PREIZKUŠANJE HIPOTEZ

1. UVOD

Tako pogosto kot vrednost parametrov osnovne množice nas zanima vprašanje, ali se osnovne množice razlikujejo. Ko govorimo o razlikah med množicami, mislimo na njihove parametre. Ali se aritmetične sredine osnovnih množic razlikujejo? Največkrat se omejimo na razliko dveh osnovnih množic. To je namreč bolj preprosto, kot če bi hoteli odgovoriti na vprašanje o več množicah hkrati. Tudi če imamo več množic, si lahko primerjavo med parametri vedno razdelimo na po dve in dve množici in te dvojice primerjamo.
Če bi imeli podatke za cele osnovne množice, bi, npr. izračunali njihove aritmetične sredine in jih primerjali »na pogled«. Že najmanjša razlika med sredinama pomeni, da se množici razlikujeta. Te razlike ne bi bilo treba na nikakršen dodatni način dokazovati. Dokazuje se sama!
Kadar pa imamo podatke le za dva vzorca iz teh osnovnih množic, postopek ni več tako preprost. Tudi če sta vzorčni aritmetični sredini različni, še ne moremo vedeti, ali se razlikujeta tudi aritmetični sredini osnovnih množic. Da res ne, si lahko razložimo s preprostim premislekom: že, če bi oba vzorca izbrali iz iste osnovne množice, bi se njuni sredini razlikovali, kaj šele, če ju izberemo iz različnih množic. Zato bo potreben preizkus v tem smislu: »Ali sta vzorčni sredini zadosti različni, da to potrjuje različnost osnovnih množic, ali ne?«. Takšne statistične preizkuse imenujemo preizkusi ničelnih hipotez.

2. Preizkušanje NIČELNE hipotezE o razliki med aritmetičnimi sredinami z malimi vzorci

Sprašujemo se, ali sta aritmetični sredini osnovnih množic različni. Vzorčni sredini se praktično vedno razlikujeta; o sredinah osnovnih množic pa tega še ne vemo. Glede tega postavimo prvo hipotezo:

Aritmetični sredini osnovnih množic se ne razlikujeta.

To je ničelna hipoteza. Zapišemo jo lahko matematično na dva načina:

$$M_1 - M_2 = 0$$

ali tudi:

$$M_1 = M_2$$

Zaenkrat še ne vemo, ali je pravilna. Šele preizkus bo pokazal, kolikšna je verjetnost, da je pravilna in kolikšna je verjetnost, da ni pravilna. Nasproti tej hipotezi postavimo še nasprotno hipotezo:

Aritmetični sredini osnovnih množic se razlikujeta.

Iz prve osnovne množice izberemo vzorec z n_1 in iz druge vzorec z n_2. Na razliki aritmetičnih sredin teh vzorcev bo temeljil preizkus. Ta razlika je $x_1 - x_2$.
Sedaj za kratek hip moramo vpeljati prej zamišljeno množico vseh možnih vzorcev (tokrat: vseh možnih parov vzorcev). Če so vzorci dovolj veliki, se bodo vzorčne razlike porazdeljevale približno normalno, pri malih vzorcih pa v t-porazdelitvi.
Naj velja ničelna hipoteza in je razlika aritmetičnih sredin osnovnih množic enaka nič. Potem se bodo vzorčne razlike porazdeljevale po zakonitosti t-porazdelitve okoli ničle. Spomnimo se, da so se aritmetične sredine velikih vzorcev porazdeljevale približno normalno.
Razložimo si to dejstvo kar se da preprosto. Če je razlika sredin osnovnih množic deset (npr. deset cm, če gre za telesne višine), potem bo večina vzorčnih razlik okoli deset. Približno polovica vseh parov vzorcev bo imela razliko manjšo od deset in polovica večjo od deset. Če bo populacijska razlika petdeset tisoč, potem bo tudi v večini parov vzorcev razlika okoli petdeset tisoč (npr. učiteljske plače v Sloveniji in na Hrvaškem preračunane v tolarje). In na koncu, če se populacijski sredini ne razlikujeta (razlika je nič), potem je tudi v večini parov vzorcev razlika blizu ničle. Približno polovica parov ima negativno razliko

