinkinio* $(\omega_{i_1},\ldots,\omega_{i_k})$ *pasirodymo dažnis (arba tikimybė) (105) sekoje.*

19 apibrėžimas *(109) seką vadinsime k-seka, jei* $\forall u_1, v_1, u_2, v_2, \ldots, u_k, v_k \in [0,1], u_1 < v_1, u_2 < v_2, \ldots, u_k < v_k,$

$$(114)\quad P(u_1 \le x_n < v_1, u_2 \le x_{n+1} < v_2, \ldots, u_k \le x_{n+k-1} < v_k) = (v_1 - u_1)(v_2 - u_2)\ldots(v_k - u_k).$$

Čia $P(u_1 \le x_n < v_1, u_2 \le x_{n+1} < v_2, \ldots, u_k \le x_{n+k-1} < v_k)$ *yra (109) sekos paeiliui einančių k elementų patekimo atitinkamai į intervalus* $[u_1, v_1), [u_2, v_2), \ldots, [u_k, v_k)$ *dažnis (arba tikimybė).*

Ar egzistuoja ∞-sekos? Taip.

29 teorema (J.Franklin) *Seka*

$$x_n = \theta^n \mod 1$$

yra ∞*-seka beveik visiems realiems* $\theta > 1$.

Bet iki šiol nei vieno tokio θ nėra surasta.

1 pastaba *Spėjama, kad skaičiaus* π *skaitmenys* $3, 1, 4, 1, 5, 9, \ldots$ *yra* ∞*-seka. Bet nėra įrodyta, kad tai bent 1-seka.*

Egzistuoja algoritmai, kurių pagalba galima apskaičiuoti ∞-sekas.

8.2 Atsitiktinumo samprata. I

Kokias sekas reikėtų vadinti atsitiktinėmis sekomis? Jeigu seka yra 1-seka, tai turbūt dar neatsitiktinė.

Pavyzdys.

Jeigu jau 2-seka, tai labiau atsitiktinė. Aišku,kad (žr. brėž....)

Paanalizuokime ∞-sekas. Jeigu $k = 10^{10}$, tai sekoje pasitaikys serijos iš 10^{10} nuliukų. Šios serijos tikimybė labai ir labai maža. Bet jeigu toks atvejis pasitaiko eksperimente – problema ("tikriausiai negeras generatorius"!). Jeigu tokio atvejo negali pasitaikyti net teoriškai, tai tokia seka jau negali būti "labai" atsitiktinė.

Prie šios problemos dar grįšime vėliau.

8.3 Ryšys tarp diskrečiųjų ir tolydžiųjų k-sekų

Tegul $Y_0, Y_1, \ldots$ – b-tainė seka ($\forall Y_n$ yra vienas iš skaičių $0, 1, \ldots, b-1$). Aišku, kad b-tainė seka yra k-seka, jeigu visiems b-tainiams skaičių rinkiniams $(y_1, y_2, \ldots, y_k)$

$$P((Y_n, Y_{n+1}, \ldots, Y_{n+k-1}) = (y_1, y_2, \ldots, y_k)) = \frac{1}{b^k}.$$

30 teorema *Tegul $x_0, x_1, \ldots$ – k-seka iš intervalo $[0,1)$. Tuomet $[bx_0], [bx_1], \ldots$ – b-tainė k-seka.*

Įrodymas. Pasirenkam $y_1, \ldots, y_k$. Tegul

$$u_j = \frac{y_j}{b}, \quad v_j = \frac{y_j + 1}{b}, \quad Y_n = [bx_n].$$

Tuomet

$$\begin{aligned} &P((Y_n, Y_{n+1}, \ldots, Y_{n+k-1}) = (y_1, y_2, \ldots, y_k)) \\ &\quad = P(y_1 \le bx_n < y_1 + 1, \ldots, y_k \le bx_k < y_k + 1) \\ &\quad = P(bu_1 \le bx_n < bv_1, \ldots, bu_k \le bx_{n+k-1} < bv_k) \\ &\quad = P(u_1 \le x_n < v_1, \ldots, u_k \le x_{n+k-1} < v_k) \\ &\quad = (v_1 - u_1) \ldots (v_k - u_k) = \frac{1}{b^k}. \end{aligned}$$

