

1. UVOD

POPULACIJA → množica vseh elementov, na katere se nanašajo naše ugotovitve. Sestoji iz posameznih "elementov" (osebe, gospodinjstva, kmetije, šole, podjetja ipd.), ki so osnovna enota naše analize. Priporočljivo je, da proučevano populacijo na samem začetku opredelimo v idealni obliki kot željeno ciljno populacijo. Vendar pa se anketirana populacija pogosto razlikuje od ciljne populacije, zato je natančnost pri njeni opredelitvi – kot tudi pri spremljanju morebitnih odstopanj – izredno pomembna.

VZOREC → ko je populacija opredeljena, se lahko začnemo ukvarjati z vzorcem, ki je del ciljne populacije, na podlagi katerega izvedemo sklepanje o celotni populaciji. Zbiranje podatkov iz dela populacije je cenejše, pridobimo na času, hkrati pa zagotovimo povsem zadostno natančnost ocen. Zaradi večje kakovosti anketiranja pridobimo celo boljše podatke v primerjavi s popisom celotne populacije.

VERJETNOSTNI VZOREC → vsak element v populaciji ima vnaprej znano in neničelno verjetnost, da se pojavi v vzorcu in samo pri verjetnostnih vzorcih lahko uporabimo statistično sklepanje, ki je podlaga za izračun intervalov zaupanja. Prvi predpogoj za korektno statistično sklepanje je verjetnostna izbira elementov – izbira, za katero je med drugim značilno tudi to, da je neodvisna od anketarja (verjetnostne vzorce zaradi objektivnosti izbire in kvantificiranja potencialnega tveganja za napako včasih imenujemo tudi znanstveni vzorci).

Vzorci, ki ne zadostijo pogojem o vnaprej znani in neničelni verjetnosti izbora, se imenujejo neverjetnostni vzorci, ki torej ne omogočajo statističnega sklepanja, oblikovanja strokovnih trditev oz. statističnih interpretacij, pri katerih je tveganje strokovnih trditev natančno opredeljeno.

Prvi pogoj za izvedbo verjetnostnega vzorčenja je obstoj vzorčnega okvira, iz katerega izberemo elemente v vzorec (prostorsko vzorčenje, ki poskrbi za vzpostavitev posebnega nadomestka za vzorčni okvir, ko nimamo celotnega seznama elementov v populaciji – uporabljamo pri gospodinjstvih, stanovanjih in poslovnih prostorih).

2. ENOSTAVNO SLUČAJNO VZORČENJE ~ Simple Random Sampling

Enostavno slučajno vzorčenje (**SRS**), kot izhodišče za obravnavo verjetnostnih vzorcev, je najpreprostejša metoda, ki vsebuje vsa načela kompleksnejših metod (kot kompleksne vzorčne načrte obravnavamo vse vzorce, ki niso SRS vzorci).

N – velikost populacije, n – velikost vzorca; pri SRS vzorčenju ima vsaka podmnožica n različnih elementov iz populacije z N elementi enak izbor. S kratico **EPSEM** (Equal Probability Selection Method) pa označimo enako verjetnost vsakega elementa v populaciji za izbor v vzorec. Poleg SRS vzorcev imajo lastnost EPSEM tudi nekateri kompleksni vzorčni načrti. Posebnost SRS vzorcev je torej v tem, da imajo poleg siceršnje lastnosti EPSEM vsi SRS vzorci določene velikosti tudi enako verjetnost, da jih izberemo. Slednje, kot rečeno, ne velja za vse EPSEM vzorčne načrte (eden od načinov za izbiro vzorca je lahko metoda loterije, tabela slučajnih števil). Enostavno slučajno vzorčenje (SRS) pa brez dodatnih oznak razumemo kot vzorčenje brez ponavljanja.

→ glej stran 15 – 16 (1), (1a), (1b) ⇒ izraza σ^2 (teoretično) in S^2 (primernejša oblika za statistično uporabo) se uporabljata enakovredno

VZORČNA VARIANCA → pojavlja se vprašanje, kako dobra cenilka populacijske aritmetične sredine \bar{Y} je vzorčna aritmetična sredina \bar{y} . Cenilko, to je funkcijo, ki izbranemu vzorcu vsakič dodeli določeno vrednost, včasih imenujemo tudi statistika. Ker populacijske aritmetične sredine \bar{Y} običajno ne poznamo, je videti, da na podlagi enega samega vzorca ni mogoče odgovoriti na vprašanje o kakovosti cenilke \bar{y} . Za odgovor bi namreč potrebovali ponavljajoče se izbiranje vzorcev.

Ocena označuje določeno vrednost (npr. vzorčno aritmetično sredino) iz posameznega vzorca, medtem ko izraz cenilka označuje funkcijo – to je slučajno spremenljivko – ki izhaja iz splošnega postopka za izračun ocene populacijskega parametra in zavzema določeno vrednost pri vsaki ponovitvi vzorca. Dobljeni niz vzorčnih aritmetičnih sredin oz. ocen pa je slučajna spremenljivka in ima posebno porazdelitev, ki jo imenujemo vzorčna porazdelitev.

- kadar velikost vzorca n ni premajhna (dovolj je že 20), je statistična teorija ugotovila, da je porazdelitev vzorčnih aritmetičnih sredin približno normalna, in to ne glede na porazdelitev spremenljivke v osnovni populaciji;
- vzorčne ocene povprečij na vseh vzorcih določene populacije se porazdeljujejo normalno; poleg tega je aritmetična sredina teh povprečij enaka populacijski aritmetični sredini (»centralni limitni izrek« ali »zakon velikih števil«);
- !!! v primeru vzorčne porazdelitve opazujemo porazdelitev vzorčnih aritmetičnih sredin v populaciji vseh vzorcev in ne porazdelitve elementov v osnovni populaciji !!!
- kadar je aritmetična sredina vzorčnih ocen enaka populacijskemu parametru, ki ga ocenjujemo, potem je taka cenilka nepristranska za ocenjevanje populacijskega parametra (tako je pri SRS vzorcu cenilka \bar{y} nepristranska cenilka populacijskega parametra \bar{Y} , cenilka s^2 pa nepristranska cenilka parametra S^2); čeprav je središče vzorčne porazdelitve slučajne spremenljivke \bar{y} enako \bar{Y} , pa se bodo posamezne ocene, ki jih zavzema cenilka v določenem vzorcu, seveda razlikovale od populacijske aritmetične sredine; zato potrebujemo mero, ki bo izražala variiranje posameznih ocen okoli \bar{Y} , torej variiranje slučajne spremenljivke \bar{y} ; v statistiki je splošno sprejeta mera variabilnosti slučajnih spremenljivk **varianca** oz. **standardni odklon** (kvadratni koren variance); $S^2 =$ elementarna varianca; $V(\bar{y}_0) =$ vzorčna varianca; $SD(y) =$ elementarni standardni odklon; $SE(\bar{y}_0) =$ standardni odklon vzorčne porazdelitve oz standardna napaka.;

Varianco cenilke aritmetične sredine v SRS vzorcu (empirični izračun) je mogoče izračunati tudi analitično $V(\bar{y}_0) = (1 - f) * \frac{S^2}{n}$. Tedaj je odvisna od treh dejavnikov:

1. FPC = 1 - f (final population correction), ki kaže, da je populacija končna in ne neskončna, kot predpostavlja klasična obravnava ter da je bilo vzorčenje izvedeno brez ponavljanja \Rightarrow pri neskončni populaciji ali pri vzorčenju s ponavljanjem zato popravek FPC ne nastopa $V(\bar{y}_0) = \frac{S^2}{n}$ (poleg tega je tudi razlika med vzorčenjem s ponavljanjem ali brez njega pri velikih populacijah povsem zanemarljiva, saj je pri izbiranju s ponavljanjem verjetnost ponovnega izbora istega elementa izredno majhna \Rightarrow faktor FPC običajno zanemarjamo oz. poenostavljamo na vrednost 1, če je vzorčni delež manjši kot npr. $f = \frac{n}{N} = \frac{1}{20} = 0,05$); glej str.

19

2. velikost vzorca $n \rightarrow$ velikost variance se manjša z velikostjo vzorca – pri večjem vzorcu je zato vzorčna varianca manjša; ker v izrazu za varianco vrednost N sploh ne nastopa, ostaja pri velikih populacijah velikost N enako nepomembna kot pri majhnih populacijah in se zato vzorčenje veliko bolj izplača v velikih populacijah;

3. elementarna varianca S^2 za spremenljivko y v populaciji (S^2 lahko povsem enakovredno nadomestimo s σ^2 - oba izraza sta populacijska parametra, na katera ne moremo vplivati in sta običajno neznanili količini); ena od prednosti izraza variance, kjer uporabljamo S^2 (in ne σ^2) je dejstvo, da je vzorčna cenilka s^2 nepristranska cenilka za parameter S^2 (ne pa tudi za σ^2);

$$V(\bar{y}_0) \rightarrow v(\bar{y}_0) = (1 - f) \frac{S^2}{n} \quad SE(\bar{y}_0) \rightarrow se(\bar{y}_0) = \sqrt{\frac{(1 - f)s^2}{n}} \quad S^2 \rightarrow s^2 = \frac{\sum (y_i - \bar{y})^2}{n - 1}$$

INTERVAL ZAUPANJA \rightarrow pri vsaki normalni porazdelitvi je 95% vseh elementov znotraj intervala, ki ga v obe smeri začrtuje $\pm 1,96$ standardnega odklona od aritmetične sredine \Rightarrow

interval zaupanja (na vzorcu) vsebuje populacijsko vrednost \bar{y} v 95% vseh primerov vzorcev. S 5 – odstotnim tveganjem trdimo, da interval zaupanja vključuje pravo populacijsko vrednost.

- interpretacija: s 5 – odstotnim tveganjem trdimo, da je populacijska aritmetična sredina v intervalu od $\bar{y}_0 - 1,96 * SE(\bar{y}_0)$ do $\bar{y}_0 + 1,96 * SE(\bar{y}_0)$;

Meje intervala zaupanja so torej slučajna spremenljivka, zato je pomembno, da interval zaupanja, ki opredeljuje populacijski parameter, ločujemo od **VERJETNOSTNEGA INTERVALA**, ki določa gibanje cenilke. Tako bi lahko zapisali, da se npr. s 95% verjetnostjo vzorčna aritmetična sredina nahaja v verjetnostnem intervalu. Za razliko od intervala zaupanja, so meje verjetnostnega intervala fiksne, saj so odvisne samo od populacijskih vrednosti \bar{y} in $SE(\bar{y}_0)$.