(aritmetična sredina drugega vzorca je malo večja od aritmetične sredine v prvem vzorcu) in druga polovica pozitivno razliko (prvi vzorec ima malo večjo sredino). To se smiselno povsem sklada s tistim, kar smo že prej ugotovili o porazdelitvi aritmetičnih sredin posameznih vzorcev. Ilustrirajmo vse to še z nekoliko nenavadnimi a poučnimi primeri. Zaradi razumljivosti razlage (boljše berljivosti besedila) bomo kar na pogled ocenjevali rezultate v vseh primerih (potrudili smo se biti kar najbolj realni; bralec lahko presodi, koliko je to uspelo). Iz osnovne množice dijakov srednje šole smo izbirali vzorce in dijakom izmerili telesno višino v cm. Enako smo izbirali vzorce iz osnovne množice učencev osnovne šole. Če je aritmetična sredina v osnovni množici dijakov npr. M_1= 169 cm, bodo večinoma vzorci dijakov imeli sredino med 165 in 173 cm. Če je aritmetična sredina v osnovni množici učencev osnovne šole npr. M_2 =150 cm, bodo sredine večine vzorcev med 147 in 153 cm. In če sedaj jemljemo pare: po en vzorec iz množice dijakov in po en vzorec iz množice učencev, bomo dobili razlike približno petnajst do petindvajset točk (takšna je približno razlika med osnovnima množicama). In še krajši primer: med povprečnimi plačami v vzorcih slovenskih učiteljev in v vzorcih nemških učiteljev bo približna razlika 2000 evrov (ker je približno takšna razlika med povprečjem plač vseh učiteljev v Sloveniji in Nemčiji). Zaradi slučajnostnega izbora učiteljev v vzorce bodo razlike nekje malo večje od 2000 in drugje malo manjše od 2000; toda nasplošno bodo okoli 2000.
Vrnimo se k naši glavni razlagi. Če torej velja ničelna hipoteza, potem se bodo vzorčne razlike porazdeljevale okoli ničle. Sedaj bomo uporabili zakonitosti t-porazdelitve. Pri velikih vzorcih smo na tem mestu uporabili z-preizkus. Sedaj bomo uporabili t-preizkus.

$$t = \frac{\bar{x}_1 - \bar{x}_2}{se_{\bar{x}_1 - \bar{x}_2}}$$

Za naš par vzorcev izračunamo razliko in iz nje vrednost »t«. Ta je lahko manjša ali večja od kritične vrednosti t iz tabele.
Kritično vrednost »t« bomo označili takole:

$t\ (a=2P=0{,}05;\ g)$

Vrednost g določimo po obrazcu:

$g = n_1 + n_2 - 2$

Na koncu t-preizkusa se lahko zgodi dvoje:

1. Dobljena vrednost t je večja od kritične vrednosti

Če bi držala ničelna hipoteza, bi smelo biti takšnih parov največ 5 odstotkov. Ker pa smo mi takšen par dobili, lahko sklepamo takole. Verjetnost, da se je to zgodilo je le 5 odstotna (ničelna hipoteza drži, mi smo dobili tako veliko razliko). Bolj je verjetno, da smo dobili tako veliko razliko, ker že ničelna hipoteza ne drži. Verjetnost tega drugega sklepa je 95 %.
Seveda bolj verjamemo v to drugo možnost, saj je verjetnost zanjo neprimerno večja. Sprejeli bomo sklep, da pač ničelna hipoteza ne drži in jo bomo **zavrnili.** Hkrati bomo sprejeli nasprotno hipotezo. S tem smo dokazali, da se aritmetični sredini osnovnih množic razlikujeta.
Ne smemo pa pozabiti, da vendarle obstaja 5% možnosti, da je ničelna hipoteza pravilna. Pri zavrnitvi tvegamo 5%, da je vendarle ničelna hipoteza pravilna in da se aritmetični sredini osnovnih množic v resnici ne razlikujeta.