Teorema įrodyta. ◀

31 teorema *Tegul*

$$x_0, x_1, \cdots \in [0,1). \tag{115}$$

Jeigu seka $[c^j x_0], [c^j x_1], \ldots$ yra c^j-tainė k-seka kokiam nors $c \in \mathbb{N}, c > 1$, ir $\forall j \in \mathbb{N}$, tai (115) seka yra k-seka.

32 pavyzdys *Pateiksime pavyzdį, parodantį, kad teoremos formulavime sąlyga $\forall j \in \mathbb{N}$ yra būtina.*

Nagrinėkime seką

$$0.000,\ 0.100,\ 0.200,\ \ldots,\ 0.900,\ 0.000,\ 0.100,\ \ldots$$

Tokia seka nėra net 1-seka. Tačiau paėmę $c = 10, j = 1$, gausime seką $0, 1, 2, \ldots, 9, 0, 1, \ldots$. Ji yra 10-tainė 1-seka. ◀

31 teoremos įrodymas. Reikia įrodyti, kad

$$P^* = P(u_1 \le x_n < v_1, \ldots, u_k \le x_{n+k-1} < v_k) = \prod_{t=1}^{k} (v_t - u_t).$$

1. Tarkime, kad visi $u_t, v_t, t = 1, \ldots, k$, yra b-tainiai racionalūs skaičiai, t.y. trupmenos su vardikliu lygiu b, o seka $[bx_0], [bx_1], \ldots$ yra b-tainė k-seka. Tuomet

$$u_t = \frac{m_t}{b}, \quad v_t = \frac{s_t}{b}, \quad t = 1, \ldots, k,$$

o

$$P([bx_n] = y_1, \ldots, [bx_{n+k-1}] = y_k) = \frac{1}{b^k}$$

visiems $y_j \in \{0, 1, \ldots, b-1\}$. Žr. brėž..... Taigi, turėsime,kad

$$\begin{aligned}
P^* &= P\left(\frac{m_1}{b} \le x_n < \frac{m_1+1}{b}, \cdot\right) + P\left(\frac{m_1+1}{b} \le x_n < \frac{m_1+2}{b}, \cdot\right) \\
&\quad + \cdots + P\left(\frac{s_1-1}{b} \le x_n < \frac{s_1}{b}, \cdot\right) = \ldots \\
&= P\left(\frac{m_1}{b} \le x_n < \frac{m_1+1}{b}, \ldots, \frac{m_k}{b} \le x_{n+k-1} < \frac{m_k+1}{b}\right) \\
&\quad + \cdots + P\left(\frac{s_1-1}{b} \le x_n < \frac{s_1}{b}, \ldots, \frac{s_k-1}{b} \le x_{n+k-1} < \frac{s_k}{b}\right) \\
&= P\left([bx_n] = m_1, \ldots, [bx_{n+k-1} = m_k\right) \\
&\quad + \cdots + P\left([bx_n] = s_1, \ldots, [bx_{n+k-1} = s_k\right) \\
&= \frac{1}{b^k} + \cdots + \frac{1}{b^k} = \prod_{t=1}^{k}(v_t - u_t).
\end{aligned}$$

2. Tarkime,kad bent vienas iš skaičių u_t, v_t nėra b-racionalus. Pažymėkime (žr. brėž....) u_t' ir v_t' artimiausius b-racionalius skaičius tokius, kad

$$u_t' \le u_t < u_t' + \frac{1}{b}, \quad v_t' \le v_t < v_t' + \frac{1}{b}.$$

Be to, b (dėka j) galime paimti tokį didelį, kad

$$u_t' + \frac{1}{b} < v_t'.$$

Tuomet

$$\begin{aligned}
P^* &\le P(u_1' \le x_n < v_1' + \frac{1}{b}, \ldots, u_k' \le x_{n+k-1} < v_k' + \frac{1}{b}) \\
&= \prod_{t=1}^{k}\left(v_t' - u_t' + \frac{1}{b}\right) = \prod_{t=1}^{k}\left(\frac{v_t' - u_t' + \frac{1}{b}}{v_t - u_t}\right)(v_t - u_t) \\
&\le \prod_{t=1}^{k}(v_t - u_t)\prod_{t=1}^{k}\left(1 + \frac{2}{(v_t - u_t)b}\right).
\end{aligned}$$