- pri statističnem sklepanju je običajno 5% tveganje, pri manj pomembnih ocenah uporabimo tudi tveganje $\alpha = 0,1$ in takrat imamo (namesto vrednosti $z_{\alpha/2} = 1,96$, ki določa širino intervala zaupanja pri $\alpha = 0,05$) manjšo vrednost $z_{\alpha/2} = 1,65$ in zato tudi natančnejšo oceno oz. ožji interval zaupanja (poleg aritmetičnih sredin pogosto ocenjujemo tudi odstotke oz. deleže populacije z določeno lastnostjo; stran 21 – 22);

3. SISTEMATIČNO VZORČENJE

Tako kot pri SRS vzorčenju ima tudi pri sistematičnem vzorčenju vsak element v populaciji enako verjetnost izbora v vzorec. Imamo torej vzorec enakih verjetnosti (EPSEM), ki se od SRS vzorca razlikuje po tem, da verjetnost za vključitev v vzorec ni enaka za vse množice elementov oz. čeprav imajo vsi elementi enako verjetnost za izbor, pa očitno nimajo vsi enake verjetnosti, da se pojavijo v vseh vzorcih. Zato tudi niso vsi vzorci enako verjetni. EPSEM lastnost sistematičnega vzorčenja sicer daje vtis, da je vzorčna aritmetična sredina dobra cenilka populacijske aritmetične sredine, vendar neenake verjetnosti vzorcev pomenijo, da izrazi za izračun standardne napake SRS vzorcev brez dodatnih predpostavk (npr. urejenost elementov v populaciji) pri sistematičnem vzorčenju niso uporabni. Ker imamo pri sistematičnem vzorčenju skromno število elementov za ocenjevanje vzorčne variance, se vzorčne aritmetične sredine zato lahko močno razlikujejo, vzorčna varianca pa je lahko izredno velika (saj so tudi razlike med vzorčnimi deleži izjemno velike). Poleg tega znotraj posameznega vzorca vrednosti sploh ne variirajo in je zato variabilnost enaka 0 ($s^2 = 0$) \Rightarrow na podlagi enega vzorca zato vzorčne variance sploh ne bi mogli oceniti, kar je še dodatna neugodnost. Torej sama ciklična urejenost elementov v seznamu ni problematična, če ne sovпада z večkratnikom vzorčnega intervala; vendar je to le redko, za to se sistematično vzorčenje uporablja brez posebnih zadržkov

4. STRATIFIKACIJA

Vnaprej poznane informacije (tip naselja, število oseb itd.) uporabimo že pri načrtovanju vzorcev in izboljšamo vzorčni načrt, poleg tega pa z njimi lahko v fazi analize izboljšujemo natančnost ocenjevanja (najbolje je, da z dodatnimi informacijami izboljšamo tako vzorčni načrt kot cenilko).

Stratifikacija temelji na predhodni razdelitvi populacije na podpopulacije oz. stratum, kar opravimo s pomočjo dodatnih populacijskih informacij (regije kot stratumi). Vzorec se poteka ločeno in neodvisno v vsakem stratumu posebej (v vsakem stratumu izbiramo elemente povsem neodvisno, torej ne glede na vzorčni delež, s katerim izbiramo končne elemente v stratumih). Ena od prednosti stratifikacije je zato v tem, da so velikosti vzorcev v stratumih nadzorovane in niso določene slučajno v samem procesu vzorčenja \rightarrow pogosto so velikosti vzorca v stratumih enostavne sorazmerne ustrezni velikosti populacije. V takem primeru imamo **PROPORCIONALNO STRATIFIKACIJO** (pogosto se uporablja, ker so cenilke v tem primeru relativno enostavne in ker jamči, da ocene niso manj natančne kot pri SRS vzorcu

enake velikosti), kar pomeni, da v vseh stratumi uporabimo enak vzorčni delež $f = \frac{n_h}{N_h} = \frac{n}{N}$.

Če smo za izbor vzorca uporabili SRS vzorčenje, lahko rezultate SRS vzorčenja uporabimo znotraj vsakega stratuma \Rightarrow cenilke \bar{y}_h so zato nepristranske za \bar{Y}_h in tudi njihove variance in standardne napake lahko ocenimo z znanimi izrazi SRS vzorčenja.

- proporcionalna stratifikacija je EPSEM vzorčni načrt;
- varianca vzorčne aritmetične sredine, ki temelji na proporcionalnem stratificiranem vzorcu, je podobna varianci v SRS vzorcu; razlika je le v tem, da je populacijska elementarna varianca S^2 v izrazu za SRS vzorčenje tokrat zamenjana z utežno aritmetično sredino elementarnih varianc znotraj stratumov S_w^2 ;
- $S^2 \cong S_w^2 + \sum W_h (\bar{Y}_h - \bar{Y})^2$; ker vsota kvadratov (drug člen na desni) ni nikoli manjša od 0, proporcionalno stratificirani vzorec ne more biti manj natančen kot SRS vzorec;
- učinek proporcionalne stratifikacije bo večji pri večjih razlikah med aritmetičnimi sredinami stratumov – torej si pri proporcionalni stratifikaciji prizadevamo za čim bolj heterogene stratumne, ki so notranje kar najbolj homogeni;

Za primerjavo vzorčnih načrtov pogosto uporabljamo izraz vzorčni učinek (*design effect*) – **Deff**. Je splošna mera za primerjavo variance enostavnega slučajnega vzorčenja (SRS) z varianco kompleksnih vzorcev enakih velikosti, kjer primerjamo dve varianci za isto spremenljivko: *Deff* = varianca cenilke kompleksnega vzorca / varianca cenilke SRS vzorca enake velikosti oz. *Deff* meri učinkovitost izbranega vzorčnega načrta v primerjavi z SRS vzorcem);

$deff < 1$ – bolj učinkovit vzorčni načrt (torej stratificiran vzorčni načrt) oz. $deff > 1$ – manj učinkovit vzorčni načrt (torej stratificiran vzorčni načrt).

-

Npr.; $n = 250$ (velja za SRS vzorec in za stratificiran vzorec), ... $deff = 0,87 \Rightarrow$ dosegli smo večjo natančnost s proporcionalno stratifikacijo – da bi dosegli z SRS vzorcem enako natančnost kot pri proporcionalno stratificiranem vzorcu, potrebujemo velikost SRS vzorca $n = 250 / 0,87 = 287$.

- večje, ko so odstotne razlike med stratumi (npr. pod 10% in nad 90%), manjša je izboljšava

Poznamo pa tudi **DISPROPORCIONALNO STRATIFIKACIJO** (OPTIMALNA STRATIFIKACIJA), kjer imajo stratumi različne vzorčne deleže, zato v takem primeru struktura stratumov v vzorcu ne odseva strukture stratumov v populaciji. Cilj – doseči razmestitev, ki kar najbolj poveča natančnost cenilke populacijske aritmetične sredine v okviru finančnih ali drugih omejitev. Bolj heterogeni stratumi in stratumi z nižjimi stroški bodo vzorčeni z višjimi vzorčnimi deleži, s čimer dosežemo optimalno razmestitev.

Neymanova razmestitev – stroški med stratumi se ne razlikujejo (so fiksni), zato velja namesto

$f_h \propto \frac{S_h}{\sqrt{c_h}}$ formula $f_h \propto S_h$ (težava = nepoznavanje S_h in $c_h \rightarrow$ zadostujejo grobe ocene).

Disproporcionalna stratifikacija lahko prinese izjemno velike izboljšave v natančnosti, po drugi strani pa lahko v nekaterih spremenljivkah prinese tudi manj natančne ocene kot SRS vzorec enake velikosti – slednje se pri proporcionalni stratifikaciji ne more zgodite in je zato varna za vse spremenljivke. Pogosto se tako uporablja pri načrtovanju vzorcev, ki zagotavljajo dovolj natančne ocene za vsak stratum – vzorčne ocene namreč pogosto potrebujemo ne le za celotno populacijo, ampak tudi za podpopulacije, ki jih imenujemo *domene proučevanja* (samostojen predmet proučevanja) ter, kadar nas zanimajo primerjave ocen med posameznimi stratumi, pri čemer je velikost populacije v stratumi nepomembna – pomembna postane šele izračunu ocene za celotno populacijo.

Izbiranje stratumov; za učinkovito uporabo postopka stratifikacije morata biti izpolnjena dva osnovna pogoja – znana morajo biti populacijska razmerja W_h , možno mora biti ločeno in neodvisno izbiranje vzorcev po stratumih (če pogoja nista izpolnjena si pomagamo s podstratifikacijo ali dvofaznim vzorčenjem).

- za izračun standardne napake moramo v vsakem stratumu izbrati vsaj dva elementa;
- da dosežemo večjo natančnost cenilke pri ocenjevanju parametrov za celotno populacijo, morajo biti stratumi oblikovani tako, da so glede ciljne spremenljivke notranje čim bolj homogeni, hkrati pa morajo med stratumi obstajati čim večje razlike v aritmetičnih sredinah \bar{Y}_h , kar omogoča učinkovito proporcionalno stratifikacijo, ter v variancah S_h^2 , kar je pogoj za dodatne izboljšave v vzorčni varianci pri disproporcionalni stratifikaciji;
- pri premajhnih stratumih izbiramo elemente s povečanim vzorčnim deležem, da bi dobili velikost vzorca, ki zagotavlja potrebno natančnost;
- merilo pri oblikovanju stratumov je predvsem njihova notranja homogenost, ki ugodno vpliva na standardno napako cenilk v stratificiranih vzorcih – če se kateri od stratumov izkaže za premajhnega, kar pomeni, da bi v njem izbrali manj kot dva elementa, ga lahko združimo s sosednjim stratumom;
- pri majhni velikosti vzorcev zaokroževanje na najbližje celo število ne predstavlja zanemarljivega učinka, zato v tem primeru uporabimo *implicitno stratifikacijo* (seznam najprej uredimo po stratumih in nato sistematično izberemo vzorec), s čimer rešimo problem decimalne velikosti vzorca;

5. SKUPINE IN VEČSTOPENJSKO VZORČENJE

VZORČENJE V SKUPINAH → kadar so v vzorec vključeni vsi elementi izbrane skupine, govorimo o vzorčenju v skupinah (tipični primeri skupin elementov so podjetja, šole, razredi, družine, naselja ipd.). Če pa izberemo v vsaki skupini vzorec elementov, govorimo o *dvostopenjskem vzorčenju* (v skupinah); o večstopenjskem vzorčenju pa govorimo, ko zaporedno vzorčimo na več stopnjah (pri tem uporabljamo hierarhijo skupin – najprej slučajno (SRS) izberemo večje skupine, potem v njih manjše in tako naprej – uvajamo dodatne stopnje vzorčenja, dokler ne pridemo do stopnje izbiranja elementov);

