Atsitiktiniai įvykiai ir dydžiai. Tikimybės supratimas. Atsitiktiniai skaičiai

1. Atsitiktinybė

Žodžiais *atsitiktinumas* ar *atsitiktinybė* kasdieninėje kalboje dažniausiai apibūdiname tokius nutikimus, įvykius ir jų rezultatus, kurių iš principo kaip ir *neįmanoma* tiksliai *nuspėti*, prognozuoti. Todėl tie žodžiai dažnai patenka į vieną gretą su kitais kiek panašios reikšmės žodžiais, tokiais kaip chaosas, netvarka, anarchija ir pan., ir iš jų tam tikra dalimi netgi perima neigiamą konotacinį atspalvį. Kita gi vertus, atsitiktinybės priešprieša yra tvarka, reguliarumas, dėsningumas, priežastingumas – žodžiu, ta sfera, kurioje vieni įvykiai su kitais yra susiję aiškiais, vienareikšmiais priežasties–pasekmės tipo ryšiais. Vėlgi susiduriame su gana aiškia tipologine įvykių sąryšingumo *poliarizacija*: vienareikšmiai priežastiniai ryšiai viename poliuje, visiškos atsitiktinybės valdos – kitame, o visas tas "tarpas" tarp jų – gana problemiškas, nes čia pasireiškia ir vienokio, ir kitokio pobūdžio sąryšingumas, kur dažniausiai veikia ne viena aiški, o daugelis įvairiai tarpusavy susipynusių priežasčių, ir faktiškasis priklausomybės laipsnis bei tipas – labai įvairuoja. Tad paprastai žiūrima *galimumo prognozuoti*: ten, kur įvykius iš esmės galima numatyti, sakoma, jog vyrauja priežastiniai ryšiai, o ten, kur tegalimos tik apytikrės prognozės, kur spėjamasis įvykis gali arba tikrai įvykti, arba ir neįvykti, sakoma, jog dominuoja atsitiktinumas.

Įvykių sąryšingumo, priežastingumo bei atsitiktinumo klausimai gali būti sprendžiami net filosofijos plotmėje, ir įvairios filosofijos kryptys į juos žiūri nevienodai. Teistinio pobūdžio filosofiniai mokymai iš esmės visų įvykių bei nutikimų priežastimi laiko Dievo valią, o materialistinių pakraipų filosofija stengiasi iškelti objektyvųjį, pasakytum, pačios natūros sąlygojamą priežastingumą. Pabrėžtina, kad tarpais priežastingumas traktuojamas perdėm supaprastintai, schematiškai (mechanistinis determinizmas racionalizmo epochoje, kai atrodė, jog Niutono ar pan. dėsniai paaiškina jau viską – nuo obuolio kritimo iki planetų judėjimo; darvinistinių modelių tiesmukas taikymas socialiniams reiškiniams ir pan.). Įvairiais pavidalais gana dažnai sutinkama mintis, kad įvykiai mums tiktai *atrodo esą* atsitiktiniai – pirmiausia dėl to, kad mes nepažįstame ir nemokame įžvelgti tikrųjų jų priežasčių, – nėra iš esmės priimtina: atsitiktinybė iš tikrųjų nėra kokia nors priežastingumo priešybė, o greičiau jau – viena iš jo formų ar pavidalų. Taip galvodami, ir susidurtume su labai specifišku ir kaip tik šiai formai labiausiai būdingu matmeniu – su *tikimybe*.

2. Atsitiktiniai įvykiai ir dydžiai

Atsitiktiniu vadinamas toks įvykis, kurį sąlygoja ne viena kokia, o daugelis įvairių kartu veikiančių priežasčių, kurių kiekvienos, atskirai paimtos, įtaka įvykiui, jo baigmei tėra palyginti nedidelė ir lemiamos svarbos neturi. Tokiomis sąlygomis įvykio baigmė ir priklauso nuo visos to priežasčių komplekso visumos. O kadangi tas kompleksas yra nepastovus, dinamiškas ir kartojantis įvykiams įvairiai kinta bei nuolatos mainosi, tai tiksliai numatyti tokių įvykių baigmės ir pasidaro neįmanoma, t. y. tie įvykiai darosi atsitiktiniai, pavaldūs atsitiktinybei. Todėl ir akcentuojama, kad *atsitiktinis* yra toks įvykis, kuris, nors ir kartodamasis iš esmės tomis pačiomis sąlygomis, *kiekvieną sykį* vyksta vis *šiek tiek kitaip* (dėl to jo baigmės iš anksto vienareišmiškai numatyti neįmanoma).

Atsitiktinių dalykų – įvykių ar procesų, ar dydžių – pavyzdžių apstu kone kiekviename žingsnyje: metereologijos reiškiniai, iš esmės tapačių biologinių individų (vienu metu sodintų vienos rūšies augalų, vienos vados gyvuliukų ir pan.) augimas vienodomis sąlygomis, vaikų protinis bei fizinis vystymasis, tų pačių daiktų kainos turguje, per dieną parodą aplankančių žmonių skaičius, medžių kiekis miško hektare (ar kitokio pastovaus dydžio sklype), įvairių sporto varžybų rezultatai...

