STATISTINIŲ KRITERIJŲ TAIKYMAS
NEPARAMETRINIAI KRITERIJAI

Įsidėmėtina: Statistiniai (parametrų) palyginimai – tai ne „aritmetiniai“ fiksuotų (griežtai apibrėžtų, nesikeičiančių) dydžių palyginimai, bet „specializuoti“ varijuojančių (fluktuojančių, *atsitiktinai* besikaitaliojančių) dydžių (parametrų) palyginimai, ir jų tikslas – nustatyti, kiek yra *tikėtina*, kad tyrimo duomenyse „pasireiškiantys“ tų dydžių nesutapimai galį būti sąlygoti atsitiktinumo, atsitiktinybės veikimo.

I. PARAMETRINIŲ KRITERIJŲ TAIKYMAS (TĘSINYS)

Prielaidos apie proporcijų lygybę tikrinimas

Proporcijomis čia vadinamos imties *dalys*, tenkančios rūpimą požymio reikšmę turintiems objektams. Pvz., jeigu imtin iš viso atrinkta 325 žodžiai ir 97 iš jų yra daiktavardžiai, tai daiktavardžių proporcija šioje imtyje yra 97 / 325 = 0,29846... . Taigi, proporcija iš esmės yra tas pat, kas rūpimos požymio reikšmės *santykinis dažnumas*. Iš principo galima apskaičiuoti ir kokybinių, ir ranginių, ir netgi kiekybinių diskrečiųjų požymių reikšmių – ar visų galimų, ar tik kai kurių iš jų, rūpimųjų – proporcijas (santykinius dažnumus), tačiau kokybinių ir, iš dalies, ranginių požymių atveju proporcijos tampa itin svarbios dėl to, kad tuomet jos pasidaro bene vieninteliai empirinį reikšmių skirstinį apibūdinantys parametrai, nes tokių įprastų parametrų kaip vidurkis, dispersija ir pan. apskaičiuoti tuomet neįmanoma. O kadangi filologai su kokybiniais tiriamųjų objektų požymiais susiduria itin dažnai, tai jiems kaip tik proporcijų statistinio palyginimo uždaviniai yra labai aktualūs.

Aptariant proporcijų palyginimo „mechanizmus“ statistikos knygose paprastai yra orientuojamasi į imtis, fiksuojančias *binariškų* kokybinių požymių reikšmes (pvz., anketą užpildžiusių asmenų lytį ar pan.), tačiau tie patys „mechanizmai“ iš principo tinka ir tada, kai tie požymiai „iš prigimties“ yra daugiareikšmiai, o imtys „binarizuojamos“ kiek dirbtinai – priešpriešinant rūpimąją reikšmę su visomis kitomis, likusiomis (pvz., jeigu turima iš teksto paimta tam tikro dydžio vardažodžių imtis, kurioje užfiksuota, kuriuo linksniu kuris vardažodis pavartotas, tai į ją galima žiūrėti eilės „binarizuotų“ požymių – vardininkas / nevardininkas, kilmininkas /nekilmininkas ir t.t. – imčių „rinkinį“). Svarbu tik tai, kad rūpimoji požymio reikšmė yra būdinga tam tikrai (žinomai) imties objektų daliai ir nebūdinga visiems kitiems imties objektams, sudarantiems likusią jos dalį.

Kadangi proporcijos yra ne kas kita kaip santykiniai dažnumai, tai daugelyje statistikos knygų jos žymimos taip pat kaip ir tikimybės – raide p (pvz., p_1, p_2, p_x, p_k ir t.t.). Šiame šio konspekto skyrelyje jas bus stengiamasi žymėti „mnemoniškiau“ – raide d (pvz., d_1, d_2, d_x, d_k ir t.t.), t.y. žodžio (santykinis) *dažnumas* pirmąja raide. Proporcija yra „aritmetinis“ *santykis*: objektų, turinčių rūpimąją požymio reikšmę, skaičiaus santykis su visų imties objektų skaičiumi: jei imtyje iš viso yra n objektų, ir k iš jų turi rūpimąją požymio reikšmę, tai tos reikšmės santykinis dažnumas arba proporcija $d = k / n$.

