



REPUBLIKA E SHQIPËRISË
INSTITUTI I STATISTIKAVE

Udhëzimet metodologjike për vrojtimet me zgjedhje



REPUBLIKA E SHQIPËRISË
INSTITUTI I STATISTIKAVE

Tabela e përmbajtjes

I. Hyrje	6
II. Qëllimi.....	6
1. Rishikimi i udhezimeve	7
2. Vrojtimet statistikore	8
2.1 Hapat e vrojtimeve me zgjedhje	9
2.1.1 Objektivi i vrotimit apo qëllimi i vrotimit.....	10
2.1.2 Regjistrat dhe rëndësia e tyre në procesin e Vrojtimeve me Zgjedhje.	12
2.1.3 Përcaktimi i madhësisë së kampionit.....	12
2.1.4 Hartimi i Pyetësorit	13
2.1.5 Mbledhja e të Dhënave	13
2.1.6 Kapja e të dhënave dhe kodimi	14
2.1.7 Editimi dhe Imputimi i të dhënave.....	15
2.1.8 Vlerësime	16
2.1.9 Analiza e të dhënave	16
2.1.10 Publikimi.....	17
2.1.11 Dokumentimi.....	17
2.2 Cikli i Jetës së një Vrojtimi.....	17
2.2.1 Planifikimi i Vrotimit.....	18
2.2.2 Modelimi dhe Zhvillimi	18
2.2.3 Zbatimi.....	18
2.2.4 Vlerësimi i Vrotimit	18
3. Dizajnimi i Vrotimit.....	19
3.1 Hyrje	19
3.2 Baza e zgjedhjes	21
3.3 Gabimet e kampionimit	23
4. Metodatat e mbledhjes së të dhënave	24
4.1. Intervistë personale.....	25
4.2 Vrojtimit nëpërmjet telefonit.....	25



REPUBLIKA E SHQIPËRISË
INSTITUTI I STATISTIKAVE

4.3 Vrojtimt nëpërmjet web-it.....	25
4.4 Skanimi i Pyetësorëve.....	26
4.6 ë Intervista personale me ndihmën e kompjuterit (CAPI).....	26
5. Dizenjimi i Pyetësorit.....	27
5.1 Specifikimi i nevojave	27
5.2 Rishikimi i pyetësorëve të mëparshëm.....	27
5.3 Hartimi i pyetësorit.....	28
5.4 Shqyrtimi i pyetësorit	28
5.5 Testimi.....	28
5.6 Finalizimi	28
6. Dizenjimi i kampionëve probabilitarë	29
6.1 Kampionet jo-probabilitare.....	29
6.1.1 Avantazhet e kampioneve jo-probabilitare	30
6.1.2 Disavantazhet e kampioneve jo-probabilitare.....	31
6.2 Kampionet probabilitare.....	31
6.3 Teknika të selektimit dhe të vlerësimit të të dhënave	32
6.3.1 Përdorimi i një informacioni ndihmës në kampionim dhe vlerësim.....	32
6.3.2 Parametrat, vlerësuesit dhe matësit e cilësisë.....	34
6.4 Teknikat bazë të kampionimit.....	37
6.4.1 Zgjedhja e rastit e thjeshtë.....	37
6.4.2 Metoda e kampionimit PPS – Probabilitet proporcional me përmasën.....	40
6.4.3 Kampionim me strata.....	42
6.4.4 Kampionimi me Clastera.....	43
7. Përcaktimi i përmasës së kampionit dhe alokimi i tij.....	44
7.1 Faktorët që ndikojnë në përcaktimin e përmasës së kampionit.....	44
7.2 Disa nga formulat më të përdorshme në caktimin e përmasës së kampionit	46
7.3 Alokimi i kampionit.....	49
8. Vlerësimi i të dhënave.....	51