Keletas "lingvistinių" pavyzdžių: sakinių ilgis tekste x; žodžių, prasidedančių balse, kiekis kiekviename (pilname, to paties formato) teksto puslapyje, žodžių ilgis skiemenimis ar raidėmis, pagrindinio tono dažnis hercais tariant tą patį garsą įvairiems žmonėms, nekirčiuotų skiemenų kiekis tarp dviejų gretimų kirčių rišliame tekste ir t. t.

3. Įvykių „įvykimas", jų klasifikacija ir struktūra

Įvykius šiaipjau yra įprasta grupuoti ir klasifikuoti pagal tai, kokie yra jų rezultatai, kaip jie pasibaigia. Tai visiškai tinka ir atsitiktiniams įvykiams. Paprasčiausi yra 2 paskaitoje aptarti įvykiai, kurie tegali baigtis dvejopai: arba vienaip, arba kitaip: gimsta arba berniukas, arba mergaitė; metant kapeiką atsiverčia arba skaičius, arba herbas; iš kaladės traukiant kortą, ji būna arba raudonos, arba juodos spalvos, oro prognozė arba pasitvirtina, arba ne... Jeigu įvykis *pasibaigia taip, kaip spėjame* arba laukiame, sakome, kad įvykis *įvyko*: gimė mergaitė, atsivertė herbas, prognozė pasitvirtino. Jei pasibaigia priešingai, sakoma, jog *įvyko priešingas įvykis*: gimė berniukas, prognozė nepasitvirtino, atsivertė skaičius. Tad binariškų įvykių atveju tarsi gali įvykti *du skirtingi* įvykiai: arba spėjamasis, arba jam priešingas. Įvykus gi priešingam įvykiui, apie spėjamąjį kartais sakoma, jog jis ***neįvyko***: ta prasme bet kuris binariškas įvykis (atsiverčia herbas, gimsta berniukas ir pan.) gali arba *įvykti*, arba *neįvykti* (vadinasi, tada įvyksta jam priešingas įvykis).

Kebliau yra su įvykiais, kurie gali pasibaigti ne dvejopai, o keliariopai, įvairiais būdais: metant ne kapeiką, o šešiabriaunį žaidimų kauliuką gali atsiversti bet kuri viena iš šešių jo pusių, pirmoji teksto puslapio raidė gali būti bet kuri iš abėcėlės raidžių, pirmasis (ar paskutinysis, ar kuris kitas) paimtas sakinio žodis gali būti bet kuri iš galimų kalbos dalių ir pan.

Tačiau ir šiuo atveju nesunku yra įvykį kaip ir dirbtinai *binarizuoti*, t.y. vieną kurią spėjamą jo baigmę supriešinti su visomis kitomis likusiomis: *iškrito vienas taškas/iškrito ne vienas taškas*; tiriamojoje pozicijoje *rasta raidė 'l'/ ne raidė 'l'*; daiktavardis pavartotas *naudininko linksniu/ ne naudininko linksniu*; sakinio ilgis yra keturi žodžiai/ne keturi žodžiai ir t.t. Tokiu būdu įvykis, galintis turėti daug įvairių baigmių, tarsi suskaldomas į daug (faktiškai į tiek, kiek yra galimų baigmių) binariškų įvykių, iš kurių kiekvienas atskiro įvykio metu gali įvykti arba neįvykti. Dėl to tie tikimybiniai įvykių modeliai, kurie tinka binariškiems įvykiams, gali būti pritaikomi ir daugiau baigmių turintiems įvykiams aprašyti – svarbu tik, kad juos būtų galima binarizuoti. Kita vertus, tokie „daliniai“ įvykiai , apie kokius čia kalbama, yra ir savaip elementarūs ta prasme, kad jų jau nebeįmanoma išskaidyti dar smulkiau, į dar „paprastesnius“ įvykius. Dėl to dažnai juos vadina *elementariaisiais* įvykiais. Tokia sąvoka leistų kalbėti taip pat ir apie *įvykių struktūrą* (sąrangą, sanklodą): kiekvienas atsitiktinis įvykis tarsi yra sudarytas iš tam tikro kiekio elementariųjų įvykių, kurių kiekvienas iš principo gali įvykti arba neįvykti ir kurių *tik vienas* tikrai ir būtinai *įvyksta* tam „sudėtiniam“ įvykiui *kartojantis*. Galimų elementariųjų įvykių kiekis atitinka „sudėtinio“ įvykio galimų baigmių skaičių.

ĮSIDĖMĖTINA:

A. Elementariųjų įvykių „rinkinys“, apimantis visas galimas „sudėtinio“ įvykio baigmes, vadinamas *elementariųjų įvykių erdve*. Ji ir atspindi nagrinėjamojo bandymo (eksperimento) ar įvykio įprastine prasme „binarizuotą“ struktūrą. Sakysim:

„Sudėtinis“ įvykis („bandymas“):	Elementariųjų įvykų erdvė:
a) Gimusio vaiko lytis yra	berniukas (1)
	mergaitė (2)
b) Metant kauliuką iškrito	1 taškas (1)
	2 taškai (2)
	3 taškai (3)
	4 taškai (4)
	5 taškai (5)
	6 taškai (6)
c) Sakinys prasideda	daiktavardžiu (1)
	būdvardžiu (2)
	skaitvardžiu (3)
	įvardžiu (4)
	veiksmažodžiu (5)
	prieveiksmiu (6)
	prielinksniu (7)
	jungtuku (8)
	jaustuku (9)
	ištiktuku (10)
	dalelyte (11)