Proporcijų statistinio palyginimo mechanizmai teoretikų yra sukonstruoti remiantis tuo, kad rūpimąją požymio reikšmę turinčių objektų kiekis k imtyje, kurioje iš viso yra n objektų, yra atsitiktinis dydis, turintis binominį skirstinį, apibūdinamą parametrais n (bandymų kiekis = imties dydis) ir p (tikimybė, kad pavienis, atskirai paimtas populiacijos, iš kurios imta imtis, objektas turės rūpimąją požymio reikšmę; ši tikimybė paprastai nežinoma, išskyrus tuos retus atvejus, kai ją galima kaip nors apskaičiuoti – tarkime, remiantis požymio reikšmių „generatoriaus“ simetriškumu ar pan.). Tad į $d = k / n$ reikia žiūrėti kaip į k binominį skirstinį „valdančios“ nežinomos tikimybės p **empirinį įvertį**. O palyginimo esmė tokia: turime dvi imtis – vienoje iš viso yra n_1 objektų, ir k_1 iš jų turi rūpimąjį požymį, o kitoje iš viso yra n_2 objektų, ir tą patį rūpimąjį požymį iš jų turi k_2 objektų. O klausimas gi – apie to rūpimojo požymio *tikimybę populiacijose*, iš kurių šios imtys paimtos: ar esama patikimo pamato laikyti, kad šiuos konkrečius kiekius k_1 ir k_2 apsprendusi tikimybė yra tokia pati, vienoda (tuomet imtyse stebimas $d_1 = k_1 / n_1$ ir $d_2 = k_2 / n_2$ nesutapimas būtų laikomas neesminiu, atsitiktiniu dalyku, ir proporcijos d_1 bei d_2 būtų traktuojamos kaip statistiškai lygios, „esmingai nesiskiriančios“), ar atvirkščiai – tas pamatas per menkai patikimas, ir dėl to laikytina, kad kiekius k_1 ir k_2 apsprendusi tikimybė esanti skirtinga (tuomet ir proporcijos d_1 ir d_2 būtų traktuojamos kaip nelygios, besiskiriančios iš esmės).

Skirami du proporcijų palyginimo situacijų atvejai (tipai): vieną jų turime tada, kai k_1 ir k_2 yra palyginti nemaži skaičiai (kelių dešimčių ar dar didesnio dydžio), ir dėl to yra leistina k_1 bei k_2 binominio skirstinio *aproksimacija* normaliuoju skirstiniu, o antrą – kai k_1 ir k_2 yra palyginti maži skaičiai, ir jų skirstinio aproksimacija normaliuoju skirstiniu būtų per grubi. Labiau svarbus pirmasis atvejis, nes jis praktiškai kur kas dažnesnis, todėl apie jį – pirmiausiai.

A. Dvi imtys ir nemaži k_1 bei k_2

Situacija. Turimos dvi nepriklausomos imtys, jų dydžiai yra n_1 ir n_2. Pirmojoje imtyje rasta k_1, o antrojoje – k_2 objektų, turinčių rūpimąją požymio reikšmę.

Statistinės prielaidos. Nulinė hipotezė teigtų, kad rūpimąją požymio reikšmę turinčių objektų proporcijos populiacijose, iš kurių imtys paimtos, yra tokios pačios, vienodos, tad vienodos ir tikimybės, pavienis atsitiktinai paimtas objektas turės rūpimąjį požymį. Alternatyva tvirtina, kad tos tikimybės nelygios, nevienodos. Užrašius glaustai:

$$H_0:\ p_1 = p_2$$
$$H_1:\ p_1 \neq p_2$$

Kriterijaus statistika (paprastai žymima raidėmis z arba u) įvairiose statistikos knygose ganėtinai įvairuoja, pateikiama skirtingų jos apskaičiavimo būdų ir formulių; patys teoretikai dar iki galo nesutaria, kuris KS apskaičiavimo būdas esąs geriausias. Tačiau visais atvejais KS konstruojama taip, kad H_0 esant teisingai ji būtų *standartinį normalųjį* skirstinį turintis atsitiktinis dydis. Čia nurodomas KS = u apskaičiavimo būdas, pateiktas V. Čekanavičiaus ir G. Murausko knygoje „Statistika ir jos taikymai“ (I tomas, p. 189):

$$u = \frac{d_1 - d_2}{\sqrt{\left(\frac{k_1+k_2}{n_1+n_2}\right)\cdot\left(1-\frac{k_1+k_2}{n_1+n_1}\right)\cdot\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$$

Jeigu pažymėtume $d = (k_1 + k_2) / (n_1 + n_2)$, tai formulės vaizdas tarsi supaprastėtų:

$$u = \frac{d_1 - d_2}{\sqrt{d\cdot(1-d)\cdot\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$$

Mūsų filologų darbuose itin paplitęs yra kitas kriterijaus statistikos u apskaičiavimo būdas, A. Pakerio kažkada „išskaitytas“ iš biologų darbų, iš bene septintajame XX a. dešimtmetyje išėjusios tokio Urbacho knygos „Biometričesskije metody“. Ten siūloma KS = u skaičiuoti taip:

$$u = \left(2\arcsin\sqrt{d_1} - 2\arcsin\sqrt{d_2}\right)\sqrt{\frac{n_1\cdot n_2}{n_1+n_2}}$$

Abiem šiais būdais apskaičiuota u reikšmė konkretiems k_1, n_1, k_2 ir n_2 paprastai tesiskiria nežymiai, ir sunku pasakyti, kuri iš tų aproksimacijų yra tikslesnė.

Prielaidų vertinimas. Kai alternatyva dvipusė (H_1: $p_1 \neq p_2$), gautąją u reikšmę atitinkanti p-reikšmė su *Excel* apskaičiuojama taip:

p = 2 * (1–NORMSDIST(ABS(u)))

p-reikšmė šiaipjau reiškia tikimybę rizikos, kad atmetant H_0 bus padaryta pirmos rūšies klaida, todėl H_0 pagrįstai atmesti galima tik tada, kaip p-reikšmė gaunama nedidelė, nežymi, mažesnė už įprastinius, tradininius reikšmingumo lygmenis (0,1; 0, 05; 0,01 ar 0,001). Kita gi vertus p-reikšmė išreiškia ir pačios H_0 tikėtinumą (priimtinumą), t.y. tikimybę, kad joje išsakytas teiginys atitinka tikrovę. Todėl kuo didesnė *p*-reikšmė, tuo labiau nulinė hipotezė pasikliautina.

Jeigu norimas reikšmingumo lygmuo α pasirenkamas iš anksto, tai jį atitinkančios kritinės reikšmės randamos taip:

KR(kair.) = NORMSINV(α/2)
KR(deš.) = NORMSINV(1–α/2)

Jas nustačius žiūrima, ar gautoji *u* reikšmė su jomis prasilenkia, ar ne: jei prasilenkia, tai nulinė hipotezė atmetama su reikšmingumo lygmeniu α, o jei jose išsitenka – neatmetama.

Pavyzdys. Iš vieno eilėraščių rinkinio paimta 412 asmenuojamųjų veiksmažodžio formų imtis, ir 163 iš jų pasirodė beesančios vienaskaitos pirmasis asmuo („ [aš].... “). Iš kito rinkinio asmenuojamųjų formų iš viso paimta 675, ir vienaskaitos pirmasis asmuo pasitaikė 251 kartą. Ar priimtina prielaida, kad „poetinio solipsizmo“ laipsnis abiejuose rinkiniuose vienodas, t.y. kad ir viename, ir kitame rinkinyje *vienoda* yra *tikimybė* atsitiktinai paėmus asmenuojamąją veiksmažodžio formą rasti ją esant „ [aš] ... “ – vienaskaitos pirmuoju asmeniu?