REPUBLIKA E SHQIPËRISË
INSTITUTI I STATISTIKAVE

8.1 Peshimi.....	52
8.1.1 Trajtimet e mospërgjigjeve nëpërmjet peshimit të të dhënave.....	53
8.1.2 Faktorët e rregullimit të peshave për Mospërgjigjet.....	54
8.1.3 Përdorimi i Informacioneve ndihmëse për të rregulluar peshat	55
8.1.4 Vlerësimet e raporteve	56
8.1.5 Pesha të rregulluara më komplekse: kalibrimi dhe regresion i shumëfishtë	56
8.2 Prodhimi e vlerësimeve të thjeshta (Totalët, mesataret, përqindje).....	58
8.2.1 Vlerësues për tipe të dhënash të ndyshme.	58
8.3 Çështjet e vlerësimit në grupime të vogla (domains).....	61
8.3.1 Outliers (Njësi të veçuara).....	62
8.3.2 Llogaritja e gabimit të kampionimit të Vlerësimeve të Anketës	62
8.4 Llogaritja e variancës.....	63
9. Organizimi i vrojtimeve me zgjedhje.....	70
9.1 Procedurat e mbledhjes së të dhënave	70
9.3 Atributet që duhet të plotësojnë anketuesit.....	72
9.4 Monitorimi.....	73
10. Përpunimi i të dhënave.....	73
10.1 Klasifikimi dhe kodifikimi.....	74
10.2 Editimi dhe imputimi i të dhënave	74
10.3 Metoda imputimi të të dhënave.....	74
10.4 Derivimi i variablave të reja.....	75
10.5 Të dhënat e agreguara	75
10.6 Përgatitja e skedarëve finale	75
11. Shpërndarja e të dhënave.....	76
11.1 Raporti kryesor i anketës.....	76
11.2 Raporti i Cilësisë së të Dhënave	77
11.3 Raporte të tjera.....	77
11.4 Kontrolli i konfidencialitetit dhe zbulimit.....	78



REPUBLIKA E SHQIPËRISË
INSTITUTI I STATISTIKAVE

11.5 Përmbledhje	78
Referencat	79



REPUBLIKA E SHQIPËRISË
INSTITUTI I STATISTIKAVE

I. Hyrje

Për të siguruar cilësinë e statistikave zyrtare dhe besueshmërinë e publikut në to, prodhuesit e statistikave zyrtare udhëhiqen nga parimet e Kodit të Praktikës së Statistikave Europiane dhe një nga parimet e këtij kodi është Pavarësia profesionale. Pavarësia profesionale nënkupton që statistikave duhet të zhvillohen, të prodhohen dhe të shpërndahen në mënyrë të pavarur, veçanërisht për sa i takon zgjedhjes së teknikave, përkufizimeve, metodologjive dhe burimeve të të dhënave, kohës dhe formës së shpërndarjes, pa trysinë e grupeve politike, grupeve të interesit apo autoriteteve shtetërore.

Zgjedhja e teknika për vërtetimet statistikore është një komponente kritike, e nevojshme për të mundur dhe për të siguruar cilësinë, integritetin si edhe disponueshmërinë e të dhënave. Vërtetimet statistikore furnizojnë me informacione politikëbërës në shërbim të ndërmarrjes së politikave të rëndësishme në ekonominë dhe sociale të një vendi, furnizojnë me informacione kërkuesit shkencor në arritjen e konkluzioneve shkencore.

Udhëzimet metodologjike për vërtetimet me zgjedhje bazohen në udhëzimet e Eurostat¹ për Vërtetimet me Zgjedhje. Fokusi kryesor në material do të jenë aspekte të kampionimit si edhe aspektet e tjera të vërtetimit që lidhen me gabime jo-kampionimi por që ndikojnë në cilësinë e të dhënave të mbledhura.

II. Qëllimi

Qëllimi kryesor i këtij dokumenti është të përcaktojë hapat për realizimin e vërtetimeve me zgjedhje si dhe zbatimin e tyre. Dokumenti ka për qëllim të udhëheqë prodhuesit e statistikave zyrtare në lidhje me metodologjinë e vërtetimeve statistikore, përzgjedhjen e njësive të synuara etj.