B. Iš tai pačiai erdvei priklausančių elentariųjų įvykių galima išvesti („sudėlioti“) įvairius sudėtinius ***tarpinio*** pobūdžio įvykius, „įtelpančius“ į tą bandymą (sudėtinį įvykį), kurį apibūdina visa elementariųjų įvykių erdvė. Tiesa, tai darosi įmanoma tik tada, kai bandymai (sudėtiniai įvykiai) turi pakankamai daug galimų baigmių, kitaip tariant, yra palyginti sudėtingi. Pavyzdžiui, kauliuko metimo atveju rūpi tarpinio pobūdžio sudėtinis įvykis *iškrito nelyginis taškų skaičius.* Iš b) pavyzdžio aiškiai matyti, kad tokį įvykį turėsime, jeigu įvyks kuris nors iš 1, 3 ar 5 elementariųjų įvykių. Vadinasi, šių trijų konkrečių elementariųjų įvykių visuma (aibė, susieta griežtosios disjunkcijos ryšiu) ir suformuoja rūpimą sudėtinį įvykį. Tokie elementarieji įvykiai, kuriems (tiksliau – kurių vienam) įvykus įvyksta ir rūpimas sudėtinis įvykis (pavadinkime jį įvykiu *X*), vadinami jam (įvykiui *X*) *palankiais* elementariaisiais įvykiais. Taigi:

	Elementarieji įvykiai:	Palankūs sudėtiniam įvykiui:
b) pavyzdys:	1, 3 ir 5	*Iškrito nelyginis taškų skaičis*;
	2, 4 ir 6	*Iškrito lyginis taškų skaičius*
	3 ir 6	*Iškritusių taškų skaičius be liekanos dalijasi iš 3*
	3, 4, 5 ir 6	*Iškrito daugiau kaip du taškai*
c) pavyzdys:	1, 2, 3 ir 4	*Sakinys prasideda vardažodžiu*
	7, 8, 8, 10 ir 11	*Sakinys prasideda nesavarankišku žodžiu*

Atsimintina, kad:
* Vieno bandymo metu gali įvykti ***tik vienas*** kuris nors iš tai pačiai erdvei priklausančių elementariųjų įvykių (jei konkretus sakinys prasideda, tarkime, prieveiksmiu tai jis neprasideda jokia kita kalbos dalimi). Kitaip tariant, vienai erdvei priklausantys elementarieji įvykiai tarpusavy yra susieti *griežtosios disjunkcijos* ryšiu (kai elementariųjų įvykių

erdvėje tėra tik du, t.y. kai tiriamas įvykis yra binariškas, ta elementariųjų įvykių disjunkcija irgi yra dvinarė, o kitais atvejais tampa daugianarė). Dėl to jie dažnai apibūdinami kaip tarpusavy *nesuderinami* elementarieji atsitiktiniai įvykiai.
* Įvykiai (nebe elementarūs!), kurie neturi *nė vieno* sau *palankaus* elementariojo atsitiktinio įvykio, vadinami ***negalimais*** įvykiais. Kauliuko metimo atveju toks būtų, sakysim, įvykis *Iškrito daugiau kaip šeši taškai*: nėra nė vieno jam palankaus elementriojo įvykio, o ir sveiku praktiniu protu suvokiame, kad toks įvykis negalimas.
* Įvykis, kuriam *palankūs visi* tam tikros erdvės elementarieji įvykiai, vadinamas ***būtinuoju***, nes atliekant bandymą jis įvyksta kiekvienąkart, – būtinai, neišvengiamai. Kauliuko metimo bandymuose toks būtų, tarkim, įvykis *Iškrito ne daugiau kaip šeši taškai*: jam palankūs yra visi b) pavyzdžio elementarieji įvykiai, ir iš tikro suvokiame, jog kad ir kuri briauna beatsiverstų, toks įvykis iš tikro įvyks kiekvienąkart.
* Negalimieji įvykiai ir būtinieji įvykiai yra savaip išskirtiniai: tai tarsi dvi priešingos atsitiktinių įvykių ribos, priešpriešiniai jų poliai. ***Atsitiktiniais*** ir vadinami visi tokie įvykiai, kurie už negalimuosius „labiau galimi“, bet „mažiau galimi“ už būtinuosius.

3. Tikimybės supratimas

Tikimybė yra atsitiktinio įvykio *galimybės mastas* (laipsnis, didumas, šansas ir pan.). Kai sakome, jog atsitiktinis įvykis gali įvykti ar neįvykti, kartu postuluojame tam tikrą kiekvienos iš šių alternatyvų *galimybę*, potencialumą, o tos galimybės *didumas* (lauktinumas, pasikliautinumas, „intensyvumas“ ar pan.) ir yra atitinkamos alternatyvos tikimybė. *Tikimybę*, t. y. atsitiktinio įvykio galimybės mastą, galima *nusakyti žodžiais*, apibūdinti aprašomuoju būdu – pasitelkiant atitinkamos reikšmės „įprastinius“ pasakymus ir galima *išreikšti* specialiais jos *matavimo vienetais*, kurie, kaip ir visi matavimo vienetai, yra iš esmės susitarimo dalykas (o ne, sakysim, pačios gamtos padiktuoti). Žodinis tikimybės apibūdinimas yra paplitęs buitinėje, kasdieninėje kalboje, o jos raiška specialiai tam skirtais mato vienetais – tiksliuosiuose moksluose.