- čeprav so tako stratumi kot skupine množice elementov, pa se njihova vloga v vzorčenju precej razlikuje; vsi stratumi so namreč vedno vključeni v vzorec in običajno povečujejo natančnost, medtem pa za skupine velja: 1. v vzorec ne vključimo vseh skupin, ampak le njihov vzorec, zato morajo izbrane skupine predstavljati tudi preostale (neizbrane) skupine – ugodno je, če so skupine notranje heterogene in ne homogene tako kot pri stratumih; 2. razlika je tudi v vplivu na natančnost – proporcionalna stratifikacija v primerjavi s SRS vzorcem enake velikosti namreč vedno izboljša natančnost, podobno velja tudi za večino drugih postopkov stratifikacije; po drugi strani pa vzorčenje v skupinah običajno manjša natančnost v primerjavi s SRS vzorcem iste velikosti (kljub slabši natančnosti se zaradi velikih prihrankov pri stroških anketiranja vzorčenje v skupinah pogosto uporablja predvsem pri terenskem osebnem anketiranju);

- vzorčni učinek $Deff$ je odvisen od razmerja S_a^2 in S^2 (pri tem predpostavljamo, da imajo skupino enako velikost B); če so skupine notranje bolj homogene kot pri slučajni razmestitvi, bodo aritmetične sredine skupin bolj heterogene, zato bo S_a^2 večji od S^2 / B in $Deff$ bo večji od 1;

$\rho = \frac{-1}{(B-1)}$ predstavlja posebno mero za homogenost skupin, imenovano **interklasna**

korelacija, ki torej meri stopnjo podobnosti elementov v skupinah – velja;

→ kadar so skupine pri velikih populacijah oblikovane povsem slučajno, je zato $\rho = 0$ in $Deff(\bar{y}_c) = 1$, kar je enakovredno SRS vzorcu;

→ negativna vrednost ρ bi pomenila, da so skupine notranje bolj heterogene, kot če so oblikovane povsem slučajno; takšne vrednosti ρ so sicer možne, vendar nadvse redke – ena

takih izjem je npr. spolna struktura populacije; zaradi negativnosti ρ je $Deff < 1$ in vzorčenje v skupinah je v takem primeru natančnejše kot SRS vzorčenje;

→ pozitivna vrednost ρ (pa čeprav majhna – pod 0,15) povzroči, da je $Deff > 1$;

→ $\rho = 1$ se pojavi takrat, ko imajo znotraj vsake skupine vsi elementi enako vrednost in zato je tudi $Deff$ enostavno enak številu elementov v skupini;

$Deff$ je torej odvisen od dveh faktorjev: interklasne korelacije ρ in od velikosti skupin B (velja, da kadar pri enaki stopnji homogenosti zmanjšamo velikost skupine B, se skoraj linearno zmanjša tudi $Deff$ ⇒ torej bo $Deff$ pri manjših skupinah v splošnem nižji);

KONČNE SKUPINE → vzorčni načrt je povsem podoben postopku dvostopenjskega vzorčenja v skupinah. Razlikuje se le v tem, da pri dvostopenjskih vzorcih izbiramo po eno končno skupino iz vsake vzorčene populacijske skupine, medtem ko pri modelu končnih skupin te omejitve ni in bi lahko v določeni skupini med B elementi izbrali dvakrat po b elementov. Opisani način je s tem dober približek dvostopenjskega vzorčenja v skupinah. Pri tem je interklasna korelacija ρ v končnih skupinah, b pa velikost končnih skupin. Kadar so končne skupine izbrane s pomočjo SRS vzorčenja, lahko privzamemo, da je njihova homogenost v grobem enaka kot v celotni skupini.

Velja; vzorčni učinek $Deff$ pada, če se manjša velikost b končnih skupin, to je velikost izbranih podvzorcev v skupinah. Pri fiksni velikosti vzorca $n = ab$ torej manjšanje velikosti podvzorcev (vrednost b) – in hkratio večanje števila skupin (vrednost a) – izboljšuje natančnost ocen za aritmetično sredino.

NATANČNOST IN STROŠKI → da bi dosegli optimalno razmerje števila skupin a in števila elementov b, ki smo jih izbrali znotraj populacijske skupine velikosti B, moramo uravnorežiti natančnost in stroške - optimalni vzorec bo bolj razpršen pri večji homogenosti skupin (večji ρ), večjih stroških na element anketiranja (c) in manjših stroških na skupino (C_a); $C = aC_a + nc$
Večstopenjsko vzorčenje uporabljamo predvsem zaradi stroškovne učinkovitosti, kar je najbolj aktualno pri osebnem anketiranju. Namreč, če je populacija velika in razpršena, je enostopenjski vzorec elementov neprijetno razpršen, medtem ko večstopenjski vzorec koncentrira intervjuje v manjše število lokacij.

Do zdaj smo vedno predpostavljali, da so bile skupine in elementi zaradi enostavnosti izbrani s pomočjo SRS vzorčenja. V praksi pa so lahko skupine in tudi elementi izbrani s stratifikacijo ali sistematičnim vzorčenjem. Izkaže se, da je stratifikacija celo pomembnejša pri vzorčenju v skupinah kot pri vzorčenju elementov, saj omogoča večje izboljšave v natančnosti (v primeru, da iz stratumu izberemo samo eno vzorčno enoto prve stopnje, variance v takem stratumu ne moremo oceniti; vendar pa že v osnovnem načrtu stratifikacije predvidimo po dve enoti v vsakem stratumu, kar imenujemo *stratifikacijsko izbiranje parov*).

6. SORAZMERNO VZORČENJE

(do sedaj smo pri vzorčenju v skupinah predpostavljali, da so vse skupine enake velikosti, kar pa je v praksi izredno redko, saj se elementi pogosto povezujejo v skupine različnih velikosti) Velika variabilnost v končni velikosti vzorca je deloma posledica majhnega števila izbranih skupin. Vendar pa bi končna velikost vzorca variirala tudi pri večjem številu izbranih skupin, če bi obstajale razlike pri njihovi velikosti. Očitno bi torej potrebovali večji nadzor velikosti končnega vzorca. Čeprav le redko potrebujemo povsem natančno velikost vzorca, jo moramo vseeno obdržati znotraj zmernih meja. Ena od možnosti za zmanjšanje variabilnosti je stratifikacija skupin po velikosti, ki nam prinese bistveno večji nadzor nad končno velikostjo vzorca, vendar pa s tem omejimo uporabo stratifikacije po drugih faktorjih.

VZORČENJE PPS → izbiro PPS lahko razmeroma enostavno dosežemo s pomočjo komulativne vsote velikosti 'blokov'. Pri takem postopku je neugodno le to, da ima vsak 'blok'

možnost, da je izbran večkrat, zato za izbiro PPS vzorca brez ponavljanja raje uporabljamo sistematično vzorčenje. PPS izbiro sistematično izvedemo tako, da skupno velikost vzorca razdelimo glede na število enot prve stopnje, s čimer dobimo vzorčni interval določene velikosti (korak), ki ga v nadaljevanju prištevamo.

S PPS vzorčnim načrtom lahko tudi metodo "končnih skupin" (že opisano) \Rightarrow SRS vzorec končnih skupin je enak PPS vzorcu brez ponavljanja (razlika med vzorčnima načrtoma povsem je zanemarljiva, saj je verjetnost izbire dveh končnih skupin iz ene PSU majhna).

Ker ima PPS načrt lastnost EPSEM, je velikost končnega vzorca povsem določena in s tem konstantna $n \Rightarrow$ enostavna vzorčna aritmetična sredina je nepristranska cenilka populacijske aritmetične sredine.

VZORČENJE PPES \rightarrow PPS vzorčenje je le redko izvedljivo, saj so velikosti vzorčnih enot običajno neznane. Pogosto pa lahko dobimo določene ocene velikosti, ki jih nato lahko uporabimo namesto pravih velikosti pri PPS izbiri. Pri tem moramo seveda razlikovati med uporabo resničnih velikosti skupin in uporabo ocenjenih velikosti.

PPS(B_α) – uporabimo, ko razpolagamo z dejansko velikostjo, PPES (M_α) – izraz za izbiro, ki je sorazmerna z ocenjeno velikostjo ("PPES izbira").

-da bi imel vzorec s PPES izbiro lastnost EPSEM, moramo na drugi stopnji zagotoviti ustrezen vzorčni delež;

-pričakovana velikost vzorca bo variirala med enotami prve stopnje in bo enaka želeni velikosti b , samo kadar bosta prava in ocenjena velikost enaki;

-ocenjene velikosti morajo biti čim natančnejše, da ne pride do neskladja (npr. $\sum M_\alpha = 360$, $\sum B_\alpha = 315$);

RAZMERNOSTNA CENILKA \rightarrow posledica PPES vzorčenja je tudi to, da velikost vzorca ni konstantna, ampak slučajna spremenljivka, ki je odvisna od izbire enot prve stopnje. Tedaj aritmetično sredino imenujemo razmernostna aritmetična sredina ali **razmernostna cenilka**, ker je izračunana kot razmerje dveh slučajnih spremenljivk. Vendar ni nepristranska cenilka populacijske aritmetične sredine. Njena pristranskost bo zanemarljiva, če bo variabilnost v velikosti vzorca x dovolj majhna (koeficient variacije x manjši od 0,1).

Izračun variance za razmernostno cenilko je bolj zapleten, ker je v imenovalcu slučajna spremenljivka. Dober približek je mogoče izdelati le za velike vzorce na podlagi Taylorjeve vrste in ko je koeficient variacije x manjši od 0,2, še boljše 0,1.

glej formulo (23); splošni izraz za varianco razmernostne cenilke $v(r)$ velja za vsak EPSEM vzorčni načrt ne glede na verjetnosti, ki se uporabljajo za izbiro PSU, in ne glede na to, kako smo izbrali vzorec na drugi stopnji, tj. znotraj izbranih PSU. Prav tako ga lahko uporabimo pri vzorcih, pri katerih enote prve stopnje (PSU) niso stratificirane in pri katerih se velikost vzorca ne spreminja. Z določenimi prilagoditvami pa lahko izraz uporabimo pri vzorcih, ki nimajo EPSEM vzorčnega načrta.

Problemi sorazmernih vzorcev; 1. določena PSU tako velika, da ima možnost večkratnega izbora \Rightarrow verjetnost za vključitev v vzorec večja od 1; 2. premajhne PSU, kar rešimo z združitvijo s sosednjo enoto in obravnavamo kot eno (tudi v primeru, ko je $M_\alpha = 0$, saj je to le ocenjena velikost, ki morda izhaja iz zastarelih podatkov);

7. DRUGI VERJETNOSTNI NAČRTI

DVOFAZNO VZORČENJE (popis prebivalcev v ZDA) \rightarrow v prvi fazi izberemo elemente v začetni vzorec, v drugi fazi pa v podvzorec izberemo le določene elemente in prve faze (lahko je tudi več faz, vendar sta običajno le dovolj le dve). Dvofazno vzorčenje uporabljamo tudi takrat, ko za različne cenilke potrebujemo različne ocene natančnosti in tudi različne velikosti

zorcev. Z dvofaznim vzorčenjem zmanjšamo stroške in čas zbiranja podatkov ter pomembno lahko tudi zmanjšamo breme anketirancev.