Pasirinkime $\epsilon > 0$. b (j dėka) galime paimti tokį didelį, kad antroji sandauga

$$\prod_{t=1}^{k}\left(1 + \frac{2}{(v_t - u_t)b}\right) \le 1 + \epsilon.$$

Todėl

$$P^* \le \prod_{t=1}^{k}(v_t - u_t)(1+\epsilon) \le \prod_{t=1}^{k}(v_t - u_t) + \epsilon.$$

Analogiškai samprotaudami turėsime

$$\begin{aligned} P^* &\ge P(u_1' + \frac{1}{b} \le x_n < v_1', \dots, u_k' + \frac{1}{b} \le x_{n+k-1} < v_k') \\ &= \prod_{t=1}^{k}\left(v_t' - u_t' - \frac{1}{b}\right) = \prod_{t=1}^{k}\left(\frac{v_t' - u_t' - \frac{1}{b}}{v_t - u_t}\right)(v_t - u_t) \\ &\ge \prod_{t=1}^{k}(v_t - u_t)\prod_{t=1}^{k}\left(1 - \frac{2}{(v_t - u_t)b}\right) \\ &\ge \prod_{t=1}^{k}(v_t - u_t)(1-\epsilon) \le \prod_{t=1}^{k}(v_t - u_t) - \epsilon. \end{aligned}$$

Taigi, $\forall \epsilon > 0$

$$\left|P^* - \prod_{t=1}^{k}(v_t - u_t)\right| \le \epsilon.$$

Vadinasi

$$P^* = \prod_{t=1}^{k}(v_t - u_t).$$

Kadangi galime imti $b = c, c^2, \dots, c^n \to \infty,$ kai $n \to \infty$, tai teorema įrodyta. ◀

2 pastaba *3 teoremos tvirtinimas reiškia,kad realių skaičių iš intervalo* $[0,1)$ *seka bus k-seka, jei jų c-tainių skleidinių pirmieji skaitmenys sudaro k-seką, pirmieji du skaitmenys sudaro k-seką ir .t.t.*

33 pavyzdys *Tarkime,* $0.2453731...; 0.4678902...; \dots$ *yra seka iš intervalo* $[0,1)$.

Ji bus k-seka, jei

seka $2, 4, \dots$ yra k-seka,
seka $24, 46, \dots$ yra k-seka,
seka $245, 467, \dots$ yra k-seka, ir t.t.

◀

8.4 k-sekų savybės

32 teorema *Tegul $x_0, x_1, \ldots$ yra k-seka iš intervalo $[0,1)$, $f(y_1, \ldots, y_k)$ – integruojama Rymano prasme funkcija. Tuomet*

(116)
$$\lim_{n\to\infty} \frac{1}{n} \sum_{0\le j<n} f(x_j, x_{j+1}, \ldots, x_{j+k-1}) = \int_0^1 \ldots \int_0^1 f(y_1, \ldots, y_k)\, dy_1 \ldots dy_k.$$

Įrodymas. Iš k-sekos apibrėžimo išplaukia, kad teorema teisinga kai

(117)
$$f(y_1, \ldots, y_k) = \begin{cases} 1, & \text{kai } u_1 \le y_1 < v_1, \ldots, u_k \le y_k < v_k, \\ 0 & \text{kitais atvejais,} \end{cases}$$

nes šiuo atveju

$$\begin{aligned} &\lim_{n\to\infty} \frac{1}{n} \sum_{0\le j<n} f(x_j, x_{j+1}, \ldots, x_{j+k-1}) \\ &= \lim_{n\to\infty} \frac{n([u_1, v_1), [u_2, v_2), \ldots, [u_k, v_k))}{n} \\ &= (v_1 - u_1)(v_2 - u_2) \ldots (v_k - u_k) \\ &= \int_0^1 \ldots \int_0^1 f(y_1, \ldots, y_k)\, dy_1 \ldots dy_k. \end{aligned}$$