Tikimybės matas yra parinktas atsižvelgiant į jau minėtas kraštines atsitiktinių įvykių ribas – į *negalimuosius* ir *būtinuosius* įvykius. Natūralu, kad *negalimo įvykio tikimybė* prilyginama *nuliui* („įvertinama“ nuliu). *Būtinojo* gi *įvykio tikimybė* prilyginama arba *vienetui*, arba *šimtui*. Tuomet bet kurio „tarpinio“ atsitiktinio įvykio – t.y. galinčio arba įvykti, arba neįvykti – galimybė įvertinama arba *intervalo* [0;1] arba [0; 100] skaičiumi: to skaičiaus dydis tepriklauso tik nuo pačios galimybės „didumo“ (masto) ir tiesiogiai jį išreiškia. Pirmu atveju sakoma, kad tikimybė yra reiškiama ***vieneto dalimis***, antru – jog ji reiškiama ***procentais*** (nuošimčiais). Dažnesnis ir labiau įprastas yra tikimybės reiškimas vieneto dalimis. Natūralu ir savaime suprantama, jog tarp tų dviejų tikimybės reiškimo „sistemų“ esama vienareikšmio tiesioginio ryšio: padauginę vieneto dalimis išreikštą tikimybę iš 100, gausime ją išreikštą procentais. Tad tikimybės reiškimas vieneto dalimis ir procentais yra visiškai ekvivalentūs ir vienodai tikslūs, o jos apibūdinimas žodžiais – mažiau preciziškas, bet užtat artimesnis įprastai kalbėsenai. Palyginkime:

Tikimybė:	Apibūdinta žodžiais	Išreikšta vieneto dalimis	Išreikšta procentais
	jokiu būdu	*0*	*0%*
	beveik neįmanoma	*0,0001*	*0,01%*
	sunku patikėti	*0,01*	*1%*
	menkai tikėtinas dalykas	*0,05*	*5%*
	gal ir galimas daiktas	*0,1*	*10%*
	šiaip įmanoma	*0,15*	*15%*
	sunku pasakyti	*0,25*	*25%*
	gali būti	*0,5*	*50%*
	lauktume, kad	*0,75*	*75%*
	visai gali būti	*0,85*	*85%*
	greičiausiai	*0,9*	*90%*
	nėra abejonės	*0,95*	*95%*
	tikrų tikriausia, kad	*0,99*	*99%*
		0,9999	*99,99%*
		1	*100%*

Čia pridurtina, kad abiejų skaičių stulpeliai eilutėse yra užrašytas *tiksliai tas pats* tikimybės dydis, o žodiniai apibūdinimai jį teperteikia tik gerokai *apytiksliai* (pasakytum, „maždaug“, „į tą pusę“ ir pan.).

5. Kaip įvertinti atsitiktinio įvykio (ar dydžio) tikimybę?

Tai vienas iš pačių esmingiausių klausimų, kurį aiškinsimės per visus mūsų užsiėmimus ir teišsiaiškinsime tik nedidelę visų jo niuansų dalelę. Bėda ta, kad patys atsitiktiniai įvykiai yra labai įvairūs, nevienodi, dažnai – net nepalyginamos „vidinės“ struktūros, todėl vieno kokio ar ir kelių „universalių“ tikimybės „išmatavimo būdų“ praktiniams reikalams dažniausiai nepakanka. Pats pagrindinis dalykas tikimybinių metodų „taikytojui“ todėl ir yra labai gerai išmanyti bei mokėti išsiaiškinti tų konkrečių atsitiktinių įvykių ar dydžių, su kuriais jis dirba, struktūrą ir pagal tai parinkti kuo labiau tinkamus jų tikimybių nustatymo („matavimo“) būdus. O tam skirtų aparatų (fizinių instrumentų) – nėra iš viso.

Apskaičiuoti „sudėtinių“ įvykių tikimybes nebūtų sunku, jei žinotume juos sudarančių *elemetariųjų* įvykių tikimybes: pagelbėtų Būlio algebra. O elementariųjų atsitiktinių įvykių tikimybes įvertinti yra lengviausia tuomet, kai visų jų, sudarančių pilną elementariųjų įvykių erdvę, *galimybės yra vienodos*, kitais žodžiais tariant, kai bandymo (ar “sudėtinio” įvykio) visos įmanomos baigmės vienodai galimos. Tačiau taip tebūna tik tais palyginti retais atvejais, kuomet bandymų įtaisas (*atsitiktinių įvykių* šaltinis arba, kaip labiau įprasta sakyti – *generatorius*) pasižymi *simetriškumu*: jei metama *nesulankstyta* kapeika, galimybės atsiversti ir vienai, ir kitai jos pusei yra visiškai vienodos; vienodos yra galimybės atsiversti bet kuriam taškų skaičiui nuo 1 iki 6, kai metamas *taisyklingas* šešiabriaunis žaidimų kauliukas; kai iš *gerai išmaišytos* kortų kaladės traukiama korta, galimybės ištraukti bet kurią iš jų irgi yra vienodos ir t. t.