2. Dobljena vrednost t je manjša od kritične vrednosti

Tudi sedaj bi lahko podobno sklepali kot prej, le tveganje ob zavrnitvi bi bilo večje od 5%. Ker pa velja praksa, da večjega tveganja od 5% ne dopuščamo, ne bomo zavrnili ničelne hipoteze. V takem primeru jo bomo **obdržali.** Kaj to pomeni in kaj nam to o osnovnih množicah pove? Pravzaprav nič! V primeru, da obdržimo ničelno hipotezo, nismo dokazali, da se množici razlikujeta. Ali to pomeni, da se ne razlikujeta? Na prvi pogled se zdi, da je tako. Zdrav razum nam narekuje: če nismo dokazali, da sta množici različni, potem sta pač enaki. Vendar je to prenagljen sklep. V primeru, da obdržimo ničelno hipotezo, nismo o osnovnih množicah dokazali nič. Bralec naj sam temeljito premisli zadevo in se prepriča, da je tako. Ker o osnovnih množicah nismo dokazali ničesar (niti, da se razlikujejo niti, da se ne!) ni tudi nikakršnega tveganja. Ni smiselno reči, npr. *»ničelno hipotezo smo obdržali s tveganjem 5%«* ali kaj podobnega. Preprosto zaključimo, da smo ničelno hipotezo obdržali in da o osnovnih množicah ne moremo ničesar trditi.

Takšen preizkus imenujemo kratko t-preizkus. Podobno kot pri z-preizkusu kot izhodišče jemljemo tri kritične vrednosti iz tabele t-porazdelitve. Ponekod jih označujejo takole:

za 5 % tveganja	t (α=0,05)
za 1 % tveganja	t (α=0,01)
za 0,1 % tveganja	t (α=0,001)

Grška črka α v tem zapisu označuje raven statistične pomembnosti (če α pomnožimo s sto, dobimo odstotek tveganja). Poleg α se za podoben namen uporablja tudi oznaka P. Z oznako P se označuje le nezajeti del ploščine na eni strani porazdelitve. Pri našem t-preizkusu upoštevamo obe strani krivulje. Ker imamo en del ploščine pod krivuljo na levi strani in drugega na desni strani, so potemtakem oznake za kritične vrednosti takšne:

za 5 % tveganja	t (2P=0,05)
za 1 % tveganja	t (2P=0,01)
za 0,1 % tveganja	t (2P=0,001)

Ali še popolneje:

za 5 % tveganja	t (α=2P=0,05)
za 1 % tveganja	t (α=2P=0,01)
za 0,1 % tveganja	t (α=2P=0,001)

Pri tistih postopkih in porazdelitvah, kjer je upoštevana le ena stran krivulje, je v oklepaju samo P (npr. P=0,05 itd.). Raje bomo uporabljali oznako α; oznaki P in 2P omenjamo predvsem zaradi jasnosti pri uporabi drugih statističnih priročnikov.
Ker je treba izide t-preizkusa natančno interpretirati, zapišimo in poglejmo še enkrat možne primere.

1. |t| < t (a=2P=0,05; g)

Razlika med vzorčnima aritmetičnima sredinama ni statistično pomembna. Ničelno hipotezo obdržimo. O osnovnih množicah ne moremo trditi ničesar.

2. |t| ‡ t (a=2P=0,05; g)

Razlika med vzorčnima aritmetičnima sredinama je statistično pomembna na ravni α=0,05. Ničelno hipotezo zavrnemo s tveganjem 5%. Sprejmemo nasprotno hipotezo. Dokazali smo, da se aritmetični sredini osnovnih množic razlikujeta.

Če je »t« precej večji od kritičnega, lahko dopolnimo prejšnji sklep takole.