Dvofazni vzorci nastopajo tudi pri uporabi populacijskih podatkov za izdelavo učinkovitega vzorčnega načrta, kadar so stroški zbiranja podatkov na celi populaciji preveliki – prvo fazo uporabimo samo za pridobitev stratifikacijske informacije, za ocenjevanje velikosti enot pri izvedbi PPES vzorčenja ter za oceno stroškov vzorčnega načrta druge faze; na drugi fazi pa anketiramo ustrezne elemente. Pogosto se uporablja pri vzorčenju redkih populacij, za katere ni ločenega vzorčnega okvira (manjšine, vojni veterani, ...).

Pri ocenjevanju učinkovitosti dvofaznega vzorčnega načrta moramo vnaprej poznati stroške izvedbe obeh faz, pri čemer je vzorec v drugi fazi nujno manjši od vzorca v prvi fazi \Rightarrow najučinkovitejši je, če so stroški v prvi fazi bistveno manjši od stroškov v drugi fazi.

VZORČENJE Z REPLIKACIJAMI \rightarrow celoten vzorec je sestavljen iz niza podvzorcev, ki so narejeni na enak način. Vzorec torej obravnavamo kot zaporedje ponovitev povsem enakih podvzorcev. Pri tem je pomembno, da vsak podvzorec generira neodvisno oceno za populacijski parameter. Na tej osnovi tudi izračunamo tudi skupno oceno.

Pri vzorčenju z replikacijami je pomembno ocenjevanje vpliva anketarjev, pri čemer uporabljamo standardno enosmerno analizo variance. Visoka vrednost F pokaže prisotnost vpliva anketarjeve variance, kritična vrednost F je nad 1. dober kazalec variance anketarja je tudi interklasna korelacija.

Posledice variiranja med anketarji so podobne vplivu, ki ga običajno dobimo pri prostorskih skupinah. Na anketiranca namreč v določenem obsegu vpliva tudi anketar. Podobno kot vzorčni učinek $Deff$ pri vzorčenju v skupinah obstaja vpliv anketarjeve variance pri repliciranem SRS vzorčnem načrtu. Podobno kot pri vzorčenju v skupinah lahko celo majhne vrednosti R^2 pomenijo precejšnje povečanje variance v primerjavo s SRS varianco, kadar je število r (velikost podvzorca) anket na posameznega anketarja visoko (običajna ocena variance aritmetične sredine na osnovi SRS vzorca seveda ne vključuje vpliva anketarjev, zato je v primeru, ko podatki to omogočajo, ugodno preveriti izračun vzorčne variance še z metodo replikacij, ki na osnovi variiranja med podvzorci upošteva tudi vpliv anketarja).

Druga uporaba repliciranega vzorčenja je ocenjevanje običajne vzorčne variance, ki jo lahko uporabo pri vsakem parametru, vzorčni načrt podvzorcev pa je lahko tudi kompleksen. Nastaja pa manjša težava, ker izraz ocenjuje varianco aritmetičnih sredin repliciranih vrednosti \bar{z} , ki v splošnem ni enaka siceršnji cenilki z' , ki jo računamo kot enostavno aritmetično sredino vseh elementov v skupnem vzorcu (vendar je razlika med njima zanemarljivo majhna).

Več težav pa povzroča izbira parametra c , tj. število replikacij. Namreč, če izberemo premajhno število, bo ocena variance nenatančna, kar lahko povzroči preširok interval zaupanja za oceno parametra (običajno mora biti vsaj 20 ali 30 replikacij). Po drugi strani pa lahko pri večjem številu replikacij uporabimo manj podrobno stratifikacijo.

\Rightarrow prednost ocenjevanja variance z metodo enostavnih replikacij je predvsem v enostavni uporabi, čeprav se lahko pri tem zmanjša natančnost (če je c – število replikacij – majhen, je natančnost ocene variance manjša zaradi manjšega števila prostostnih stopenj, če pa je c velik, se natančnost zmanjšuje tudi zaradi manj podrobne stratifikacije). Zato enostavnih replikacij v praksi ne uporabljamo pogosto, ampak raje uporabimo ti. psevdoreplikacije, ki omogočajo stratifikacijo in dopuščajo več prostostnih stopenj.

PANELNE RAZISKAVE (na področju delovne sile) \rightarrow raziskave, pri katerih podatke na istih elementih zbiramo po večkrat. Eden od najpomembnejših razlogov za ponovno anketiranje je raziskovanje spremembe v času. Pri tem moramo ločiti med skupnimi spremembami (spremembe na ravni elementov) in spremembami v razliki (spremembe na ravni agregatov). Drugi razlog za izvedbo panelne raziskave se nanaša na zbiranje informacij, ki so dostopne le v določeni časovni točki, ko so anketiranci o njih sposobni natančno poročati (dohodki gospodinjstev, povezava med predšolsko vzgojo otrok in kasnejšim uspehom v šoli).

Paneli ali longitudinalne študije, pri katerih dobivamo podatke o elementih v več časovnih točkah, prinašajo tudi nekaj novih težav; mobilnost elementov v vzorcu, spreminjanje populacije v času, prilagajanje anketirancev → zaradi naštetih težav se vključenost anketiranca v panelu pogosto časovno omeji s ti. rotirajočimi paneli

Panelna raziskava pa lahko pomembno izboljšuje ocenjevanje razlik → če je korelacija visoka, so pridobitve panela pri ocenjevanju razlik lahko precejšnje. Pri rotirajočem panelu pa lahko z uporabo kompleksnih cenilk še dodatno povečamo natančnost. Pri negativnem korelacijskem koeficientu \bar{R} (nakup večje potrošne dobrine) je ocena razlike lahko celo manj natančna. Velja, da vzorčne načrte, pri katerih ohranimo v panelu iste enote prve stopnje (PSU), v njih pa vsakič izberemo druge elemente (vzorec stanovanj namesto gospodinjstev), lahko prav tako uporabimo za izboljšanje ocenjevanja razlik kot, če bi imeli v panelu iste elemente. Razlika je v \bar{R} , ki je v splošnem nižji kot pri primerjavi istih elementov.

8. VZORČNI OKVIRI

Anketne raziskave se pogostokrat ne osredotočajo na natančnost (izjema so mednarodne itd. analize), se pa samo načrtovanje vzorcev v anketni praksi pogosto zoži na vprašanje vzorčnega okvira, tj. seznama (finančne omejitve, predpisane sheme, ki urejajo velikost vzorca), s katerega izbiramo elemente v vzorec. Izbor vzorčnega okvira je zato osrednje vprašanje vzorčenja in hkrati najpomembnejši dejavnik, ki določa kakovost anketne raziskave. Temeljna vloga vzorčnega okvira je identifikacija in lociranje vseh elementov ciljne populacije, omogočiti pa mora tudi še dodatne informacije, ki jih je moč uporabiti pri stratificiranju in načrtovanju večstopenjskega vzorčenja. Idealen vzorčni okvir vsak populacijski element uvrsti na seznam le enkrat, vendar je to le redko, saj nastajajo različna odstopanja, ki jih je Kish strnil v štiri skupine:

1. *manjkajoči elementi*: so eno najresnejših vprašanj in se pojavijo, če je okvir neprimeren, torej da ni namenjen naši ciljni populaciji (izključevanje izrednih študentov) ali pa nepopoln, torej da je ustrezen, vendar ne vključuje vseh elementov (zastareli podatki). Torej manjkajoči elementi so elementi, ki sodijo v ciljno populacijo, vendar nimajo možnosti vključitve v vzorec in zaradi tega ne moremo govoriti o verjetnostnem vzorcu, kjer imajo vsi elementi znano in pozitivno verjetnost za vključitev. Rešitev: postopki povezovanja (v nove skupine), ki potekajo krožno (elemente, ki jih na seznamu povežemo z elementi na seznamu – polodprti interval).

2. *skupine elementov*: vzorčenje skupin lahko pripelje k neugodnemu vzorčnemu učinku *Deff*, še posebno, če je povprečna velikost skupine velika in je interklasna korelacija visoka (če je *Deff* velik lahko uporabimo podvzorčenje znotraj skupin, pomemben razlog za uporabo podvzorčenja je tudi medsebojno vplivanje elementov znotraj skupine, kar prinaša večje stroške in nobene nove informacije). Za izbor anketiranca se pogosto uporablja Kishova metoda izbora osebe, ki daje enake možnosti izbora vsem znotraj izbranega gospodinjstva, saj anketar izbira osebe po urejenem seznamu (še dve metodi sta Troidahl – Carter in Bryant). V primeru, ko so skupine vzorčene EPSEM, samo izbira vseh elementov v skupini daje EPSEM vzorec elementov, torej izbira enega elementa v skupini (s čimer se izognemo medsebojnemu vplivanju in prevelikih skupin ⇒ potrebno uteževanje) ne daje EPSEM vzorec elementov.

3. *neustrezni elementi*: sestojijo iz praznih elementov, ki jih v populaciji sploh ni več in tuje elemente, ki ne sodijo v ciljno populacijo. Rešitev: jih ne upoštevamo, česar posledica je zmanjšanje velikosti vzorca (neustrezne elemente je tako potrebno vnaprej predvideti, vsaj v grobem oceniti njihov delež in ustrezno povečati začetni vzorec). Pogosta napaka, posebej pri sistematičnem vzorčenju, je nadomeščanje neustreznih elementov s sosednjim ali z naslednjim na seznamu, čemur se moramo izogniti. Ločiti pa moramo: položaj, ko lahko prazne elemente vnaprej odkrijemo iz vzorčnega okvira (in jih tako izpustimo) in položaj, ko jih

vnaprej ni mogoče identificirati (potrebno prej še poiskati). Poseben primer neustreznih elementov so ti. redke populacije, kjer uporabimo dvofazno vzorčenje.

4. *podvojeni zapisi*: prihaja, kadar je vzorčni okvir sestavljen iz več seznamov in se nekateri elementi pojavijo na več kot enem seznamu. Rešitev: 1. odstranitev vseh dvojnikov iz celotnega vzorčnega okvira, kar pa običajno ni izvedljivo; 2. uporaba identifikacije, ki povezuje vsak element ciljne populacije samo z enim od njegovih zapisov, in to na enoličen način, preostale zapise tega elementa pa jemlje kot neustrezne elemente; 3. alternativna rešitev je, da sprejmemo vse podvojene elemente in nato uporabimo postopke uteževanja.