Vartojant įprastus žymenis, pavyzdžio situaciją reikėtų užfiksuoti taip:

$k_1 = 163$; $n_1 = 412$
$k_2 = 251$; $n_2 = 675$

H_0: $p_1 = p_2$
H_1: $p_1 \neq p_2$

Tuomet kriterijaus statistikos reikšmė būtų $u = 0{,}783258$ (jeigu skaičiuojama pagal knygoje „Biometričesskije metody“ randamą formulę, tai $u = 0{,}782228$). O ją atitinkanti *p*-reikšmė būtų lygi 0,433475. Pirmos rūšies klaidai tai labai didelė tikimybė, todėl nulinės hipotezės nėra pagrindo atmesti, o joje suformuluoto teiginio tikėtinumas yra apie 43,3%. Kitais žodžiais tariant, kone kas antru atveju nulinės hipotezės teiginys iš tikrųjų turėtų pasitvirtinti, būti teisingas. Todėl pavyzdžio duomenys *neįrodo* (su nedidele klaidos rizika), kad „poetinio solipsizmo“ laipsnis šiuose rinkiniuose būtų iš esmės skirtingas.

Požymio reikšmės proporcijų palyginimo dviejose imtyse uždaviniai savo esme šiaipjau yra *homogeniškumo patikrinimo* uždaviniai: juk čia keliamas klausimas, ar priimtina prielaida, jog tiriamosios požymio reikšmės atžvilgiu imtys yra homogeniškos, jog populiacijose, iš kurių tos imtys paimtos, ši reikšmė yra pasklidusi vienodomis proprocijomis, tokiu pačiu santykiu, tad ir pačios populiacijos laikytinos triamosios reikšmės atžvilgiu „vientisomis“. Todėl tam pačiam tikslui kartais naudojami ir kitokie, neparametriniai kriterijai, kurių statistikos nebesiejamos su binominio skirstinio parametru *p*.

B. Viena imtis, *k* nemažas (prielaidos apie proporcijos lygybę skaičiui tikrinimas)

Situacija. Ji kitokia, nei aptartoji A skyrelyje: turima tik viena *n* dydžio imtis, kurioje yra *k* objektų, turinčių rūpimąją požymio reikšmę, o likusieji (t.y. $n - k$) objektai jos neturi; *n* – palyginti nemažas skaičius (kelios dešimtys ar daugiau). Reikia patikrinti prielaidą, kad objektų, turinčių rūpimąją reikšmę, proporcija populiacijoje, iš kurios imtis paimta, esmingai nesiskiria nuo kokio nors iš anksto žinomo skaičiaus (iš intervalo [0, 1]) – tas skaičius (jį čia žymėsime *c*) gali būti, pvz., teoriškai išvesta rūpimojo požymio tikimybė, ankstesniais tyrimais nustatytas dydis ir pan. Iš imties duomenų galima gauti tiktai tikrinamos proporcijos (*p*) empirinį įvertį: $d = k / n$.

Statistinės prielaidos:

H_0: $p = c$
H_1: $p \neq c$

Kriterijaus statistika (paprastai žymima *z* arba *u*) irgi sukonstruota taip, kad nulinei hipotezei esant teisingai ji būtų *standartinį normalųjį* skirstinį turintis atsitiktinis dydis:

$$u = \frac{k - nc}{\sqrt{nc(1-c)}} \quad \text{arba} \quad u = \frac{d - c}{\sqrt{\frac{c(1-c)}{n}}}$$

Prielaidų vertinimas remiantis gautąja *u* reikšme – lygiai toks pat, kaip ir ankstesniame (A) skyrelyje.