3. |t| ‡ t (a=2P=0,01; g)

Razlika med vzorčnima aritmetičnima sredinama je statistično pomembna na ravni α=0,01. Ničelno hipotezo zavrnemo s tveganjem 1%. Sprejmemo nasprotno hipotezo. Dokazali smo, da se aritmetični sredini osnovnih množic razlikujeta.

4. |t| ‡ t (a=2P=0,001; g)

Razlika med vzorčnima aritmetičnima sredinama je statistično pomembna na ravni α=0,001. Ničelno hipotezo zavrnemo s tveganjem 0,1%. Sprejmemo nasprotno hipotezo. Dokazali smo, da se aritmetični sredini osnovnih množic razlikujeta.

Poudarimo še enkrat: možna sta le dva vsebinsko različna izida: ničelno hipotezo obdržimo in ničelno hipotezo zavrnemo.

Preizkus, ki smo ga opisali jer nekoliko bolj občutljiv na okoliščine v katerih ga izvajamo (bolj kot z-preizkus). Temeljna okoliščina, ki odločilno vpliva na potek preizkusa je razpršenost v obeh osnovnih množicah. Običajni t-preizkus lahko opravimo, če je izpolnjen naslednji pogoj:

Razpršenost v obeh osnovnih množicah mora biti enaka.

Tega pogoja neposredno ne moremo preveriti, saj nimamo podatkov o osnovnih množicah. Preverjamo ga posredno – s pomočjo podatkov za oba vzorca.
Vsebina preizkusa bo naslednja: če se varianci vzorcev ne razlikujeta preveč, potem se verjetno ne razlikujeta niti varianci osnovnih množic (vsaj ne zelo). Če se že varianci vzorcev (preveč) razlikujeta, potem se verjetno razlikujeta tudi varianci osnovnih množic. Torej naredimo preizkus ničelne hipoteze o razliki varianc osnovnih množic. Ker služi le kot uvod v t-preizkus, ga imenujemo tudi predpreizkus.
Izračunamo varianci za oba vzorca. Postavimo ničelno hipotezo: varianci osnovnih množic se ne razlikujeta. To hipotezo preizkusimo z F-preizkusom:

$$F = \frac{s_{več}^2}{s_{man}^2}$$

legenda

$s_{več}^2$	Večja varianca
s_{man}^2	Manjša varianca

F-porazdelitev je določena z dvojimi stopinjami prostosti. Oboje določimo iz numerusov vzorcev. Toda pozor: za g1 vzamemo vzorec, katerega varianca je v števcu (torej tistega z večjo varianco), za g2 pa vzamemo vzorec, katerega varianca je v imenovalcu (torej tistega z manjšo varianco).
Torej ne bomo napisali $\mathbf{g_1=n_1-1}$, ker bi to bilo napačno. Ne bomo napisali niti $\mathbf{g_1=n_{večji}-1}$, saj bi tudi to bilo napačno (podobno tudi za g_2). Zaradi enostavnosti bomo zaplet z zapisovanjem enačb za g1 in g2 rešili tako:

$\mathbf{g_1=n-1}$ (za vzorec z večjo varianco)
$\mathbf{g_2=n-1}$ (za vzorec z manjšo varianco)

Izračunano vrednost F primerjamo s kritično vrednostjo F iz tabele (za 5% tveganja, ali za a=0,05). Tabelo lahko bralec najde v Sagadin 1992 (v prilogah). Možna sta naslednja izida:

1. F < F (a=2P=0,05; g_1; g_2)

Razlika med vzorčnima variancama ni statistično pomembna. Ničelno hipotezo obdržimo. Ni dokaza, da se varianci osnovnih množic razlikujeta.
V tem primeru lahko uporabimo običajni t-preizkus, kot smo ga zgoraj prikazali.