Ky manual është kryesisht një udhëzues praktik i cili ndihmon në planifikimin e vërtetimit, modelimin dhe zbatimin e tij. Ai përfshin shumë çështje që lidhen me marrjen e kampioneve dhe shumë nga metodat themelore që mund të përfshihen në modelimin dhe zbatimin e një vërtetimi. Manuali nuk zëvendëson ekspertizën e njohur dhe gjykimin e shëndoshë, por përkundrazi ka për qëllim të ndihmojë në ndërtimin e tyre duke siguruar një pasqyrë të asaj që kërkohet për të ndërtuar vërtetime efikase dhe me cilësi të lartë.

¹ <https://ec.europa.eu/eurostat/documents/3859598/5901961/KS-RA-08-003-EN.PDF/833f7740-0589-47e1-99a5-c14878a2c1a8>



REPUBLIKA E SHQIPËRISË
INSTITUTI I STATISTIKAVE

Objektivat e përgjithshëm të këtij udhëzuesi janë përfshirja e aspekteve të përgjithshme të modelimit të një vrojtimi, duke përfshirë:

- një hyrje në konceptet e vrojtimit dhe planifikimin e një vrojtimi;
- si të formulohen objektivat e vrojtimit;
- konsiderata të përgjithshme në modelimin e vrojtimit;
- nëse duhet të bëhet një vrojtim me kampion apo një regjistrim i popullsisë së synuar dhe si të përcaktohet popullsia që do të vëzhgohet;
- lloje të ndryshme të kornizave të vrojtimit (sampling frame);
- burimet e gabimit në një vrojtim;
- metodat e mbledhjes së të dhënave të vrojtimit.

Gjithashtu ky udhëzim përmban më shumë aspekte teknike të modelimit të një vrojtimi me kampion:

- si zgjidhet një kampion;
- si vlerësohen karakteristikat e popullatës;
- si të përcaktohet madhësia e kampionit dhe si shpërndalet kampioni nëpër shtresa;

Në këtë dokument do të gjeni të ilustruar formulat e përdorura për të përcaktuar madhësinë e kampionit, që kërkojnë më shumë njohuri teknike. Udhëzuesi gjithashtu përfshin operacionet kryesore të përfshira në mbledhjen e të dhënave dhe mënyrën se si mund të organizohen operacionet për mbledhjen e të dhënave. Në të trajtohet mënyra se si përpunohen përgjigjet ndaj një pyetësi për të marrë një skedar të plotë të të dhënave të vrojtimit.

1. Rishikimi i udhëzimeve

Spektori i metodologjisë është përgjegjës për rishikimin e udhëzimeve metodologjike për vrojtimet me zgjedhje. Rishikimet janë të nevojshme kur:

- Ka ndryshime në rregulloret e Eurostat, të cilat sjellin implikime të metodave të kampionimit;
- Planifikohet dhe implementohet një teknologji e re ose ka ndryshime të konsiderueshme;



REPUBLIKA E SHQIPËRISË
INSTITUTI I STATISTIKAVE

- Raportet e auditimit statistikor ose rekomandime të tjera, të cilat identifikojnë nevoja për përmirësime metodologjike;
- Prezantohen standarde të reja kombëtare ose ndërkombëtare për metodologjine e vrojtimeve me zgjedhje, ose standardet ekzistuese rishikohen në mënyrë të konsiderueshme, për t'ju përgjigjur nevojave teknologjike dhe të institucioneve;

2. Vrojtimet statistikore

Çfarë është një vrojtim? Vrojtim është aktiviteti që lidhet me mbledhjen e informacionit nga një grup njësisht të një popullate (vrojtim me zgjedhje) apo gjithë popullata (regjistrim) rreth një ose disa karakteristikave të mirëpërcaktuara të saj. Informacioni i mbledhur në formën e një database, ndjek metoda dhe procedura të cilat lejojnë të përgjithësojnë rezultatin për gjithë popullatën. Pra një vrojtim statistikor nis me nevojën për një informacion. Kjo nevojë për informacion formalizohet dhe standartizohet përmes një definicioni të mirëpërcaktuar duke marrë kështu formën e një treguesi të matshëm.