(116) lygybė yra teisinga ir laiptuotai funkcijai

$$f = a_1 f_1 + \cdots + a_m f_m;$$

čia f_j yra funkcijos, įgyjančios tik dvi reikšmes 0 ir 1 (kaip formulėje (117)), o a_j yra pastovūs daugikliai.

Jeigu f yra integruojama Rymano prasme funkcija, o $\epsilon > 0$, tai egzistuoja dvi laiptuotos funkcijos $\underline{f}$ ir $\overline{f}$, $\underline{f} \le f \le \overline{f}$ ir skirtumas tarp integralų $\int \overline{f}, \int \underline{f}$ yra nedidesnis negu ϵ. Taigi

(118)
$$\begin{aligned} &\frac{1}{n} \sum_{0\le j<n} \underline{f}(x_j, \ldots, x_{j+k-1}) \\ &\le \frac{1}{n} \sum_{0\le j<n} f(x_j, \ldots, x_{j+k-1}) \\ &\le \frac{1}{n} \sum_{0\le j<n} \overline{f}(x_j, \ldots, x_{j+k-1}). \end{aligned}$$

Integralams, aišku, yra teisinga nelygybė

$$\begin{aligned}&\int_0^1 \dots \int_0^1 \underline{f}(y_1, \dots, y_k)\, dy_1 \dots dy_k \\ &\le \int_0^1 \dots \int_0^1 f(y_1, \dots, y_k)\, dy_1 \dots dy_k \\ &\le \int_0^1 \dots \int_0^1 \overline{f}(y_1, \dots, y_k)\, dy_1 \dots dy_k.\end{aligned}$$

Kadangi pirmosios sumos iš (118) formulės riba, kai $n \to \infty$, yra pirmasis integralas, o trečiosios sumos riba – trečiasis integralas. Be to, šie integralų skirtumas ne didesnis už ϵ, ir ϵ galime pasirinkti bet kokį teigiamą skaičių, tai vidurinės sumos iš (118) formulės riba, kai $n \to \infty$, yra vidurinis integralas, t.y.

$$\lim_{n\to\infty} \frac{1}{n} \sum_{0\le j<n} f(x_j, \dots, x_{j+k-1}) = \int_0^1 \dots \int_0^1 f(y_1, \dots, y_k)\, dy_1 \dots dy_k.$$

Teorema įrodyta. ◀

1 išvada *Jeigu seka $x_1, x_2, \dots$ iš intervalo $[0, 1)$ yra k-seka, tai ji tenkina k-tos eilės kėlinių testą:*

$$P(x_{n+p_1-1} < x_{n+p_2-1} < \dots < x_{n+p_k-1}) = \frac{1}{k!}\,.$$

Čia $p_1, \dots, p_k$ bet koks skaičių $1, \dots, k$ kėlinys.

Įrodymas. Išvada išplaukia iš 4 teoremos, paėmus

$$f(y_1, \dots, y_k) = \begin{cases} 1, & \text{kai } y_{p_1} < y_{p_2} < \dots < y_{p_k}, \\ 0 & \text{kitais atvejais,} \end{cases}$$