Be abejo, logiška ir natūralu yra tokiais atvejais visą lygias galimybes turinčių elementariųjų atsitiktinių įvykių erdvę projektuoti į visą skaičių intervalą, skirtą tikimybei reikšti, ir tą intervalą elementariesiems įvykiams *padalinti po lygiai*, lygiomis dalimis (juk jų galimybės čia vienodos!), kitaip tariant – į tiek dalių, kiek erdvėje yra elementariųjų įvykių: tuomet visų jų tikimybės bus tokios pačios, vienodos:

Bandymas	Elementariųjų įvykių kiekis erdvėje	Kiekvieno iš elementariųjų įvykių tikimybė
Kapeikos mėtymas	2	1/2 = 0,5
Kauliuko mėtymas	6	1/6 = 0,16666...
Kortos traukimas iš 36 korų kaladės	36	1/36 = 0,027777...
"Sportloto" kamuoliuko iškritimas iš lototrono	49	1/49 = 0,0204
Egzamino bilieto traukimas (64 bilietai)	64	1/64 = 0,015625

Tai *klasikinė*, iš 17–18 a. atėjusi *tikimybės samprata*, tinkanti labiausiai žaidimams ar dirbtiniams bandymams, kur lygias elemtariųjų įvykių galimybes užtikrina akivaizdi, dirbtinai kuriama bandymų (žaidimų) įrenginio būsenų ***simetrija***. O pats toks tikimybės radimo būdas vadinamas *tiesioginiu tikimybės apskaičiavimu*. Tačiau – dar kartą pabrėžtina – jis tetinka tik tiems atistiktiniams įvykiams, kuriems būdinga simetrija. O praktikoje, ypač – kalbos tyrinėjimuose tokios galimų variantų simetrijos beveik niekuomet nebūna, tad ir tiesioginis tikimybių apskaičiavimo būdas čia netinka. Sakysim, pažvelgus į anksčiau pateiktą c) pavyzdį (žr. 2 psl.), nesunku suprasti, jog 1, 2, 3 ir t. t. elementariųjų įvykių galimybės iš tikrųjų *nėra vienodos*, jų vienodumo neužtikrina jokios išankstinės aplinkybės ar sąlygos, tad ir jų tikimybių *laikyti vienodomis* bei apskaičiuoti tiesiogiai – nėra jokio pamato. Kaip tik atvirkščiai – jos priklauso nuo to, kaip dažnai kalboje pasitaiko daiktavardžių, būdvardžių, skaitvardžių, įvardžių ir t. t. , o skirtingos kalbos dalys *nevienodai* dažnos. Tokiais ir panašiais atvejais (o praktiškai jie ir yra patys dažniausi!) belieka mėginti tikimybes nustatyti netiesiogiai. Ir čia labai pagelbsti dėmesys atsitiktinio įvykio *dažnumui*.

Atsitiktinių įvykių modeliavimas urnų schemomis

Teoretikai (o kai kada ir praktikai, „taikytojai“) atsitiktiniams įvykiams modeliuoti neretai pasitelkia vadinamąsias *urnų schemas* – arba tikras, fizines, arba – kur kas dažniau – įsivaizduojamas, virtualias *urnas*, t. y. dėžes ar puodus, į kuriuos dedama tam tikras kiekis visiškai *vienodo dydžio bei formos*, bet *įvairių spalvų* rutuliukų. Uždengtoje urnoje rutuliukai gerai išmaišomi ir nežiūrint traukiamas vienas iš jų. Taip imituojamas (modeliuojamas) *elementarusis atsitiktinis įvykis*. Po to rutuliukas metamas atgal į urną, vėl gerai išmaišoma ir bandymas kartojamas tiek, kiek reikia.Kai kada rutuliukai būna vienos spalvos, bet *numeruoti*, ir traukiant žiūrima jų numerių.

Visų *simetriškųjų* bandymų (kapeikos mėtymo, kauliuko mėtymo ir pan.) imitacijoms būtina sąlyga yra, kad skritingų spalvų rutuliukų būtų *po vienodą kiekį*, o spalvų skaičius atitiktų skirtingų elemntariųjų įvykių (kitaip – galimų baigmių) skaičių. Jei urnoje yra, sakysim, 20 juodų ir 20 baltų rutuliukų, traukdami rutuliuką ir žiūrėdami jo spalvos, puikiai imituosime kapeikos mėtymą. Norėdami imituoti kauliuko mėtymą, turėsime į urną įdėti po, tarkime, 10 (ar po kokį kitą *vienodą* skaičių, nesvarbu kokį konkrečiai) raudonų, žalių, geltonų, baltų, mėlynų ir juodų rutuliukų: tuomet vienos kurios iš šių šešių spalvų rutuliuko ištraukimas atitiks vieno kurio taškų skaičiaus iškritimą. Aišku, jog spalvos irgi gali būti bet kokios, svarbu tik, kad jų būtų *šešios skirtingos*.