Pavyzdys. Toksai veikėjas skelbėsi gebąs rankomis jausti spalvas... Surengtas eksperimentas: uždengton urnon įdėta 12 raudonų ir 18 žalių vienodo dydžio ir formos rutulių, jie kaskart išmaišomi ir „gebėtojas“ prašomas ištraukti būtent raudoną rutulį. Užfiksavus traukimo rezultatą rutulys grąžinamas atgal į urną. Bandymas kartotas 160 kartų, raudonas rutulys ištrauktas 59 kartus. Ar eksperimento rezultatai leistų patikimai atmesti prielaidą, kad šio veikėjo traukiamų rutulių ‘populiacijoje’ raudonos spalvos rutulių proporcija yra $p = 12 / (12 + 18) = 0{,}4$? Trumpai:

$$k = 59,\ \ n = 160$$
$$H_0:\ p = 0{,}4$$
$$H_1:\ p \neq 0{,}4$$

Apskaičiavę kriterijaus statistiką (KS) gauname $u = -\ 0{,}80687$, ir ją atitinkanti p-reikšmė yra 0,41974. Todėl nulinės hipotezės „išsakomą“ prielaidą atmetus, rizika padaryti pirmos rūšies klaidą būtų pavojingai didelė, netol 42%. Tad sakytina, jog eksperimento rezultatai specifinių šio veikėjo gebėjimų neįrodo.

Teoretikai nėra vieningos nuomonės, kokioms n bei c reikšmėms ši aproksimacija (tuo pačiu – ir šis u apskaičiavimo būdas) jau laikytina pakankamai tikslia. Apskritai ji tikslesnė tada, kai imtis didesnė ir kai c artimesnis 0,5. Pvz., ai kuriose statistikos knygose nurodoma, jog tokia aproksimacija jau pakankamai tiksli, kai bent viena iš sandaugų nc arba $n(1 - c)$ viršija 30, kitur – reikalaujama, kad imties dydis būtų didesnis už didžiausią iš dydžių: $5/c$, $5/(1 - c)$ ir $25(1 - 2c)^2/c(1 - c)$. Tad jeigu turima imtis yra nedidelė ir šių sąlygų neatitinka, reikia vadovautis kitokiais, mažoms imtims skirtais proporcijų palyginimo „instrumentais“. Jų aprašymų galima rasti statistikos knygose.

C. Viena imtis, *k* mažas (prielaidos apie proporcijos lygybę skaičiui tikrinimas: aproksimacija Puasono skirstiniu)

Tada, kai rūpimąją požymio reikšmę teturi tik labai maža dalis populiacijos objektų, ir k išlieka mažas skaičius net ir imtims esant palyginti didelėms, normalioji aproksimacija taip pat nebetinka. Tuomet racionaliau yra laikyti, kad H_0 esant teisingai k turi *Puasono* skirstinį su parametru $\lambda = np$. Todėl turimas k lygintinas su „įsivaizduojamu“ diskrečiuoju atsitiktiniu dydžiu (tarkime A), turinčiu Puasono skirstinį su parametru $\lambda = nc$. O prielaida apie proporcijos lygybę skaičiui c tuomet vertinama remiantis tikimybėmis, kad šis „įsivaizduojamas“ A bus *ne mažesnis* už k ir kad jis bus *ne didesnis* už k; formaliai sakant, tikimybėmis $T_1 = P_{(A >= k)}$ ir $T_2 = P_{(A <= k)}$. Apskaičiuoti šias tikimybes su *Excel* gana paprasta, prisiminus, jog bet kurio diskrečiojo skirstinio atveju $P_{(x >= s)} = 1 - F(s - 1)$, o $P_{(x <= s)} = F(s)$ – apie tai kalbėta per ankstesnes paskaitas. Tad su *Excel* šias tikimybes skaičiuotume:

T_1 = 1 – POISSON(k – 1; n*c; TRUE)
T_2 = POISSON(k; n*c; TRUE)

Šios tikimybės *tiesiogiai* ir apibūdina 1 rūšies klaidos riziką, kai nulinė hipotezė teisinga, t.y. įgauna panašią prasmę kaip ir p–reikšmė; tikriau sakant, jos faktiškai ir tampa p-reikšmėmis, kai alternatyva buna vienpusė:

Kai H_1: $p > c$, p-reikšmė = T_1
Kai H_1: $p < c$, p-reikšmė = T_2

O kai alternatyva dvipusė ($p \neq c$), taip pat praverstų nustatyti p-reikšmę, ir, nors tai gana keblu, tačiau galima išprotauti, kad su *Excel* ją racionalu rasti taip:

Kai H_1: $p \neq c$, p-reikšmė = IF(MIN(T_1, T_2) < 0,5; 2 * MIN(T_1, T_2); 1)

Nustatę gi p-reikšmę, pagal jos (bei, tuo pačiu, pirmos rūšies klaidos) dydį ir sprendžiame, ar nulinė hipotezė atmestina, ar ne. Statistikos knygose dažniausiai rekomenduojamas paprastesnis kelias: dvipusės alternatyvos atveju H_0 *atmesti*, jeigu nors viena tikimybė iš T_1, T_2 gaunama *mažesnė už pusę* pasirinkto reikšmingumo lygmens (esant vienpusei alternatyvai H_1: $p > c$, H_0 *atmetama*, kai T_1 gaunama *mažesnė*, o esant alternatyvai H_1: $p < c$, – kai T_2 *maženė* už pasirinktą reikšmingumo lygmenį).

Pavyzdys. Rizika susirgti xxx-titu šalyje apskritai yra 0,0004 (t.y. juo vidutiniškai serga 4 žmonės iš 10 000). Ištyrus 7000 atsitiktinai parinktų N miesto gyventojų, sergančių rasta 5. Ar tai leistų atmesti prielaidą, jog, lyginant su visa šalimi, šiame mieste rizika susirgti xxx-titu nėra pakitusi?

Turime:

$k = 5,\; n = 7000$
$H_0\colon\; p = 0{,}0004$
$H_1\colon\; p \neq 0{,}0004$

Su Excel apskaičiuotos $T_1 = 0{,}15235$, o $T_2 = 0{,}93489$; vadinasi, alternatyvai esant dvipusei, *p*-reikšmė būtų p = $2T_1 = 0{,}30465$. Taigi duomenys pakankamo pagrindo nulinei hipotezei atmesti *neduoda*: ją atmetus pirmos rūšies klaidos tikimybė būtų aiškiai per didelė. Per didelė (virš 15%) ji išliktų ir vienpusės alternatyvos $H_1\colon\; p > 0{,}0004$ atveju.

NEPARAMETRINIAI KRITERIJAI IR JŲ TAIKYMAS

Taip vadinami tie statistiniai kriterijai, kuriais naudojamasi tada, kai keliamos ir vertinamos statistinės hipotezės liečia ne empirinių skirstinių atskirus *parametrus*, bet kokius nors specifinius tyrinėtojui rūpinčius turimų duomenų aibių ypatumus. Paprastai tai – irgi įvairūs *palyginimo* uždaviniai, tik dabar – orientuoti į pačias duomenų aibes kaip visuminius objektus, o ne į atskirus jų skirstinių parametrus. Paminėtini trys dažnesni tipai uždavinių, kuriuos sprendžiant plačiai taikomi neparametriniai kriterijai: tai skirstinių *suderinamumo* uždaviniai, *homogeniškumo* uždaviniai ir požymių *nepriklausomumo* uždaviniai. Sprendžiant net ir to paties tipo (ar dar siauriau imant – tos pačios atmainos) uždavinius, galima naudotis gana įvairiais neparametriniais kriterijais, tad iškyla svarbus klausimas, kurį iš jų rinktis racionaliausia. Kriterijaus pasirinkimo motyvai gali būti gana įvairūs. Šiame konspekte paminėta tik vieno iš dažnai naudojamų neparametrinių kriterijų – kriterijaus *chi-kvadrat* – taikymo keli pavyzdžiai.