2. 1. F ‡ F (a=2P=0,05; g; g_2)

Razlika med vzorčnima variancama je statistično pomembna. Ničelno hipotezo zavrnemo. Torej se varianci osnovnih množic razlikujeta (s tveganjem največ 5%). V tem primeru ne moremo (ali ne smemo) uporabiti prikazanega t-preizkusa. Namesto njega moramo izbrati nek nadomestni preizkus (glej Sagadin 2003). V primeru računalniške obdelave s programom SPSS bo v tabeli z rezultati na izbiro običajni t-preizkus in nadomestni. Izberemo glede na rezultate F-preizkusa (ki ga računalniški program tudi naredi. Treba je samo iz tabele izbrati ustreznega.

3. POROČANJE O PREIZKUSIH V PRIMERU RAČUNALNIŠKE OBDELAVE

Ker v novejšem času opravljamo statistične obdelave večinoma računalniško, bo končno poročilo o preizkusu nekoliko drugačno (samo bistvo pa ne!). Računalnik ne potrebuje tabele, temveč za preizkušanje ničelne hipoteze

uporablja kar enačbo t-porazdelitve. Zato namesto primerjave s kritičnimi vrednostmi ponuja natančno izračunano raven statistične pomembnosti.
Torej povsod, kjer je raven statistične pomembnosti 0,05 ali manj, zavrnemo ničelno hipotezo; tam kjer je vrednost večja od 0,05 pa jo obdržimo. V končnem poročilu navedemo vrednost "t", število prostostnih stopinj in raven statistične pomembnosti. Tako poročamo tudi v primerih, ko nismo zavrnili ničelne hipoteze. S tem povemo vse, kar je potrebno za razumevanje interpretacije.

4. STANDARDNA NAPAKA

Standardno napako za posamezne vzorce že poznamo. Vemo tudi, da prave vrednosti standardne napake ne moremo dobiti in da uporabljamo nadomestek (oceno standardne napake). Spet ne bomo zapletali imen in bomo ta nadomestek imenovali kratkomalo standardna napaka (če ne zaradi drugega, že zato, ker prave tako nikdar nimamo). Standardno napako razlike dobimo iz standardnih napak posamičnih vzorcev. Pri tem sta lahko vzorca odvisna ali neodvisna; standardna napaka pri obojih ni enaka! Standardna napaka razlike je pri odvisnih vzorcih manjša. Razlikovanje odvisnih vzorcev od neodvisnih pri velikih vzorcih nima velikega praktičnega pomena. Le v izjemnih primerih imamo velike odvisne vzorce. Odvisni vzorci so veliko bolj pogosti med malimi vzorci. Postopek izbiranja odvisnih vzorcev je zahteven, še zlasti če bi šlo za velike vzorce.

Standardna napaka za neodvisne vzorce:

$$se_{\bar{x}_1-\bar{x}_2} = \sqrt{se^2_{\bar{x}_1} + se^2_{\bar{x}_2}}$$

Standardna napaka za odvisne vzorce:

$$se_{\bar{x}_1-\bar{x}_2} = \sqrt{se^2_{\bar{x}_1} + se^2_{\bar{x}_2} - 2r_{12} \cdot se_{\bar{x}_1} \cdot se_{\bar{x}_2}}$$

legenda
r_{12} je korelacijski koeficient med rezultati v obeh vzorcih

Vidimo, da je standardna napaka za odvisne vzorce manjša kot za neodvisne. Del, ki se odšteva v enačbi standardne napake, je odvisen od korelacijskega koeficienta. Čim močnejša je korelacija med rezultati obeh vzorcev, tem manjša je standardna napaka odvisnih vzorcev. Manjša standardna napaka je seveda ugoden izid. Preprosto povedano: čim bolje nam uspe izenačevanje parov pri izbiri vzorcev, tem večjo korist imamo od odvisnih vzorcev.

Primer

Na radiotelegrafskem tečaju so merili hitrost sprejemanja znakov Morzejeve abecede.