Që një vrojtim të ketë rezultate në cilësinë e dëshiruar është shumë e rëndësishme që një vëmendje e veçantë t'i kushtohet punës përgatitore për të. Pavarësisht nga lloji i vrojtimit të gjitha kërkojnë një përgatitje dhe ndjekje të kujdesshme teknike të hapave të tij.

Fazat e kryerjes së një vrojtimi janë përshkruar në manualin e Modelit Generic Statistical Business Process Model (GSBPM). Ky manual përshkruan gjithë punët që nga përkufizimi i qëllimit kryesor të vrojtimit, strategjinë e grumbullimit të informacionit, përpunimin e të dhënave, nxjerrjen e rezultateve, vlerësimin e cilësisë deri tek arkivimi. Ne aneksin 1 jepet manuali i GSBPM.



REPUBLIKA E SHQIPËRISË
INSTITUTI I STATISTIKAVE

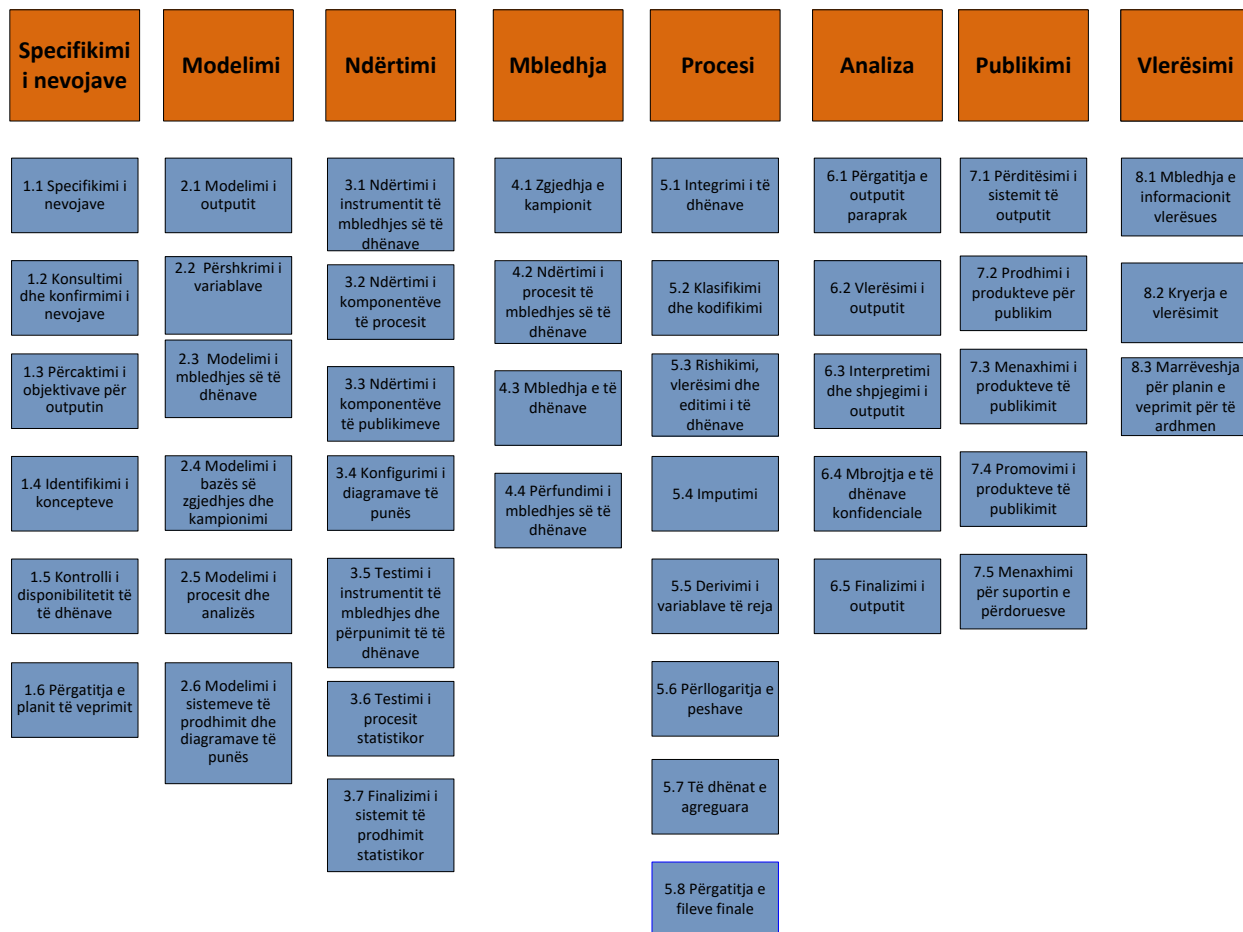


Figura 1: Fazat e kryerjes së një vrojtimi statistikor sipas GSBPM

Bërja e një vrojtimi mund të duket si një procedure e thjeshtë e bërjes së disa pyetjeve, përpunimit të tyre dhe prodhimit të statistikës. Por, që një vrojtım të jetë shkencor dhe rezultatet të jenë të sakta dhe të kuptimta duhet të ndjekë në mënyrë preçize procedura dhe formula. Me qëllim që të kuptojmë të gjithë procesin është e nevojshme të kuptojmë hap pas hapi punët në skemën e mësipërme dhe lidhjen e tyre.

2.1 Hapat e vrojtimeve me zgjedhje

Hapat e një vrojtimi mund ti përmbledhim në këtë listë:

- Formulimin e objektivit.
- Përcaktimin e bazës së kampionit, frame listës.



REPUBLIKA E SHQIPËRISË
INSTITUTI I STATISTIKAVE

- Përcaktimin e dizajnit të kampionit.
- Bërjen e pyetësorit, përcaktimi i dizajnit të pyetësorit.
- Grumbullimin e të dhënave.