Administrativni vzorčni okviri → glej stran 74 – 77

9. NEODGOVORI

Če podatkov od elementov, ki so za raziskavo ustrezni (vedno) in so bili vključeni v vzorec, ne dobimo, imamo opraviti z neodgovori, ki predstavljajo v anketnem raziskovanju resen in naraščajoč problem, saj so anketiranci vse manj pripravljeni sodelovati v raziskavah. Osnovna težava pa je, da se elementi, ki so v anketni raziskavi sodelovali, razlikujejo od ostalih ustreznih elementov, ki v raziskavi niso sodelovali. Tako prihaja do pristranskosti, ki je zaradi sklepanja na osnovi respondentov odvisna od dveh faktorjev: deleža respondentov v populaciji in razlike med aritmetičnima sredinama respondentov in nerespondentov (če sta stratuma respondentov in nerespondentov oblikovana povsem slučajno, sta aritmetični sredini enaki – v takem primeru pristranskost zaradi neodgovorov ne obstaja). Ločiti moramo;

→ *neodgovor elementa* (kadar za izbrani element ne dobimo nobene informacije): kategorije neodgovorov pri osebnem anketiranju; najpogostejša sta *zavrnitev sodelovanja* (strategije reševanja problema neodgovorov – primernejši čas, zamenjava anketarja, nagrade, primerno sestavljen vprašalnik) in *nedostopnost anketiranca* (najobsežnejša kategorija neodgovorov), ostale kategorije pa so *nezmožnost sodelovanja*, *odsotnost anketiranca*, *nezmožnost locirati ustreznega anketiranca* in *izgubljeni ali poškodovani vprašalniki*.

STOPNJE NEODGOVOROV:

Stopnja odgovorov → je razmerje med številom izpolnjenih vprašalnikov in številom vseh ustreznih elementov, ki smo jih vključili v vzorec (oz. stopnja neodgovorov je delež nerespondentov med vsemi ustreznimi elementi). V raziskavi pa moramo ločiti med neodgovori in neustreznimi elementi, pri čemer obstaja težava njihove določitve. Te ocene pa lahko pridobimo iz prejšnjih anket, dodatnih analiz ali na osnovi ekspertne presoje (vsekakor pa je potrebno stopnje neodgovorov vsakič kritično preveriti in ugotoviti, kako so bile izračunane).

Stopnja ustreznosti → ima v števcu skupno število ustreznih elementov, v imenovalcu pa število vseh elementov vključenih v vzorec. Stopnja ustreznosti torej govori o naravi vzorčnega okvira, nikakor pa ne o uspešnosti pri pridobivanju sodelovanja.

Stopnja anketiranja → ima v števcu vse elemente z odgovori, v imenovalcu pa vse elemente vključenih v vzorec. Zaradi tega je izračunavanje te stopnje enostavno, vendar pa je manj informativna in ne izraža uspešnosti pri pridobivanju sodelovanja.

Stopnja nekontaktiranja → izraža delež ustreznih elementov, ki jih nismo uspeli kontaktirati, stopnja kontaktiranja pa je temu komplementarna mera.

Stopnja zavračanja → se nanaša na delež ustreznih elementov, ki so izrecno zavrnilo sodelovanje. Včasih se stopnja zavračanja računa le med kontaktiranimi elementi, kar otežuje primerjave, zato je bolj primerno, da jo imenujemo stopnja nesodelovanja, katere komplementarni produkt je stopnja sodelovanja.

Vsota vseh stopenj za posamezne kategorije neodgovorov daje – zaradi enakega imenovalca (število ustreznih elementov) – skupno stopnjo neodgovorov. Stopnje neodgovorov so predvsem izredno raznolike, odvisne od vrste dejavnikov in se razlikujejo še posebno izrazito glede na način anketiranja.

Delni odgovori; anketiranec prekine pogovor, še preden odgovori na vsa vprašanja – v tem primeru bomo govorili o respondentu, če bo izpolnil dovolj vprašanj, v nasprotnem o nerespondentu.

Kako se boriti proti pristranskosti zaradi neodgovorov? Osnovna strategija je njihovo preprečevanje s pomočjo uteževanja, posebno kadar je mogoče identificirati jasne podskupine oz. razrede z različnimi stopnjami neodgovorov (ker pa pogosto obstajajo razlike med respondenti in nerespondenti znotraj vsake skupine, bo določena pristranskost ostala kljub uteževanju). Uteževanje pa tudi dodatno povečuje vzorčno varianco in se zato uspešnost postopka uteževanja od primera do primera razlikuje. Podobne učinke kot uteževanje imajo tudi postopki vstavljanja nadomestnih elementov, prednost je le v tem, da razpolagamo s popolneno matriko podatkov

→ neodgovor spremenljivke (kjer v fazi anketiranja sicer zberemo določene informacije o elementu, vendar za nekatere spremenljivke nimamo odgovorov): neodgovor spremenljivke se kaže kot manjkajoči podatek v zapisu določenega odgovora (spremenljivke) pri elementu, ki je sicer sodeloval v anketni raziskavi. Obseg neodgovora spremenljivke je različen in variira glede na naravo spremenljivke ter glede na način zbiranja podatkov. Rešitev: omejitev analize samo na elemente z vsemi odgovori (kar pa je vprašljivo, saj lahko odstranimo prevelik del respondentov), postopki vstavljanja, metode neposredne analize, uteževanje (s čimer se zmanjša elementarna varianca in izkrivlja porazdelitev spremenljivke).

»hot deck« postopki vstavljanja → enostavna računalniška obdelava (+), podcenjevanje variance in tudi pripis določene vrednosti večim elementom (-)

!!!ena od najpomembnejših prednosti postopkov vstavljanja je dejstvo, da na tej osnovi razpolagamo s celotno in popolneno matriko podatkov, brez vseh manjkajočih vrednosti!!!

!!!vsaka uporaba vstavljanja vpliva tudi na povečanje vzorčne variance (kot postopki uteževanja) in povzroča širitev intervalov zaupanja!!!

10. ANALIZA RAZISKAVE

UTEŽEVANJE → največkrat je potrebno zato, ker so bili elementi izbrani z različnimi verjetnostmi, pa tudi zaradi drugih razlogov (poststratifikacija, neodgovori itd.).

Neenake verjetnosti izbora: npr. če študenti obiskujejo več kot eno študijsko smer, EPSEM vzorec zapisov ne določa tudi EPSEM vzorca študentov; če so verjetnosti izbora neenake, bo ocena za populacijsko aritmetično sredino pristranska, prav tako pa enostavna ocena aritmetične sredine precenjuje populacijsko aritmetično sredino ⇒ da bi vzorčni načrt popravili, potrebujemo uteži, ki so obratno sorazmerne z verjetnostjo izbora.

→ glej (26e) – koeficient variacije $cv(w)$ se zmanjšuje z večanjem velikosti vzorca, pri dovolj velikem vzorcu ga je mogoče zanemariti; pri velikih razlikah v verjetnostih za izbiro elementov v vzorec se lahko natančnost zelo zmanjša, zato se moramo velikemu variiranju v utežeh izogibati;

Skupine elementov; uteževanje potrebujemo tudi pri odpravljanju težav z vzorčnimi okviri, posebej v primeru skupin elementov, vendar se jih v praksi ne uporablja pogosto, zato predstavljajo le zanemarljiv vpliv na rezultate raziskave.

Disproporcionalna stratifikacija; 1. način – cenilko populacijske aritmetične sredine izračunamo s pomočjo vzorčne aritmetične sredine v vsakem stratumu, nakar ocene kombiniramo v uteženo aritmetično sredino; 2. način – vsakemu izbranemu elementu pripišemo ustrezno utež, ki je tako enaka za vse elemente v istem stratumu, vendar različna med stratumi. Oceni \bar{y}_w in

\bar{y}_{st} sta enakovredni, prednost \bar{y}_w (tokrat ni razmernostna cenilka, ker je imenovalec konstanten) pa je v enostavnosti računanja.

Poststratifikacijske uteži; uporabna je predvsem takrat, kadar poznamo populacijske deleže W_h , vendar elementom ne moremo vnaprej določiti stratuma, ki mu pripadajo in zato ne moremo opraviti predhodne stratifikacije, vendar pa lahko od elementov, ki so bili vključeni v raziskavo, zberemo informacije, ki omogočajo uvrstitev v stratum in s tem uporabo postratifikacije. Postratifikacijo pa lahko uporabimo tudi za izkoriščanje dodatnih stratifikacijskih faktorjev poleg tistih, ki smo jih uporabili pri načrtovanju vzorca.

!!!tako kot pri proporcionalni stratifikaciji se tudi pri postratifikaciji natančnost povečuje, če obstaja glede na ciljne spremenljivke heterogenost med stratumi in homogenost znotraj stratumov!!!

Uteževanje zaradi neodgovorov in nepokritja; podobno kot postratifikacija zahtevajo tudi popravki zaradi neodgovorov oz. nepokritja poznavanje populacijske porazdelitve kontrolnih spremenljivk. Popravke za neodgovore izvedemo z utežmi, ki jih pripišemo vsem respondentom v razredu. Tovrstno uteževanje prilagodi porazdelitev respondentov dejanski porazdelitvi celotnega vzorca znotraj razredov, tako da respondenti v vsakem razredu predstavljajo tudi nerespondente. Vendar takšno uteževanje je moč uporabiti le v primeru neodgovorov, ne pa tudi pri nepokritju.

⇒ uteži uporabljamo za popravke neenakih verjetnosti izbora, neodgovore in prilagoditev vzorčne porazdelitve znani populacijski porazdelitvi (pogosto pa je tudi potrebna kombinacija različnih tipov uteži);

IZRAČUNAVANJE VZORČNIH NAPAK → velikost vzorčne napake je odvisna od velikosti in tudi od kompleksnosti konkretnega vzorca.