Ir visai kitaip būtų, jei rutuliukais bandytume imituoti kalbos dalis, randamas sakinio pradžioje. Čia mes težinome tik, kad įsivaizduojamoje urnoje yra *11 skirtingų spalvų* rutuliukų, bet *po kiek* katros spalvos jų ten įdėta – nežinome. Ir dar – svarbu, kas mums rūpi: ar ribotas, baigtinis sakinių kiekis, pavyzdžiui, tik konkretaus teksto sakiniai (tuomet ir *bendras* rutuliukų *kiekis* įsivaizduojamoje urnoje turėtų būti baigtinis, toks pat, kaip ir sakinių kiekis), ar sakiniai “apskritai kalboje” (antru atveju urną derėtų vaizduotis esant begalinio dydžio ir begalybę rutuliukų joje).

Modernesnė urnų schemos atmaina yra *lototronas* – tokia speciali mašinėlė, kuri pati išmaišo jos indan sudėtus rutuliukus ir po to atsitiktinai išmeta vieną iš jų. Savaip giminingos urnoms yra ir įvairios *ruletės* („laimės“ ratai).

6. Įvykių dažnumas. Absoliutus ir santykinis dažnumas

Įvykio *dažnumas* – tai jo *kartojimosi intensyvumas*: apie intensyviai besikartojančius įvykius sakoma, kad jie yra dažni, o apie retkarčiais pasirodančius – reti. Dažnumas – labai svarbi statistikos kategorija, tarsi dar vienas įvykių matmuo, leidžiantis juos, įvykius, *palyginti*, be visų kitų galimų atžvilgių, dar ir šiuo specifiniu požiūriu: kurie karojasi daugiau, o kurie – mažiau.

Skiriamos dvi dažnumo „atmainos“: ***absoliutus*** (kartais pasakoma absoliutinis) *dažnumas* (dažnis) ir ***santykinis*** *dažnumas* (dažnis).

Absoliutus dažnumas	Santykinis dažnumas
Skaičius, nurodantis, kiek iš viso kartų įvyko (pasikartojo) tiriamasis įvykis (jį įprasta žymėti *m*), kai bandymas buvo kartojamas kažkiek (sakysim, *n*) kartų.	Santykis, rodantis, kokioje visų atliktų bandymų dalyje įvyko (pasikartojo) tiriamasis įvykis. Nustatomas absoliutų dažnumą *m* dalinant iš visų bandymų skaičiaus *n* (*m/n*)

Pavyzdžiui, tiriame, kaip dažnai kartojasi triskiemeniai žodiai

Bandymas: imame žodį ir žiūrime, ar jis triskiemenis. Jei *taip* – tiriamasis įvykis *įvyko*, jei *ne* – *neįvyko*. Ištyrėme 2 tekstus (atlikome 2 bandymų serijas): pirmajame tekste iš viso rasta žodžių (*n*=): 415; triskiemeniai iš jų buvo (*m*=): 94, o antrajame iš viso rasta žodžių 712; tarp jų triskiemenių 153. Tuomet:

	Absoliutus dažnumas	Santykinis dažnumas
1 tekstas:	94	94/415 = 0.2265
2 tekstas:	153	153/712 = 0.2149

Nesunku suprasti, kad:

Palyginti *absoliučius* dažnumus logiškai įmanoma tik tada, kai atliekamos dvi ar kelios bandymų serijos, ir atliktų *bandymų skaičiai* (kiekiai) jose yra *vienodi*.

Palyginti *santykinius* dažnumus logiškai įmanoma *visada*, nes išreikšdami tiriamojo įvykio pasikartojimų *santykį* su bendru bandymų skaičiumi, jie visuomet yra bendramačiai.

Dėl tos priežasties santykinis dažnumas statistikoje yra labiau populiarus, ir dažniausiai remiamasi kaip tik juo. Tačiau tais atvejais, kai yra užtikrinamas bandymų kiekio vienodumas (pavyzdžiui, apsispręstume iš kiekvieno teksto imti, sakysim, tik po 350 ar po kitokį vienodą skaičių žodžių), visiškai galima palyginimams naudoti ir absoliutų dažnumą.

7. Santykinio dažnumo raiška, jo santykis su tikimybe

Kadangi santykinis dažnumas yra ne kas kita kaip santykis (m/n), tai jį irgi natūralu „matuoti“ arba vieneto dalimis, arba procentais, t. y. tais pačiais vienetais, kuriais matuojama tikimybė. Dar daugiau: jei tiriamasis įvykis ***būtinai*** įvyksta kiekvieno bandymo metu, kiekvienąkart (sakysim, *visi* rutuliukai urnoje yra balti, ir mums rūpi įvykio *ištrauktas baltas rutuliukas* dažnumas), tai pasidaro $m=n$ ir santykinis tokio įvykio dažnumas tampa, suprantama, lygus vienetui arba 100% (nes $m/n=n/n=1$). Ir atvirkščiai: jei tiriamasis įvykis pasidaro ***iš principo negalimas*** (tarkime, ir toliau rūpi įvykio *ištrauktas baltas rutuliukas* dažnumas, bet urnoje tėra *vien tik juodi* rutuliukai), tai m tampa lygus nuliui, o tuo pačiu ir tokio įvykio santykinis dažnumas prilygsta nuliui (nes $0/n=0$). Taigi, santykinis dažnumas labai jau primena tikimybę: būtino įvykio atveju jis lygus vienetui, negalimo – nuliui, o galinčio įvykti ar neįvykti įvykio – svyruoja tarp 0 ir 1.