nes šiuo atveju

$$
\begin{aligned}
&P(x_{n+p_1-1} < x_{n+p_2-1} < \cdots < x_{n+p_k-1}) \\
&= \int_0^1 \ldots \int_0^1 f(y_1, \ldots, y_k)\, dy_1 \ldots dy_k \\
&= \iint \ldots \int_{0 \le y_{p_1} < y_{p_2} < \cdots < y_{p_k} < 1} dy_1 dy_2 \ldots dy_k \\
&= \int_0^1 dy_{p_k} \int_0^{y_{p_k}} dy_{p_{k-1}} \ldots \int_0^{y_{p_3}} dy_{p_2} \int_0^{y_{p_2}} dy_{p_1} \\
&= \int_0^1 dy_{p_k} \int_0^{y_{p_k}} dy_{p_{k-1}} \ldots \int_0^{y_{p_3}} y_{p_2}\, dy_{p_2} \\
&= \int_0^1 dy_{p_k} \int_0^{y_{p_k}} dy_{p_{k-1}} \ldots \int_0^{y_{p_4}} \frac{y_{p_3}^2}{2}\, dy_{p_3} \\
&= \int_0^1 dy_{p_k} \int_0^{y_{p_k}} \frac{y_{p_{k-1}}^{k-2}}{(k-2)!}\, dy_{p_{k-1}} \\
&= \int_0^1 \frac{y_{p_k}^{k-1}}{(k-1)!}\, dy_{p_k} = \frac{1}{k!}\,.
\end{aligned}
$$

◀

3 pastaba *∞-sekos tenkina daug kitų testų.*

8.5 Atsitiktinumo samprata. II

Pateiksime pavyzdį, rodantį, kad ir k-sekas negalima laikyti atsitiktinėmis.

34 pavyzdys $0,0,0,1,\ 0,0,0,1,\ 1,1,0,1,\ 1,1,0,1,\ 0,0,0,1,\ \ldots$

Šios sekos periodas yra 16 ir ji yra 3-seka (nesunku patikrinti). Paimkime lyginius ir nelyginius sekos narius:

$0,1,0,1,\ 1,1,1,1,\ 0,1,0,1,\ 1,1,1,1,\ldots$ – lyginiai sekos nariai
$0,0,0,0,\ 1,0,1,0,\ 0,0,0,0,\ 1,0,1,0,\ldots$ – nelyginiai sekos nariai.

Pirmuoju atveju dažniau pasirodo vienetukai, antruoju – nuliukai. Taigi posekiai $\{x_{2n}\}$ ir $\{x_{2n-1}\}$ jau nebėra net 1-sekos. ◀

8.6 (m,k)-sekos

20 apibrėžimas *Seka $x_1, x_2, \ldots$ vadinama (m,k)-seka intervale $[0,1)$, jei*

$$P(u_1 \le x_{mn+j} < v_1, \ldots, u_k \le x_{mn+j+k-1} < v_k) = (v_1 - u_1)\ldots(v_k - u_k)$$

visiems $u_i, v_i \in [0,1]$, $u_i < v_i$, ir visiems j, $0 \le j < m$.

33 teorema (A.Nieven, X.Zuckermann) *∞-seka yra (m,k)-seka visiems natūriniams m ir k.*

Tai gražus nelauktas rezultatas.

8.7 Atsitiktinumo samprata. III

Taigi k-sekas,tur būt, nevisada galime laikyti atsitiktinėmis. O kaip bus su ∞-sekomis?

Ar ekvivalentūs teiginiai: "∞-seka" ir "atsitiktinė seka".

Dabar pateiksime daug atsitiktinių sekų apibrėžimų ir paanalizuosime juos.

21 apibrėžimas *Seka iš intervalo $[0,1)$ vadinama atsitiktine, jei ji yra ∞-seka.*

Bendrai paėmus, jeigu seka tikrai atsitiktinė, taigalimas variantas, kad ji net netolygiai pasiskirsčiusi intervale $[0,1)$. Bet, tikriausiai, su tikimybe lygia 1 ji yra ∞-seka. Kokią bepaimtume seką atsitiktinių skaičių, ji turi būti lygiavertė kitai tokiai sekai, net sekai vien iš nuliukų.

Bet gal galėtume atsisakyti minėtų sekų. Mums liktų pakanakmai daug gerų sekų.