Predpostavka o SRS vzorcu;

- pri SRS (enostavno slučajno vzorčenje) vzorcu brez ponavljanja je varianca vzorčne aritmetične sredine manjša od variance aritmetične sredine vzorca SRS s ponavljanjem za faktor končne populacije FPC ; $FPC = 1 - f$ ($f = n/N$) ⇒ pri velikih populacijah je f majhen, zato je FPC praktično 1 in zaradi tega lahko uporabimo izraz za standardno napako na podlagi SRS vzorca s ponavljanjem tudi pri SRS vzorcu brez ponavljanja;

- proporcionalno stratificiran vzorčni načrt s SRS vzorčenjem znotraj stratumov daje cenilke, ki so vsaj tako natančne kot pri SRS načrtu; cenilke so natančnejše, kadar je stratum notranje homogen glede na ciljne spremenljivke; uporaba izraza za standardno napako ob predpostavki SRS vzorčenja bo zato pri proporcionalni stratifikaciji povzročila določeno precenjevanje vzorčne napake; če zanemarimo FPC , se pa precenjevanje še poveča;

- ker je disproporcionalna stratifikacija ne – EPSEM vzorčni načrt, zahteva vsako ocenjevanje parametrov predhodno uteževanje, ki ga bo nevtraliziral in moramo elementarno populacijsko varianco oceniti z uteženo oceno, pri tem pa daje cenilke, ki so bolj ali manj natančne kot pri SRS vzorcu s ponavljanjem enake velikosti; optimalna stratifikacija bo od proporcionalne stratifikacije dala večjo natančnost v primerih, ko elementarna varianca izraziteje variira po stratumih; po drugi strani pa lahko optimalna stratifikacija, ki je prilagojena eni spremenljivki, povzroči podcenjevanje standardne napake pri ocenjevanju preostalih spremenljivk; optimalno stratifikacijo pogosto uporabimo, ko potrebujemo ločene ocene za določene domene, ki pa so pogosto vzorčene z višjimi vzorčnim deležem, s čimer se zagotovi zadostna velikost posamezne domene – vendar pa to vodi v zmanjšanje natančnosti ocene na celotnem vzorcu v primerjavi s SRS vzorcem enake velikosti (če zanemarimo učinek FPC , bo optimalna stratifikacija oz. ustrezno uteževanje povečalo varianco aritmetične sredine celotnega vzorca, v primerjavi s SRS vzorcem s ponavljanjem enake velikosti);

- vzorčenje v skupinah privede do zmanjšanja natančnosti v primerjavi s SRS vzorcem enake velikosti, če je koeficient interklasne korelacije ρ pozitiven, odvisno pa je tudi kakšne

so povprečne velikosti podvzorcev (če je povprečna velikost skupin velika, se natančnost zmanjša, kljub relativno nizki vrednosti ρ);

- kompleksni vzorčni načrt daje precej manj natančne cenilke kot SRS vzorec s ponavljanjem enake velikosti oz. vzorčni učinek $Deff$ je običajno večji od 1;

Prednosti predpostavke o uporabi vzorčenja s ponavljanjem sta bistveno zmanjšanje računanja in večja splošnost uporabe, saj en sam program zadošča za izračun standardne napake, ne glede na način (obliko) podvzorčenja znotraj enote prve stopnje (PSU).

Taylorjeva vrsta; ali linearizacija oz. metoda delta, katere uporaba za izvedbo je enostavna, osnovna ideja pa je v zapisu cenilke v poenostavljeno linearno obliko, nakar lahko uporabimo postopke za ocenjevanje variance pri (bolj enostavnih) linearnih cenilkah (to je cenilka, ki se izraža kot linearna kombinacija vrednosti spremenljivk v vzorcu, npr. aritmetična sredina).

Uporaba Taylorjeve metode omogoča tudi ocenjevanje variance za cenilko razmerja (tudi regresijski in korelacijski koeficient ipd.), ki je eden od najpogostejših primerov nelinearnih statistik (KV mora biti manjši od 0,1 oz. 0,2).

Metoda BRR (metoda uravnoveženih ponovljenih replikacij ali replikacije polovičnega vzorca oz. replikacije polvzorca); pri tej metodi obe izbrani PSU iz vsakega stratuma obravnavamo, kot da sta vzorčeni povsem neodvisno in ga zato obravnavamo v dveh replikacijah: 1. naključno izbrana PSU iz vsakega stratuma; 2. preostale PSU;

Temeljna značilnost ponovljenih replikacij: lastnost uravnoveženosti se nanaša na način, kako so izbrani polvzorci, pri tem pa skupno število polvzorcev ni izbrano povsem slučajno in neodvisno, ampak na poseben in uravnovežen način, ki daje najučinkovitejšo cenilko variance. Za doseg čim boljšega rezultata, mora biti število izbranih polvzorcev večje ali enako številu stratumov, poleg tega pa tudi mnogokratnik števila štiri.

Metoda Jack – knife; je pomembna tehnika za ocenjevanje variance pri kompleksnih vzorčnih načrtih. Gre za ponavljanje določenih replikacij, vendar jih tokrat generiramo z opustitvijo samo ene PSU, preostale PSU v stratumu pa utežimo, da obdržimo osnovno porazdelitev vzorca po stratumih (pri tem je pomembno, da iz vsakega stratuma izpustimo vsaj eno PSU, saj v nasprotnem varianca teh stratumov ne bo zastopana v celotni oceni variance).

Kadar se odločamo za eno od metod ocenjevanja variance, upoštevamo prednosti zmanjšanega obsega računanja, razpoložljivost programov in praktično primernost za ocenjevanje v konkretnem vzorčnem načrtu; Taylorjevo metodo največkrat uporabljamo pri relativno enostavnih cenilkah, BRR in JRR pa je moč brez težav uporabiti tudi za najbolj zapletene ocene. Metoda BRR je sicer omejena na stratificirani vzorčni načrt z dvema notama na stratum, kar pa v praksi – z nekoliko prilagoditve – ustreza večini vzorcev. Pri vzorcih, ki niso stratificirani pa lahko uporabimo JRR (prednost JRR pred BRR je v tem, da brez težav obravnava tudi vzorčne načrte, ki v stratumih niso izbirali po dve PSU, $a_h = 2$).

UTEŽI IN VZORČNA VARIANCA → ena od pomembnih in nadvse neugodnih posledic, ki jih ima uteževanje za statistični analizo, je povečanje vzorčne variance.

$VIF = 1 + CV^2_{(w)} \Rightarrow$ faktor povečanja variance, ki izraža povečanje vzorčne variance uteženega vzorca v primerjavi z vzorčno varianco vzorca enake velikosti, kjer nismo potrebovali uteži (v primeru, ko imamo varianco v disproporcionalno stratificiranem vzorcu - kjer je bilo potrebno uteževanje zaradi neenakih verjetnosti vključitve elementov v vzorec – VIF izraža primerjavo z vzorčno varianco pri predpostavki proporcionalno stratificiranega vzorca enake velikosti, kjer ni bilo potrebno uteževanje).

SREDNJA KVADRATNA NAPAKA → povečanje vzorčne variance lahko razumemo kot ceno, ki jo moramo plačati zaradi uporabe uteži. Vendar pa po drugi strani z utežmi zmanjšujemo pristranskost. Koliko pa je uteževanje v resnici koristno, lahko ocenimo šele z izračunom

srednje kvadratne napake, ki je vsota kvadrata pristranskosti (razlika med pričakovano vrednostjo na osnovi ponavljajočih vzorcev ter pravo populacijsko vrednostjo parametra) in variance: $MSE(\bar{y}) = Bias^2(\bar{y}) + Var(\bar{y})$.

Uteževanje (primerjava utežene ocene z neuteženo oceno) je na osnovi kriterija srednje kvadratne napake upravičeno le v primeru, ko se je le – ta, zaradi učinkovitega zmanjševanja pristranskosti, z uteževanjem zmanjšala.

Učinkovitost opravljenega uteževanja torej presodimo s pomočjo primerjave MSE za uteženo in neuteženo cenilko. V praksi pa se pogosto dogaja, da se z uteževanjem MSE poveča, kar pomeni, da uporaba uteži ni bila smiselna. Obsežno in kompleksni uteževanje namreč nemalokrat znatno poveča vzorčno varianco, ne daje pa izboljšav v pristranskosti.

!!!Lastnost cenilke z majhno MSE označujemo kot točnost. Točna cenilka torej poleg komponente natančnosti, ki se nanaša na majhnost vzorčne variance, vključuje še komponento majhne pristranskosti. V primeru večje pristranskosti so zato cenilke lahko natančne, niso pa tudi točne, saj je MSE zaradi pristranskosti še vedno lahko zelo velika!!!

!!!Pojmov točnosti in natančnosti ne gre zamenjavati z lastnostmi anketnega merskega instrumenta, kot sta veljavnost (z anketnim vprašalnikom dejansko merimo tisto, kar naj bi merili oz. smo nameravali meriti) in zanesljivost (lastnost meritve, da v enakih ponovitvah daje kar najbolj enake rezultate)!!!

11. VELIKOST VZORCA

Dva običajna postopka določanja velikosti vzorca: na podlagi standardne napake in na podlagi koeficienta variacije.

VELIKOST VZORCA IN STANDARDNA NAPAKA → za opredelitev velikosti vzorca moramo najprej opredeliti, kakšno natančnost potrebujemo, kar pa ni lahka naloga, saj hitro pride do precenitve.

Najprej vzamemo cenilko, ki bo vključevala populacijsko vrednost v intervalu in določimo tveganje. Pri enostavnem slučajnem vzorcu (SRS) on ob neupoštevanju faktorja FPC sledi:

$n' = \frac{PQ}{SE(p)^2}$, kjer lahko v grobem ocenimo dopustne vrednosti za populacijski delež →, v nasprotnem pa upoštevamo $P = Q = 50\%$ (P izberemo čim bližje 50%).

Če je začetna velikost vzorca v primerjavi z velikostjo populacije majhna, lahko zanemarimo faktor FPC in s tem bi izračun n' že dal ustrezno velikost vzorca, v nasprotnem pa faktorja FPC ne moremo zanemariti.

Za druge oblike vzorčnih načrtov pa so ocene bolj zapletene: v primeru, da obstaja vzorčni učinek $Deff$, se bo potrebna velikost SRS vzorca povečal še za faktor $Deff$, ki izraža razmerje, za katero je kompleksni vzorec manj učinkovit od SRS vzorca. V primeru npr. stratifikacije pa je lahko $Deff$ tudi manjši od 1, kar v primerjavi s SRS vzorcem zmanjšuje zahtevano velikost vzorca.

- s širitvijo intervala zaupanja se velikost vzorca zmanjšuje;
- po večini se velikost vzorca določa predvsem na podlagi grobe ocene stroškov in nadvse približnega razmerja z natančnostjo;
- pri velikosti vzorca pa moramo predvideti tudi stopnjo anketiranja, ki izraža delež respondentov med vsemi elementi, ki jih vključimo v vzorec;
- pomemben dejavnik pri določanju velikosti vzorca je tudi zagotovitev ustrezne natančnosti z ocenjevanje razlik med podskupinami;
- po določitvi velikosti vzorca, je običajno potrebno določiti še vzorčni delež in vzorčni korak oz. vzorčni interval;

VELIKOST VZORCA IN KOEFICIENT VARIACIJE:

Elementarni koeficient variacije; predstavlja koristen nadzor nad relativno razpršenostjo osnovne spremenljivke (in tudi enostavno primerjavo relativnih razlik) in se na področju družboslovja giblje med 0,2 in 0,5, medtem pa je pri binomski porazdelitvi, ki je osnova za porazdelitev deležev, bistveno večji (tudi 9, saj velja manjši ko je delež večji je koeficient variacije) \Rightarrow za ocenjevanje deležev potrebujemo bistveno večje vzorce kot pri ocenjevanju parametrov intervalnih spremenljivk.