Taip yra dėl to, kad:

> Atsitiktinio įvykio *santykinis dažnumas* yra jo *tikimybės empirinis pasireiškimas* konkrečioje situacijoje, konkrečioje bandymų serijoje, tarsi jos koks praktinis atitikmuo, analogas, *viena iš* daugybės galimų tikimybės *realizacijų.*

Taigi, tikimybė priklauso labiau idealybės sferai, abstraktybei, teorijai, "grynajai" matematikai, o santykinis dažnumas – labiau realybės dalykams, konkretiems bandymams ar įvykių srautams, jų serijoms, statistikai.

Literatūroje tikimybė ir santykinis dažnumas net ir ženklinami (žymimi) labai panašiai: didžiąja P arba mažąja p – matyt, iš angliško žodžio *probibility.* Kartai, norėdami akivaizdžiai parodyti, kur P ar p reiškia tikimybę, o kur – santykinį dažnumą, pastarąjį dar žymi P^* arba p^*.

8.Tikimybės įvertinimas santykiniu dažnumu. Statistinė tikimybės samprata.

Tais atvejais, kai tikimybę *apskaičiuoti tiesiogiai* anksčiau minėtu būdu negalima (pirmiausiai – dėl to, jog iš anksto yra aišku, kad bandymas nepasižymi simetriškumu ir visų jo potencialių baigmių galimybės ***nėra*** vienodos), tenka vietoj pačios *tikimybės* imti jos *empirinį analogą*, t. y. *santykinį dažnumą* ir jį laikyti bent jau apytikre tikimybės reikšme, jos daugiau ar mažiau patikimu *įvertinimu*.

Grįždami prie urnų schemos, galime tarti, jog žinome tik, kad įvairių spalvų rutuliukų jon įdėta ***ne po lygiai***, bet kokios konkrečios jų proporcijos ten yra – nežinia. Mums rūpi tik vienos kurios spalvos rutuliukai. Kadangi nežinome, kokį visų urnoje esančių rutuliukų procentą (dalį) šios spalvos rutuliukai sudaro, turime atsitiktinai traukti tam tikrą kiekį (kelias dešimtis ar šimtus) rutuliukų ir žiūrėti, kiek tarp jų pasitaikys rūpimos spalvos.

Labai svarbus klausimas: ar empiriniu būdu nustatytas santykinis dažnumas (o dažniausiai jis nustatomas iš palyginti riboto bandymų skaičiaus) tikimybę atspinti (įvertina) *pakankamai tiksliai* ir *kokia* gali būti to įvertinimo (prielaidos, kad ir pati tikimybė esanti maždaug tokio pat dydžio, kaip ir nustatytasis santykinis dažnumas) *paklaida*? Gaustai atsakyti galima taip: kuo iš *didesnio bandymų kiekio* santykinis dažnumas apskaičiuojamas, tuo, apskritai imant, tiksliau jis gali įvertinti (reprezentuoti) nežinomą tikimybę.

Yra nustatyta, kad kai išlaikomos dvi sąlygos:

- tiriamieji įvykiai tikrai yra atsitiktinio pobūdžio
- pati tikriamojo įvykio tikimybė kartojant bandymus netinta, išlieka visą laiką vienoda

tai santykinis dažnumas pasižymi vadinamąja *stochastinio stabilumo* savybe – daug kartų kartojat bandymų serijas, jis atsitiktinai svyruoja apie tiksliai mums nežinomą tikimybės reikšmę: vienąkart būna už ją kiek didesnis, kitąkart galbūt mažesnis, ir iš kuo gausesnės bandymų serijos jis nustatomas, tuo to svyravimo amplitudė mažesnė. Kai bandymų skaičius serijoje neribotai didėja, bendra šių svyravimų amplitudė neribotai mažėja. Tad:

> Statistiškai suprantama *tikimybė* ir yra tasai *dydis*, dažniausiai – iš anksto tiksliai nė nežinomas, *apie kurį atsitiktinai svyruoja* atsitiktinio įvykio *santykinis dažnumas*, ir – su tuo mažiau tikėtina didele paklaida (nuokrypiu), iš kuo didesnio bandymų skaičiaus jis apskaičiuotas.

Tikslia matematikos kalba tai galima būtų užrašyti tokia formule:

$$\boxed{P_{\{\,/\,p\,-\,p^*/\,<\,e\}} \rightarrow 1, \text{ kai } n \rightarrow \infty}$$

Žodžiais tariant – tikimybė įvykio, kad tikimybės ir santykinio dažnumo skirtumas absoliutiniu didumu neviršys bet kokio labai mažo dydžio e, artėja į vienetą, kai bandymų skaičius artėja į begalybę, t. y. toks įvykis, bandymų skaičiui artėjant į begalybę, darosi vis labiau *praktiškai būtinas*.

Jei atliekama daug bandymų serijų, iš kurių kiekvienos apskaičiuojama vis kiek kitokia konkreti santykinio dažnumo reikšmė, galima sakyti, kad statistinė tikimynė yra tas dydis, apie kurį šios reikšmės telkiasi (koncentruojasi), ir – tuo glaudžiau, kuuo didesni bandymų kiekiai serijose.