Bet, sakykime, turime atsitiktinę pagal 12 apibrėžimą seką. Jos posekis

$$x_0, x_1, x_4, x_9, \ldots, x_{n^2}, \ldots$$

taip pat turi būti atsitiktinis. Jeigu paimtume minėto posekio visus narius lygius nuliui, tai seka vis tiek būtų atsitiktinė pagal 12 apibrėžimą, nes šių narių yra palyginti retai ir ribiniam skaičių pasiskirstymui jie įtakos neturi.

Taigi 12 apibrėžimas nėra geras.

22 apibrėžimas *Seka iš intervalo $[0,1)$ vadinama atsitiktine, jei kiekvienas jos begalinis posekis yra ∞-seka.*

Šitas apibrėžimas tikrai per griežtas, nes nebelieka atsitiktinių sekų. Juk iš kiekvienos aprėžtos sekos galime išskirti monotoninį (ir net griežtai) posekį, kuris nebus ∞-seka.

Vadinasi reikia taip apriboti ∞-sekas, kad posekius galėtume imti ne visus, o tik pagal kokias nors taisykles.

23 apibrėžimas *Seka iš intervalo* $[0,1)$ *vadinama atsitiktine, jei bet kokiu efektyviu algoritmu apibrėžtas jos begalinis posekis yra* ∞*-seka.*

Čia efektyvus = konkretus, t.y. gali būti sudaryta programa, kuri nustato natūraliųjų skaičių posekį.

Šis apibrėžimas irgi nekoks. Pavyzdžiui, mėtome monetą: $H, H, S, S, H, \ldots$ Norime gauti posekį, kuris gaunamas, kai imami nariai po dviejų herbų pasirodymo arba kas nors panašaus.

24 apibrėžimas *Tegul funkcijos*

$$f_0, f_1(x_1), f_2(x_1, x_2), \ldots, f_n(x_1, \ldots, x_n), \ldots$$

tokios, kad

- $f_i,\ i = 1, 2, \ldots,$ *yra efektyviai apskaičiuojamos;*
- $f_i,\ i = 1, 2, \ldots,$ *įgyja tik dvi reikšmes 0 ir 1.*

Sekos $x_0, x_1, \ldots$ $i+1$*-ąjį narį paliekane, jei* $f_i = 1$*, ir* $i+1$*-ąjį narį išmetame, jei* $f_i = 0$*.*

Taip gauta nauja seka vadinama sekos $x_0, x_1, \ldots$ *retiniu.*

25 apibrėžimas *Seka iš intervalo* $[0,1)$ *vadinama atsitiktine, jei kiekvienas šios sekos retinys yra* ∞*-seka.*

35 pavyzdys *Jei*

$$f_k = \begin{cases} 1, & \text{kai } k \text{ lyginis}, \\ 0, & \text{kai } k \text{ nelyginis}, \end{cases}$$

tai paliekami tik nelyginiai sekos nariai.

36 pavyzdys *Mėtome monetą. Jei*

$$f_k(x_1, x_2, \ldots, x_k) = f_k(x_{k-99}, x_{k-98}, \ldots, x_k) = \begin{cases} 1, & x_{k-99} = x_{k-98} = \cdots = x_k = H, \\ 0, & \text{kitais atvejais}, \end{cases}$$

tai paliekami tik sekos nariai po 100 *herbų pasirodymo.*

Dažnai sekų, apibrėžtų kaip atsitiktinių su 16 apibrėžimu, pakanka, nors yra prigalvota ir kitokių apibrėžimų.

8.8 Baigtinės atsitiktinės sekos

Turbūt aišku, kad seka 011101001 daugiau atsitiktinė nei seka 010101010, o pastaroji labiau atsitiktinė nei seka 000000000. Pateikime pora baigtinės sekos atsitiktinumo apibrėžimo variantų.

26 apibrėžimas *b-tainę baigtinę seką $X_1, X_2, \ldots, X_N$ vadinsime k-seka, jeigu*

$$\left| P\left((X_n, X_{n+1}, \ldots, X_{n+k-1}) = (x_1, x_2, \ldots, x_k)\right) - \frac{1}{b^k} \right| \leq \frac{1}{\sqrt{N}}$$

visiems b-tainiams skaičių rinkiniams $x_1, x_2, \ldots, x_k$.

$$\left(Arba \quad \left| \approx \frac{\nu(N)}{N} - \frac{1}{b^k} \right| \leq \frac{1}{\sqrt{N}} . \right)$$

Kodėl $\frac{1}{\sqrt{N}}$?