Koeficient variacije za cenilke; opazujemo npr. razpršenost aritmetičnih sredin (tj. vzorčnih ocen v populaciji vseh vzorcev). Podobno kot daje elementarni koeficient variacije informacijo, kje se nahaja glavnina elementov v populaciji, omogoča koeficient variacije za cenilko vpogled, kje je glavnina vzorčnih ocen.

Majhni deleži; kadar je P dovolj majhen ($P < 5\%$), nastanejo pri izračunu koeficienta variacije cenilke p znatne poenostavitve, pri tem pa je osnovna spremenljivka porazdeljena binomsko, koeficient variacije pa je neodvisen od n , N in tudi P .

Bernoullijeva porazdelitev – določa verjetnost, da se v n ponovitvah dogodek z verjetnostjo P zgodi natanko k – krat. Ta porazdelitev pa pri majhnih verjetnostih P in pri velikem številu poskusov (velikost vzorca n) preide v Poissonovo porazdelitev, ki ima en sam parameter λ , saj je elementarna varianca te porazdelitve enaka pričakovani vrednosti. Poissonova porazdelitev je posebej uporabna v primerih, ko velikosti vzorca ne poznamo in se pogosto uporablja kot teoretični model porazdelitve pri analizi nominalnih spremenljivk.

Koeficient variacije in interval zaupanja; ker je vzorčna porazdelitev za cenilke pri velikih vzorcih vedno normalna, lahko iz koeficienta variacije podrobno sklepamo o njeni porazdelitvi, razpršenosti in s tem tudi o kakovosti cenilke.

Določanje velikosti vzorca; želeno širino intervala zaupanja oz. standardno napako pri izbiri vzorca lahko izrazimo tudi s koeficientom variacije, ki ima v primerjavi s kriterijem absolutne širine intervala zaupanja pomembno prednost, da ni odvisen od aritmetične sredine, poleg tega pa nam lahko poda razlike v natančnosti (manjši ko je delež, slabša je natančnost) \Rightarrow v praksi je zato pri načrtovanju velikosti vzorcev bolje izhajati iz koeficienta variacije in ne iz absolutne širine intervala zaupanja oz. iz standardne napake za cenilko.

Kritične vrednosti koeficienta variacije; $CV(\bar{y}) < 0,20$ – ocena, ki jo še objavimo brez posebnega opozorila; $CV(\bar{y}) < 0,33$ – ocena, ki jo objavimo le informativno, z opozorilom; $CV(\bar{y}) > 0,33$ – ocena, ki se načeloma ne objavlja;

INTERVALNE SPREMENLJIVKE \rightarrow obravnavamo jih kot normalno porazdeljene spremenljivke in da vsebuje tipično vrednost standardnega odklona $1/3$.

- v primeru dveh enako velikih in neodvisnih vzorcev za proučevanje razlike aritmetičnih sredin zadošča že velikost, ki je za faktor 1,4 večja od velikosti, ki jo potrebujemo pri siceršnjem ocenjevanju aritmetične sredine;
- ker v mnenjskih anketah npr. zadovoljni anketiranci običajno ostajajo zadovoljni in obratno, je pogosto povsem realno pričakovati vrednosti $R > 0,5$ in s tem zožitev intervala za najmanj 30% \Rightarrow z uporabo panela bi tako dosegli enako natančnost z vzorcem, ki je bistveno manjši;
- predpostavko o normalni porazdelitvi je potrebno vsakič preveriti in to nemalokrat še predno imamo podatke, pri čemer lahko uporabimo porazdelitve podobnih spremenljivk v že opravljenih anketah; če pa imamo spremenljivko, o kateri nimamo predhodnih informacij, je priporočljivo izvesti poskusno (pilotno) anketo;
- končna velikost vzorca je v največji meri odvisna od najpomembnejše ciljne spremenljivke; če je le – teh več, si izberemo spremenljivko, ki zahteva največji vzorec ali pa vrednost mediane ciljnih spremenljivk;

NOMINALNE SPREMENLJIVKE → pri ocenjevanju deležev potrebujemo bistveno večje vzorce kot pri normalno porazdeljenih spremenljivkah. V družboslovju pa se pogosto srečujemo tudi z deleži (poseben primer aritmetičnih sredin), ki omogočajo 'razne', ki nastanejo v primeru nominalnih spremenljivk.

12. KRITIČNA VELIKOST PODSKUPIN

Pri majhnih vzorcih in pri t – porazdelitvi koeficient variacije izgubi svojo siceršnja vlogo za opredeljevanje natančnosti cenilk in tudi za določanje intervalov zaupanja, saj vzorčna porazdelitev ni več normalna. Tako se bo relativna širina intervala zaupanja (20% vrednosti ocene) ujemala z mejno vrednostjo koeficienta $CV(\bar{y}) = 0,10$ samo, kadar je vzorčna porazdelitev normalna.

Pri velikih vzorcih velja centralni limitni izrek, zato je vzorčna porazdelitev normalna in neodvisna od porazdelitve obravnavane spremenljivke. Normalnost vzorčne porazdelitve omogoča tudi enostaven in standardiziran način izračunavanja intervalov zaupanja ter poenostavljene postopke preverjanja domnev. Po drug strani pa pri majhnih vzorcih centralni limitni izrek ne velja in vzorčne porazdelitve zato večinoma niso normalne, ampak so vsakič drugačne, odvisne od konkretne porazdelitve spremenljivke v osnovni populaciji in od velikosti obravnavanega vzorca. Statistično sklepanje na osnovi majhnih vzorcev je zato bistveno bolj zapleteno.

Poudariti velja, da ostra in enotna meja, ki bi ločevala velike in majhne vzorce, ne obstaja. Pri ocenjevanju aritmetične sredine oz. deležev omenjamo $n = 30$, kar pa ne velja za ocenjevanje drugih parametrov (pri nelinearnih parametrih – elementarna varianca – so za prehod vzorčne porazdelitve v normalno obliko potrebni nekoliko večji vzorci). Pri $n < 30$ uporabimo namesto t – porazdelitve t – porazdelitev (čeprav so pri 10 – 30 elementih razlike med porazdelitvama še vedno razmeroma majhne).

Velikost vzorca pa lahko določimo tudi samo na podlagi kritične velikosti podskupin v vzorcu. Pri tem v prvem koraku izračunamo potrebno velikost vzorca za podskupine, v katerih želimo dosegati predvideno natančnost, v drugem koraku pa zahtevane velikosti podskupin povežemo v skupni vzorec; npr. zanimajo nas podskupine določene s spolom, starostjo (3 kategorije) in petimi regijami; imamo torej $2 * 3 * 5 = 30$ podskupin, v katerih moramo dosegati kritično velikost n^0 ; če so vse populacijske podskupine enako velike in smo se odločili za najmanjšo kritično velikost $n^0 = 10$, potem potrebujemo skupni vzorec velikosti $n = 300$ elementov. Opisani pristop daje nadvse praktičen napotek za hitro in enostavno določanje velikosti vzorca. Vendar pa običajno podskupine niso enako velike, zato moramo v tem primeru v najmanjši celici zagotoviti določeno minimalno število elementov.

13. PROSTORSKI VZORCI

Kadar pri izdelavi verjetnostnega vzorca nimamo seznama elementov ciljne populacije, pogosto uporabimo prostorske vzorce (vzorčenje oseb, gospodinjev, stanovanj, kmetij itd.), katerega uporaba je enostavna, saj sledimo zaporednemu verjetnostnemu izbiranju prostorskih enot, na zadnji prostorski enoti pa enostavno popišemo vsa stanovanja in izberemo v vzorec ustrezne elemente.

Prostorski vzorci v ZDA predstavljajo osnovni način verjetnostnega vzorčenja za raziskave, ki se izvajajo z osebnim anketiranjem (vodilni organizaciji stan SRC in NORC). V prostorsko velikih državah se pogosto srečujemo s ti. *krovnim vzorcem*. Namreč, kadar je izdelava prostorskih vzorcev zaradi razprostranjenosti ozemlja in drugih razlogov izredno draga ali

zapletena, je ugodno na daljši rok (npr. 10 let) izdelati en sam večji krovni vzorec, iz katerega se bo za vsako posamezno raziskavo izbirali elementi (enote prve stopnje, za drugo stopnjo izbiramo glede na dane potrebe) brez dodatnega terenskega dela za potrebe vzorčenja. Glavna prednost prostorskih vzorcev je, da ne zahtevajo točnih podatkov o velikosti prostorskih enot.

Prostorski vzorec AKP (Anketa o kadrovskem potencialu); v primerjavi z vzorčenjem iz registra prebivalstva je dražji za 30% in zamuden, poleg tega pa je njegova izdelava tako težavna, da se pogosto odločamo za številne poenostavitve (opustitev oštevilčenja stanovanj). Po drugi strani pa ima določene prednosti: zajame vse osebe, ki prebivajo v Sloveniji, ne glede na državljanstvo in formalni/ administrativni status njihove prijave.

Vidiki prostorskih vzorcev v Sloveniji:

- izbor prostorskih enot; - število stopenj; - enote prve stopnje (PSU)/ $CV > 0,2 \Rightarrow$ enote niso priporočljive/; - število enot prve stopnje; - število PSU v populaciji in v vzorcu; - prostorske enote zadnje stopnje (vzorčne enote – povezani popisni okoliši); - število elementov na zadnji stopnji (5 – 30); - ocenjene velikosti prostorskih enot; - viri podatkov o prostorskih enotah.

Ločiti moramo:

1. poenostavitve na zadnji stopnji vzorčenja (verjetnostno vzorčenje je izvedeno vse do zadnje prostorske enote /npr. popisnega okoliša/, kjer se opušča verjetnostna izbira stanovanj; namesto tega se uporablja enostavno izbiranje vsakega k – tega elementa, kvotna izbira... nobena poenostavitev ne zagotavlja verjetnostnega vzorca);
2. poenostavitve na prvi stopnji vzorčenja (npr. opustitev PPS \Rightarrow izberemo preveč majhnih naselij, majhnih ulic itd. oz. enot, v katerih nato izbiramo fiksno število elementov; resno in nepredvidljivo napako lahko tudi povzroča izbor enot, za katere nimamo podatkov o verjetnosti izbora).