9.Atsitiktinių įvykių generatoriai. Atsitiktiniai skaičiai

Generatorių, generuojančių dirbtinius atsitiktinius įvykius, pavyzdžių jau buvo paminėta: tai įvairūs žaidimų kauliukai, kortos, urnos su rutuliukais, lototronas; dar galima būtų pridėti ruletę ir greitaeigius elektroninius skaitiklius. Visi jie turi savų ypatumų.

Parankus “atsitiktinybės šaltinis”, su kuriuo neretai susiduria ir į tikslumą linkstantys lingvistai, yra atsitiktiniai skaičiai. Jie parankūs tuo, kad nereikia nei turėti kokių nors specialių įtaisų, nei pačiam atlikti bandymų: jų lentelės pateikiamos labai dažname matematinės statistikos ar tikimybių teorijos vadovėlyje; be to, skaičių, labai panašių į atsitiktinius (iš tikrųjų – kvaziatsitiktinių arba pseudoatsitiktinių) lengvai galima prisigeneruoti tiek, kiek reikia, su kompiuteriu. Mes irgi jais naudosimės palyginti dažnai. Atsitiktiniai skaičiai patogūs dar ir tuo, kad jais palyginti lengvai galima modeliuoti labai įvairias atsitiktinybės formas (vadinamuosius tikimybinius pasiskirstymus). Labiausiai paplitę yra *tolygiai pasiskirstę* (kokiame nors intervale) atsitiktiniai skaičiai: vadinasi, bet kuris to intervalo skaičius bet kurioje pozicijoje gali pasirodyti su vienoda tikimybe.

Tikrus (ne - kvazi...) *tolygiai pasiskirsčiusius* atsitiktinius skaičius patogu „generuoti“ irgi panaudojant urną: užtenka į ją įdėti 10 rutuliukų, sunumeruotų nuo 0 iki 9, ir atsitiktinai traukti po vieną rutuliuką, kuris po to vėl grąžinamas atgal. Ištrauktų rutuliukų skaitmenis grupuodami po keturis gautume atsitiktinius *sveikuosius* skaičius, tolygiai pasiskirsčiusius intervale [0; 9999]; jei tuos skaitmenis užrašytume kitokio dydžio grupėmis, intervalai būtų, aišku, atitinkamai kitokie.

Tolygiai pasiskirsčiusiems atsitiktiniams skaičiams generuoti taip pat puikiai tinka ir elektroninis greitaeigis dešimtainis vienaskiltis skaitiklis, kuriame, paspaudus mygtuką, didžiuliu greičiu ima keistis skaitmenys nuo 0 iki 9 ir nustoja keistis po to, kai paspaudžiame antrą mygtuką. Kadangi laiko tarpas tarp dviejų mygtukų paspaudimų yra, lyginant su skaitiklio “apsukomis”, atsitiktinis dydis, tai tiktimybė, kad skaitiklis sustos, kai bus rodomas tas ar kitas skaitmuo, yra visuomet vienoda ir lygi 0,1.

Dažnai yra pateikiami atsitiktiniai skaičiai, tolygiai pasiskirstę tame pačiame intervale, kuriame “įsitenka” atsitiktinių įvykių tikimybės bei jų santykiniai dažnumai, t. y. intervale [0; 1]. Norėdami gauti juos, pirma turėtume generuoti atsitiktinius skaičius, tolygiai pasiskirsčiusius kokiame nors palyginti plačiame intervale (tarkime [0; 999999]), ir iš kiekvieno tokio skaičiaus dalinti vienetą.

Vėliau susidursime ne tik su *tolygiai*, bet ir su *kitaip pasiskirsčiusiais* atsitiktiniais skaičiais.

10. Truputėlis tikimybių (ir – santykinių dažnumų) "algebros"

* Jeigu kokio nors įvykio (pavadinkime jį įvykiu *A*) tikimybė yra *p,* tai jam priešingo (negacija !) įvykio (*not A*) tikimybė bus lygi (1 - *p*). Savaime suprantama.

* Jei įvykio *A* tikimybė yra $p_{A,}$ o įvykio *B* – p_B, ir *A* ir *B* yra nesutaikomi įvykiai (t. y. negali abu įvykti vieno bandymo metu) tai įvykio (*A xor B*) tikimybė bus lygi ($p_A + p_B$). Disjunkcija – loginė sudėtis!

* Jei įvykiai *A* ir *B* yra nepriklausomi (t. y. vieno iš jų tikimybė niekaip nepriklauso nuo to, įvyko ar neįvyko kitas), o jų tikimybės atitinkamai lygios p_A ir p_B, tai įvykio (*A and B*) tikimybė bus lygi (p_A x p_B). Konjunkcija – loginė daugyba!

Kiek kebliau yra su *sutaikomų* atsitiktinių įvykių disjunkcija (natūralu, jog tai būna *paprastoji* arba *silpnoji* disjunkcija, *A or B*) ir *priklausomų* atsitiktinių įvykių konjunkcija: formulės bendrai sudėtinių įvykių tikimybei apskaičiuoti tampa šiek tiek sudėtingesnės.

Pateiktosios formulės tinka ne tik dvinarėms, bet ir ***daugianarėms*** disjunkcijoms bei konjunkcijoms (mat, tos loginės jungtys iš principo gali sieti ne tiktai du, bet *ir daugiau* narių).