27 apibrėžimas *N ilgio b-tainė seka vadinama atsitiktine, jeigu ji yra k-seka visiems $k \leq \log_b N$.*

37 pavyzdys *Suraskime visas atsitiktines (pagal pateiktą apibrėžimą) dvejetaines sekas, kurių ilgiai lygūs 11.*

Jeigu nagrinėtume dvejetaines sekas, kurių ilgiai lygūs 11, tai pagal pateiktą apibrėžimą iš visų tokių sekų (iš viso jų yra $2^{11} = 2048$) neatsitiktinių bus 170. Kadangi

$$3 < \log_2 11 < 4,$$

tai reikia patikrinti ar nagrinėjamos sekos yra 1-sekos, 2-sekos ir 3-sekos su tikslumu

$$\frac{1}{\sqrt{N}} = \frac{1}{\sqrt{11}} \approx 0.3.$$

Visos sekos, kuriose yra mažiau kaip 2 vienetukai arba mažiau kaip 2 nuliukai nebus net 1-sekos. Pavyzdžiui, 11000000000. Šiuo atveju vienetuko tikimybė 2/11, todėl

$$\left| P(X_n = 1) - \frac{1}{2} \right| = \left| \frac{2}{11} - \frac{1}{2} \right| > 0.3.$$

Sekos, kuriose yra 7 vienetukai ar 7 nuliukai iš eilės nebus 2-sekomis. Pavyzdžiui, 00000001110. Šiuo atveju poros 00 tikimybė yra 6/10, todėl

$$\left| P((X_n, X_{n+1}) = (0,0)) - \frac{1}{4} \right| = \left| \frac{6}{10} - \frac{1}{4} \right| = 0.35 > 0.3.$$

◄

2 uždavinys *Ar sekos* 00000111111, 01010101010, 00011100011, *iš sekų minimų 6 pavyzdyje, yra atsitiktinės pagal pateiktą apibrėžimą?*

3 uždavinys *Pabaikite spręsti 6 pavyzdį. Suraskite visas 170 sekų, kurios nėra atsitiktinės pagal pateiktą apibrėžimą.*

28 apibrėžimas *b-tainė baigtinė seka $X_1, X_2, \ldots, X_N$ vadinama (n, ϵ)-atsitiktine algoritmų aibės $\mathbb{A}$ atžvilgiu, jeigu kiekvienam algoritmu iš aibės $\mathbb{A}$ apibrėžtam posekiui $X_{t_1}, X_{t_2}, \ldots, X_{t_m}$ teisinga viena iš nelygybių:*

$$m < n$$

arba

$$\left|\frac{1}{m}\,\nu_a(X_{t_1}, X_{t_2}, \ldots, X_{t_m}) - \frac{1}{b}\right| \leq \epsilon \quad \forall a,\ 0 \leq a < b.$$

Čia $\nu_a(x_1, \ldots, x_m)$ yra skaičiaus a pasirodymų skaičius sekoje $x_1, \ldots, x_m$.

Kitais žodžiais tariant, kiekvienas pakankamai ilgas posekis (apibrėžtas algoritmo iš $\mathbb{A}$ pagalba) turi būti apytikriai tolygiai pasiskirstęs. Aibė $\mathbb{A}$, žinoma, turėtų būti sudaryta iš palyginti nesudėtingų algoritmų.

38 pavyzdys *Nagrinėkime baigtinę dvejetainę seką $X_1, \ldots, X_8$ ir algoritmų aibę $\mathbb{A}$:*

1. *visa seka,*
2. *nelyginiai sekos nariai,*
3. *sekos nariai, einantys po* 0*,*
4. *sekos nariai, einantys po* 1*.*

Kada tokia seka yra $(4, \frac{1}{8})$-atsitiktinė?