- Kodifikimi, data entry, kontrolli
- Editimi (korigjimi) dhe imputimi i të dhënave.
- Vlerësimi, përgjithësimi i rezultatit përmes procesit të peshimit të tyre.
- Analiza
- Publikimi
- Dokumentimi.

2.1.1 Objektivi i vrojtimit apo qëllimi i vrojtimit

Që një vrojtimit të jetë i suksesshëm, është shumë e rëndësishme që objektivi kryesor i vrojtimit të jetë i përcaktuar mirë, i shprehur në mënyrë të qartë që në fillim të tij. Dikujt mund t'i duket një theksim i tepërt por j'u garantoj që nuk është i tillë.

Një objektivi i qartë i vrojtimit do të thotë një pohim statistikor i qartë në lidhje me informacionin që kërkon të japë ky vrojtimit. Një pohim i qartë në lidhje me fushën e mbulimit. Pra, popullata që ajo mbulon duhet të jetë qartë e kufizuar dhe të mos lë shkas për keq interpretime. Gjithashtu që në fillim duhet të jetë qartë e përcaktuar sesi këto rezultate do të përdoren nga përdoruesit, popullatën që ajo mbulon qoftë kjo në njësi vrojtimit apo shtrirje gjeografike. Pra pyetësori, variablat në të dhe plani i tabulimit duhet të jenë të mirëpërcaktuara që në fillim të studimit. Pra, që në fillim duhet të jenë të mirëpërcaktuara grupimet në të cilat kërkohet të nxirren rezultatet. Përgatitje e planit të tabulimit dhe përcaktimi sesi rezultatet do të prezantohen ndihmon në përcaktimin e detajuar të çdo variabli e kategorie të tij. Ky specifikim ndihmon studiuesin në fillimin e dizajnit të pyetësorit, kategorizimin e pyetjeve. Ky plan efektikon në të njëjtën kohë dhe punën e metodologjistit për përcaktimin e dizajnit të kampionit, përmasës dhe alokimit të kampionit.