Reiktų pabrėžti, kad didžiumos neparametrinių kriterijų taikymo aplinkybės yra platesnės negu kad parametrinių, nes jie paprastai nereikalauja, kad turimi duomenys kokio nors tipo skirstinius; taigi, nereikalaujama ir duomenų apriorinio normalumo.

A. Suderinamumo uždaviniai

Pradėkime nuo pavyzdžio. Atsitiktinai parinkta 400 žmonių ir tirta, kurią savaitės dieną kievienas iš jų yra gimę. Gauti duomenys suvesti į tokią empirinio skirstinio lentelę, paremtą absoliutiniais kokybinio-požymio *savaitės diena* dažnumais:

Pirmad.	Antrad.	Trečiad.	Ketvirtad.	Penktad.	Šeštad.	Sekmad.
48	64	52	46	68	55	67

Ar šie duomenys suderinami su nuostata, kad tikimybė gimti bet kurią savaitės dieną yra vienoda (t.y. lygi 1 / 7 = 0,142857)?

Tai vienas iš paprastesnių suderinamumo uždavinių, nes visus empirinius kokybinio požymio (savaitės diena) reikšmių dažnumus tetenka lyginti tik su vienu, pastoviu „predikatu“, apspręstu pastovios tikimybės 1/7. Didžiumoje kitų suderinamumo uždavinių kiekvienos iš reikšmių tikimybė būna vis kitokia, tad ir lyginimo „predikatai“ įvairuoja.

Jeigu nuostata, kad tikimybė gimti bet kurią savaitės dieną yra pastovi ir lygi 1 / 7 = 0,142857 teisinga, tai bet kurią savaitės dieną „teoriškai“ turėtų būti gimę po... 0,142857*400 = 57, 14286 žmogaus! Sprendžiant klausimą apie šių konkrečių duomenų suderinamumą su minėta nuostata, būtent su šiuo „predikatu“ ir tektų lyginti turimus empirinius absoliutinius dažnumus. Tačiau – *kaip* lyginti?

Suderinamumo uždavinių statistinis modelis bendru atveju būtų toks:

Situacija. Imtis atsitiktinė, tiriamas požymis – diskretusis (jis gali būti kokybinis, ranginis arba kiekybinis diskretusis; jei jis – kiekybinis tolydusis požymis, tai diskretizuojamas dirbtinai, paprastai – grupuojant reikšmes intervalais). Imtyje jis įgauna k skirtingų reikšmių. Eksperimentiškai yra nustatytas kiekvienos (i-tosios; $i = 1, 2 \dots k$) iš tų reikšmių absoliutinis dažnumas a_i, o taip pat esama galimybių apskaičiuoti kiekvienos iš jų teorinę tikimybę t_i. Visi šie duomenys paprastai suvedami į vadinamąją dažnių lentelę, panašią į anksčiau pateiktąją, tik turinčią dar ir trečią eilutę, skirtą teorinėms reikšmių tikimybėms t_i. Natūralu, kad taip pat yra žinomas ir visos imties dydis (n) – jis prilygsta visų reikšmių absoliutinių dažnumų sumai: n = suma(a_i), kur $i = 1, 2 \ldots k$.

Statistines prielaidas šiuo atveju suprantamiau yra suformuluoti žodžiais. Bendru atveju jas galima apibūdinti maždaug taip:

H_0: Stebimasis empirinis reikšmių skirstinys iš esmės nesiskiria nuo parinktojo teorinio modelio ir yra su juo statistiškai suderinamas

H_1: Stebimasis empirinis reikšmių skirstinys iš esmės skiriasi nuo parinktojo teorinio modelio ir yra su juo statistiškai nesuderinamas

Formalizuotas statistinių prielaidų „pavaizdavimas“ taip pat galimas: tada nulinėje hipotezėje teigtume, kad visiems $i = 1, 2 \ldots k$ požymio skirtingų reikšmių proporcijos p_i populiacijoje, iš kurios paimta imtis, esmingai nesiskiria nuo jų teorinių tikimybių t_i:

H_0: $p_i = t_i$
H1: $p_i \neq t_i$

$(f) = k - 1$

Kriterijaus statistika skaičiuojama tokių dydžių suma (pažymėkime ją x):

$$x = \sum_{i=1}^{k} \frac{(a_i - nt_i)^2}{nt_i}$$

Kai H_0 teisinga, ši KS turi chi-kvadrato skirstinį su tam tikru (f) laisvės laipsnių skaičiumi, kuris priklauso nuo to, kokiu konkrečiu būdu buvo apskaičiuojamos tikimybės t_i, tiksliau – nuo to, *kelių* parametrų empiriniais *įverčiais* naudotasi apskaičiuojant t_i. Šių įverčių kiekį pažymėję h, laisvės laipsnių skaičių gautosios x chi-kvadrato skirstiniui nustatytume šitaip:

$f = k - h - 1$

Prielaidų vertinimas remiasi apskaičiuotąja x reikšme. Jeigu remiamasi iš anksto pasirinktu reikšmingumo lygmeniu α, tai H_0 atmetama, kai x viršija turimą laisvės laipsnių skaičių atitinčią kritinę reikšmę (KR), kurią galima arba rasti statistikos knygų prieduose, arba apskaičiuoti su *Excel* remiantis žinomu laisvės laipsnių skaičiumi f ir pasirinktuoju reikšmingumo lygmeniu α:

KR = CHIINV(f; α)

Kita vertus su *Excel* lengva apskaičiuoti ir gautąjį x dydį atitinkančią p-reikšmę:

p = CHIDIST(x; f)

Jeigu p-reikšmė gaunama nedidelė (paprastai – mažesnė už tradicinius reikšmingumo lygmenis), H_0 taip pat atmetama (su tokio pat dydžio pirmos rūšies klaidos rizika) bei, tuo pačiu, verifikuojama (patvirtinama esanti teisinga) alternatyva. O jeigu p-reikšmė gaunama pakankamai žymi ar netgi didelė, tai H_0 neatmetama ir paliekama toliau „likti prielaida“.

Jeigu grįžtume prie šio skyrelio pradžioje pateikto pavyzdžio ir aptartu būdu jam apskaičiuotume kriterijaus statistiką, tai gautume $x = 8{,}765$; laisvės laipsnių skaičius būtų $f = 7 - 1 = 6$ (nes kokybinis požymis *savaitės diena* įgyja 7 skirtingas reikšmes, o tikimybes $t_i = 1/7$ „išprotavome“ nesinaudodami *nė vieno* parametro empiriniu įverčiu, taigi, $h = 0$), todėl pagal imties duomenis gaunama p-reikšmė būtų p = 0,187232. Tad nulinei hipotezei atmesti nebūtų visiškai tvirto pamato (kone 19% siekianti pirmos rūšies klaidos rizika dar yra gerokai per didelė), ir turėtume sakyti, kad šios imties duomenys *nepaneigia* nuostatos, jog tikimybė gimti bet kurią savaitės dieną yra vienoda.

Kartais tiriamojo požymio reikšmių skirstinys pavaizduojamas pateikiant ne absoliutinius, bet *santykinius* jo įgyjamų reikšmių dažnumus, kitaip tariant – empirinius reikšmių proporcijų p_i (populiacijoje, iš kurios paimta imtis) įverčius $d_i = a_i / n$. Nesunku įrodyti, kad tokiais atvejais kriterijaus statistikos x apskaičiavimo būdas irgi būtų panašus į pateiktąjį, tik „algerbiškai“ modifikuotas:

$$x = n \sum_{i=1}^{k} \frac{(d_i - t_i)^2}{t_i}$$

Suderinamumo uždavinius tenka spręsti ir siekiant patikrinti, ar eksperimento metu gautų duomenų empirinis skirstinys atitinka kokį nors hipotetinį teorinį skirstinį (pvz., normalųjį – tai dažna problema).