Odstopanje od načel verjetnostnega vzorčenja se pri prostorskih vzorcih pogosto formalizirajo v ti. *vzorčenje slučajne poti*, kjer naprej določimo slučajne začetne točke, v katerih se anketar giblje po točno določenih pravilih slučajne poti in pri tem izbira elemente v vzorec. Vendar obstajata dva problemska vidika: kot prvič je tu izbira slučajnih začetnih točk vzorčenja, saj le – te določajo tudi prostorske enote; kot drugič pa je uporaba metod pri izbiri slučajnih točk (namreč, enako verjetnost izbora oseb imajo tako enostanovanjske hiše s 5 osebami kot neka stolpnica s 500 osebami \rightarrow rešitev: povečamo verjetnost izbora večstanovanjskih stavb ali povečamo število izbranih stanovanj ali pa uporabimo uteži).

14. TELEFONSKI VZORCI

Telefonsko anketiranje predstavlja zaradi vrste razlogov (postopki vzorčenja, ko omogočajo razmeroma enostavno izdelavo verjetnostnih vzorcev) najpogostejši način anketiranja.

Waksberg – Mitofski postopek (W – M): izhajamo iz skupin oz. blokov telefonskih števil (prvih 8), ki predstavljajo enote prve stopnje (PSU) v dvostopenjskem načrtu, zadnji dve cifri pa variirata.

Vsaka uporaba dvostopenjskega vzorčenja pomeni tudi manj natančne ocene, saj je odgovarjajoči vzorčni učinek *Deff* skoraj vedno večji od 1. dvostopenjsko vzorčenje je upravičeno le, če z njegovo uporabo prihranimo glede na stroške in natančnost (običajno pa je *Deff* majhen v primerjavi s prihranki zaradi manjšega števila neuspešnih klicev).

Ker se rezidenčne telefonske številke gospodinjev gostijo v skupine, sta namreč pri W – M postopku približno dve od treh izbranih številki tudi rezidenčni telefonski številki, pri enostavni SRS shemi slučajnega generiranja pa je npr. med petimi le ena številka. Prihranek pri stroških

zato v opisanem W – M postopku običajno presega minimalno izgubo v natančnosti zaradi vzorčnega učinka.

Najenostavnejša možnost pa je enostavna slučajna (SRS) izbira blokov, v drugem koraku sledi izbira ene same telefonske številke iz vsakega bloka. Za vse izbrane telefonske številke se nato preizkusi z avtomatiziranim preverjanjem (tj. postopek vnaprejšnje izbire) njihova dejanska aktivnost, kjer se odstrani vse javne poslovne številke. Opisani postopek zagotavlja EPSEM vzorec, saj gre za običajni dvostopenjski PPS vzorec, pri katerem na drugi stopnji izbiramo le en element. Glavna prednost tega postopka pred klasično W – M metodo (ki traja bistveno dlje) je dejstvo, da ne potrebujemo sekvenčnega (zaporednega) preverjanja rezidenčnosti. Pri W – M postopku se temu izognemo s fiksno izbiro števil v vsakem bloku (fiksno število klicev), vendar končamo s precej različnim številom rezidenčnih števil v posameznih blokih, kar nadomestimo z uteževanjem, pri tem pa se moramo sprijazniti z znatnim povečanjem vzorčne variance zaradi uteževanja (ker pa gre za variabilne uteži, je povečanje manjše kot pri fiksni utežeh).

Vzorci iz tiskanih imenikov:

- problem neažurnosti
- predvideti vnaprej oceno deleža poslovnih števil in za ustrezní odstotek povečati izhodiščni vzorec (nerezidenčne in neustrezne elemente potrebno spustiti)
- poenostavitev: zaporedna izbira strani z določenim korakom in nato na izbranih straneh opravimo nadaljnjo stopnjo vzorčenja oz. izberemo določeno število cifér
- vsem poenostavitvam je skupno, da prinašajo večja ali manjša odstopanja od verjetnostnega vzorčenja in s tem bolj ali manj neznano nevarnost pristranskosti ocen

Vzorci iz elektronskih imenikov:

- poslovne številke so označene, zato njihova izločitev ne povzroča nobenih težav
- postopki slučajnega generiranja telefonskih števil (RDD) – računalniško generiranje števil – dajejo zaradi velikega števila praznih števil le izkoristek pod 20% (le vsaka 5 je rezidenčna); nekoliko boljši izkoristek dajejo metode variiranja zadnjega števila

Vidiki telefonskega vzorčenja: - velikost gospodinjstva; - število telefonskih priključkov v gospodinjstvu (v primeru dveh priključkov je potrebna utež $W = 1/2$); - izbor oseb v gospodinjstvu (problem kvot: ne anketiramo nikogar, saj so prisotne le osebe, ki pripadajo skupinam, katerih kvote smo že zapolnili, poleg tega pa kvote onemogočajo izračunavanje pravih stopenj odgovorov); - število klicev; - število anketarjev (majhno število anketarjev je lahko, zaradi vpliva svoje osebnosti na anketiranca, velik vir nenatančnosti); - neodgovori; - številke brez odziva; - uteževanje rezultatov (po eni strani zagotovi videz »reprezentativnega« vzorca, po drugi strani pa popravi ocene, ki so zaradi porušene sociodemografske strukture v telefonskih vzorcih nekoliko pristranske, poleg tega pa nam uteževanje daje koristno informacijo o smeri popravka);

15. NEVERJETNOSTNO VZORČENJE

V praksi so nadvse razširjeni, saj so mnogo cenejši in izvedbeno hitrejši, ocene pa se nemalokrat ne razlikujejo od ocen iz verjetnostnih vzorcev.

!!! Glavna prednost **verjetnostnega vzorčenja** je v mehanizmu verjetnostnega izbora, ki je predpogoj za uporabo statistične teorije. Tako lahko vnaprej izberemo cenilke z majhno ali celo brez pristranskosti. Statistična teorija omogoča pri verjetnostnih vzorcih tudi izračunavanje natančnosti, oblikovanje intervalov zaupanja in preverjanje domnev. Ker pa za **neverjetnostne vzorce** teorija statističnega sklepanja ne velja, pri teh vzorcih ne moremo izračunati intervalov zaupanja niti ne moremo preverjati domnev. Ocenjevanje kakovosti neverjetnostnih vzorcev je zato subjektivne narave, saj porazdelitev vzorčnih ocen ne temelji na statističnih zakonitostih.

Zato tudi v primeru, kadar izkušnje kažejo, da je bila neverjetnostna metoda uspešna, to še ne zagotavlja, da bo takšna ostala še naprej!!!

Priložnostno vzorčenje je ena od najenostavnejših oblik neverjetnostnega vzorčenja, ki pa je zaradi precejšnje verjetnosti pristranskosti, rezultate tovrstnih raziskav tvegano uporabljati za sklepanje o celotni populaciji.

Druga oblika neverjetnostnega vzorčenja je strokovni izbor ali ekspertna izbira, kjer vzorec, ki naj bi bil »reprezentativen«, izbere strokovnjak, ki pozna vsebino raziskovanega področja (statistika cen – v prvi stopnji izberejo v vzorec večje oz. tipične kraje, na drugi stopnji pa še trgovine), vendar spet obstaja možnost pristranskosti. Ekspertna izbira je primernejša kot verjetnostna izbira, ko izbiramo le en ali dva elementa: če je vzorec majhen, bo varianca cenilke v verjetnostnem vzorcu velika in bo zato pristranskost cenilke vzorca na podlagi ekspertne izbire lahko zanemarljiva; ko pa se velikost vzorca povečuje, se varianca cenilke verjetnostnega vzorca zmanjšuje, netočnost cenilke v ekspertni izbiri pa se ne spreminja (za manjše vzorce je primernejša ekspertna izbira, za večje vzorce pa verjetnostno vzorčenje).

Najpogostejša oblika neverjetnostnega vzorčenja pa je kvotno vzorčenje, katerega bistvo so kvote oz. fiksno število oseb z določenimi značilnostmi, ki jih morajo anketarji vključiti v vzorec, torej ni verjetnostnega izbora. Cilj uporabe kvot je nadzorovanje pristranskosti pri anketarjevem izboru oseb v vzorec. Kvote so lahko medsebojno povezane ali pa neodvisne (glej str. 171).

Pomembna izboljšava kvotnega vzorčenja je njegovo kombiniranje z verjetnostnim vzorčenjem, ki se običajno izvede pri izboru enot prve stopnje. Do razlike pride šele v zadnji fazi, saj pri verjetnostnem vzorcu anketiramo elemente izbrane na podlagi verjetnostnega mehanizma, pri kvotnih vzorcih pa anketarji zgolj izpolnjujejo svoje kvote (tipično navodilo pri kakovostnih kvotnih vzorcih zato opredeljuje postopek iskanja ustreznih anketirancev v verjetnostno izbranih blokih na podlagi enolično določene začetne točke in nadaljnjega točno predpisanega gibanja slučajne poti).

Kvotne skupine so na videz podobne stratumom, saj imamo v obeh primerih populacijske skupine elementov, iz katerih izbiramo element v vzorec. Razlika je, da so pri stratumih elementi izbrani z verjetnostnim mehanizmom, pri kvotnih skupinah pa ne.

Pri kvotnem vzorčenju je pozornost usmerjena v skupine, ki zagotavljajo, da je pristranskost pri izboru čim manjša (pri stratifikaciji to ni potrebno). Za doseg tega se morajo kvotne skupine med seboj razlikovati predvsem glede dosegljivosti in pripravljenosti na sodelovanje v anketi.

Razloga za uporabo kvotnega vzorčenja: prvi razlog je v poenostavitvah, ki izhajajo iz tega, da za izbor anketirancev ne potrebujemo vzorčnega okvira ciljnih oseb, drugi razlog pa je v zmanjšanem številu potrebnih stikov.

Kvotno vzorčenje izvaja vse več organizacij. Vendar velja, da kadar je kvotno vzorčenje poceni in hitro, potem je navadno tudi slabo. Kadar pa je bolje opravljeno, ni veliko cenejši od verjetnostnega vzorčenja. Boljše kvotne metode pa se uporabljajo vselej, ko je rezultate mogoče preveriti z zunanjimi viri. Cenovne prednosti kvotnega vzorčenja se vsekakor maščujejo v dodatni pristranskosti (ki jo pogosto sploh ni mogoče preveriti, saj največkrat proučujemo pojave, o katerih nimamo kontrolnih populacijskih podatkov) in nenatančnosti vzorčnih ocen. Postavlja pa se tudi vprašanje o vzorčni varianci, saj le te ni moč računati.

⇒ pri presoji o morebitni uporabi kvotnih vzorcev moramo pazljivo pretehtati vse bistvene vidike kvotnih in verjetnostnih vzorcev: čas, stroške in obseg odgovarjajočih napak.