Seka $X_1, \ldots, X_8$ yra $(4, \frac{1}{8})$-atsitiktinė, jei

1. $$\left|\frac{1}{8}(X_1 + X_2 + \cdots + X_8) - \frac{1}{2}\right| \leq \frac{1}{8}\,,$$
   t.y. jeigu sekoje yra 3, 4 arba 5 vienetukai;
2. $$\left|\frac{1}{4}(X_1 + X_3 + X_5 + X_7) - \frac{1}{2}\right| \leq \frac{1}{8}\,,$$
   t.y. 2 vienetukai yra nelyginėse vietose;

3. tarp $X_1, X_2, \ldots, X_7$ yra (1, 6 ir 7 nuliukų atvejus nenagrinėjame, nes tada neišpildyta 1 sąlyga)

   (a) 2 arba 3 nuliukai, tai $m < n = 4$ ir nieko tikrinti nereikia;

   (b) 4 nuliukai, tai už jų turi stovėti 2 nuliukai ir 2 vienetukai: X_{t_1}, X_{t_2}, X_{t_3}, X_{t_4}, kad

   $$\left|\frac{1}{4}(X_{t_1} + X_{t_2} + X_{t_3} + X_{t_4}) - \frac{1}{2}\right| \leq \frac{1}{8};$$

   (c) 5 nuliukai, tai už jų turi stovėti 2 arba 3 nuliukai ir kiti vienetukai: $X_{t_1}, X_{t_2}, X_{t_3}, X_{t_4}, X_{t_5}$, kad

   $$\left|\frac{1}{5}(X_{t_1} + X_{t_2} + X_{t_3} + X_{t_4} + X_{t_5}) - \frac{1}{2}\right| \leq \frac{1}{8};$$

4. visiškai analogiškai 3 sąlygai. ◀

LITERATŪRA

[1] **J.Banks, J.S.Carson, II, B.L.Nelson**, *Discrete-Event System Simulation*, Prentice-Hall, Upper Saddle River New Jersey, 1996.

[2] **W.G.Bulgren**, *Discrete System Simulation*, Prentice-Hall, Englewood Cliffs New Jersey, 1982.

[3] **I.Deák**, *Random Number Generators and Simulation*, Akadémiai Kiadó, Budapest, 1990.

[4] **L. Devroye**, *Non-Uniform Random Variate Generation*, Springer-Verlag, New York, 1986.

[5] **S.M.Ermakov**, *The Monte-Carlo Method and Adjoining Questions*, Nauka, Moscow, 1975. (Rusų kalba)

[6] **G. S. Fishman**, *Monte Carlo: Concepts, Algorithms, and Applications*, Springer-Verlag, New York Berlin Heidelberg, 1995.

[7] **G.S.Fishman**, *Principles of Discrete Event Simulation*, John Wiley & Sons, New York Brisbane Chichester Toronto, 1978.

[8] **W.J.Graybeal, U.W.Pooch**, *Simulation: Principles and Methods*, Winthrop Publishers, Cambridge Massachusetts, 1980.

[9] **D.E.Knuth**, *The Art of Computer Programing.* V 2, *Seminumerical Algorithms*, Addison-Wesley Publishing Company, London, 1969. (Yra vertimas į rusų kalbą, 1977)

[10] **A.M.Law, W.D.Kelton**, *Simulation Modeling and Analysis*, McGraw-Hill Book Company, New York and oth., 1981.

[11] **I.Manno**, *Introduction to the Monte-Carlo Method*, Akadémiai Kiadó, Budapest, 1999.

[12] **B.D.Ripley**, *Stochastic Simulation*, John Wiley & Sons, New York Chichester Brisbane Toronto Singapore, 1987.

[13] **S.M.Ross**, *Introduction to Probability Models*, Academic Press, Boston San Diego New York London Sydney Tokyo Toronto, 1993.

[14] **R.Y.Rubinstein**, *Simulation and the Monte Carlo Method*, John Wiley & Sons, New York Chichester Brisbane Toronto, 1981.