Tabela 7. Podatki za dva vzorca

	numerus	aritm. sredina	varianca (stand. odklon)
prvi vzorec	18	60,50	340,82 (18,46)
drugi vzorec	11	44,50	319,30 (17,87)

Izračunamo standardno napako razlike in nato izračunamo z-vrednost:

Tabela 8. Preizkus ničelne hipoteze o razliki aritmetičnih sredin

predpreizkus F		
	$F = \frac{340{,}82}{319{,}30}$	1,07
kritična F vrednost za 5%	$g_1 = 18 - 1 \quad g_2 = 11 - 1$ *	2,69
izid F preizkusa	1,07<2,69 Ničelno hipotezo obdržimo, zato lahko delamo običajni t-preizkus	
glavni preizkus		
standardna napaka v prvem vzorcu	$se_1 = \frac{18{,}46}{\sqrt{18}}$	4,35
standardna napaka v drugem vzorcu	$se_2 = \frac{17{,}87}{\sqrt{11}}$	5,39
standardna napaka razlike	$se = \sqrt{se_1^2 + se_2^2}$ $se = \sqrt{4{,}35^2 + 5{,}39^2}$	6,93
t-vrednost	$t = \frac{60{,}50 - 44{,}50}{6{,}93}$	2,31
kritična t vrednost za 5%	$g = 18 + 11 - 2 = 27$	t=2,052

$$z = 2{,}31 > 2{,}052$$

Razlika je statistično pomembna na ravni 0,05 in s tveganjem 5% zavrnemo ničelno hipotezo. Aritmetični sredini osnovnih množic se razlikujeta.

*pri določanju prostostnih stopinj g1 in g2 je potrebna velika pazljivost. V našem primeru je za g1 uporabljen numerus prvega vzorca (18-1), toda ne zato ker je prvi, ampak zato ker ima večjo varianco (njegova varianca je v števcu izraza za F). Za g2 je uporabljen numerus drugega vzorca (11-1), ker je to vzorec z manjšo varianco (in ne zato ker je drugi).

5. ZNAČILNOSTI t-PREIZKUSA

Ob t-preizkusu lahko premislimo še o nekaj značilnostih, ki pa se smiselno nanašajo na skoraj vse preizkuse.

1. Tveganje, ki je samo po sebi neprijetno, ima še neko senčno stran. Namreč to, da ob zavrnitvi ničelne hipoteze, nikdar ne vemo, ali se je prav tokrat zgodil redek slučaj, ko je ničelna hipoteza pravilna (mi smo jo pa kot nepravilno zavrnili!). Res je, da je takšnih primerov največ pet odstotkov (pogosto pa še dosti manj), a vendarle: katerakoli zavrnitev je lahko napačna. Ali je pravilna ali napačna nam ostane skrito.
2. Prikazani preizkus je dvostranski. Na grafičnem prikazu porazdelitve vidimo na obeh straneh ploščini, ki predstavljata tveganje. Pri tveganju 5 odstotkov je 2,5% ploščine na levi strani in 2,5% ploščine na desni strani. Zato se v preizkusu ne oziramo na predznak vrednosti "t". Obstaja tudi enostranski t-preizkus. Pri takšnem preizkusu je ničelna hipoteza malo drugačna, drugačne so tudi kritične vrednosti "t". Tukaj ne bomo podrobneje obravnavali enostranskih preizkusov; ti se na pedagoškem področju uporabljajo le izjemoma. Pri enostranskem "t" preizkusu je nezajeti del samo na eni strani porazdelitve. Pri dvostranskem preizkusu je α=2P, pri enostranskem pa α=P.
3. Izraz *statistična pomembnost* zahteva nekaj več pojasnil. Ta izraz moramo razumeti ozko statistično, saj ima prav malo skupnega s siceršnjim

pomenom izraza *pomembnost!* Če je neka razlika statistično pomembna, to še sploh ne pomeni, da je vsebinsko pomembna. Lahko, npr. ugotovimo, da je med deklicami in dečki statistično pomembna razlika v dolžini stopala, pa to dejstvo ni za učitelja slovenščine prav nič pomembno. Preveden v vsakdanji govor pomeni izraz *statistična pomembnost* približno tole: *razlika je dovolj velika* (...da jo lahko posplošimo na osnovne množice).

4. Pojem statistične pomembnosti se nanaša **samo** na vzorčne parametre; nikdar pa ne na populacijske. Zato ne moremo reči, npr.: "Aritmetični sredini osnovnih množic se statistično pomembno razlikujeta". Kadar rečemo, "tale razlika je statistično pomembna", ne smemo pomisliti na osnovne množice.
5. Po drugi strani pa se ničelne hipoteze nanašajo samo na osnovne množice! Zakaj? Najprej zato, ker o množicah res ne vemo tega, kar se sprašujemo. Zato o neznanem v osnovni množici postavimo hipotezo. Razlika v vzorcu nam je znana (lahko jo izračunamo, in to vsakokrat tudi zares naredimo) in o znanem nima smisla postaviti hipoteze. Hipoteze se postavljajo nasplošno o neznanem in ne o znanem. O tem ali bo včeraj deževalo, nima smisla postavljati hipoteze (saj to že vemo); smiselno pa je postaviti hipotezo o tem, ali bo jutri deževalo (saj tega še ne vemo). Ničelne hipoteze o vzorcih ne postavljamo tudi zato, ker nas vzorci pravzaprav ne zanimajo. V resnici nas zanimajo samo osnovne množice.

Pogosto na začetku preučevanja vzgojnih pojavov postavimo raziskovalno hipotezo. Postavimo jo na podlagi tega, kar o teh pojavih že od prej vemo. Lahko je pravilna, ali pa tudi ne. To se bo šele pokazalo. Ničelna hipoteza je praktično vedno nasprotna tistemu, kar zares predvidevamo, torej je *nasprotna raziskovalni hipotezi.* Kaj naprimer pomeni raziskovalna hipoteza "uspeh učencev, ki se vozijo v šolo, se razlikuje od uspeha učencev, ki se ne vozijo"? To preprosto pomeni, da mislimo, da je uspešnost povezana z vožnjo v šolo. Nekateri učenci veliko časa izgubijo z vožnjo, zato jim ostane manj časa in energije za učenje. Za trditvami o razliki se skoraj vedno skriva povezanost med pojavi. Razlika v nekih stališčih med dečki in deklicami enostavno pomeni, da so stališča povezana s spolom, da spol vpliva na stališča. Razlika v načinu poučevanja med bolj izkušenimi in manj izkušenimi učitelji pomeni, da izkušnje vplivajo na način poučevanja. Takšnih primerov bi lahko našli nešteto. Če bi resno mislili z ničelno hipotezo (“ni razlike...”), bi pravzaprav rekli: “Pojav A ne vpliva na pojav B”. Proučevanje nepovezanih pojavov pa nima veliko smisla. Nepovezanih pojavov je na svetu preveč, da bi bilo vredno vlagati energijo v dokazovanje njihove nepovezanosti. Dokazovanje nepovezanosti pojavov je smiselno le v izjemno redkih primerih, npr. ko vse teorije trdijo, da sta pojava povezana, nam se pa zdi, da nista. V takem primeru želimo ovreči neko znanstveno zmoto in pokazati da sta A in B nepovezana. Povrhu pa smo že videli, da tudi v primeru, ko ničelno hipotezo obdržimo, tega ne moremo šteti kot dokaz, da je res pravilna. Takšen izid preizkusa razumemo kot pomanjkanje kakršnegakoli dokaza. Zato prikazani statistični preizkus ničelne hipoteze sploh ni primeren za dokazovanje pravilnosti ničelne hipoteze (tudi, če bi z njo mislili resno). Zaključimo na kratko: ničelno hipotezo postavljamo predvsem zaradi postopka, resno pa mislimo z nasprotno hipotezo. Zato je “življenjska doba” ničelne hipoteze zelo kratka: nastane na začetku preizkusa in se konča na koncu preizkusa.