Është shumë e rëndësishme që gjatë kësaj periudhe të zhvillohen konsultime mes përdoruesit dhe prodhuesit të statistikave në përcaktimin sa më të mirë të qëllimit të vrojtimit si dhe të fushës së mbulimit të tij. Këto konsultime ndihmojnë për të arritur në një kompromis midis të dhënave të kërkuara, formës në të cilën ato do të jepen, nivelit të disagregimit të tyre dhe vendosjen e strategjive për mënyrën e shpërndarjes së këtyre të dhënave në seri kohore.



REPUBLIKA E SHQIPËRISË
INSTITUTI I STATISTIKAVE

Ka shumë kufizimeve që ndikojnë në formulimin e qëllimit të vërtetimit. Një nga këto kufizime dhe më i rëndësishmi lidhet me cilësinë e vlerësimit. Sa i saktë, sa preciz do të jetë rezultati i pritshëm. Natyrisht që kjo lidhet me përcaktimin e një marzhi gabimi të pranueshëm për variablin më të rëndësishëm. Sa më të lartë t'a kërkojmë precizionin e rezultatit (në lidhje me një variabël ndihmës kryesor) aq më të lartë do t'a kemi përmasën e kampionit. Kështu që n'a duhet që në fillim një përcaktim i precizionit të kërkuar në total apo në nivelin e disagregimit dhe një kompromis për nivelin e disagregimit të të dhënave.

Faktorët që ndikojnë në precizion dhe rrjedhimisht në përmasën e kampionit përfshijnë:

- Homogjenitetin e karakteristikës në studim, variancën (Var)
- Përmasën e popullatës në studim (N)
- Modelimin e kampionit
- Metodën e vlerësimit
- Shkallën e përgjigjes

Kufizime të tjera operationale dhe buxhetore lidhen me precizionin si:

- Sa buxhet kemi disponibel?
- Sa intervistues, kompjutera, staf ndihmës duhen?
- Sa kohë kemi në dispozicion për t'a bërë këtë punë?

Megjithatë, në këtë çështje do të ndalemi disi në kapitujt e mëposhtëm, por gjithësesi është e rëndësishme t'i përmendim për të kuptuar lidhjen e ndërsjellë që ekziston midis hapave të një vërtetimi me zgjedhje.

Pra, pa një ide shumë të qartë të informacionit që kërkohet, studiuesit që ndërmarrin këtë aktivitet riskojnë humbjen e kohës dhe burimeve financiare pa arritur një informacion të vlefshëm. Për këtë arsye objekti i vërtetimit duhet të jetë i mirëpërcaktuar që në fazën e planifikimit të vërtetimit.

Në formulimin e qëllimit kryesor të vërtetimit, janë një sërë pyetjesh me të cilat metodologjisti njihet:

- Çfarë informacioni kërkohet nga ky vërtetim?
- Kush do t'i përdorë këto të dhëna dhe si do të përdoren?
- Çfarë definicionesh do të përdoren?
- Cila është fusha specifike që do të mbulojë ky vërtetim?



REPUBLIKA E SHQIPËRISË
INSTITUTI I STATISTIKAVE

- A është përgatitur plani i analizës dhe çfarë tabelash do të kërkohen?
- Cili është indikatori apo variabli kryesor në studim?
- Me çfarë precizioni do të kërkohen të dhënat?
- A ka kufizime të tjera operationale dhe buxhetore?

Të gjitha këto pyetje kanë efektin e tyre në hapat e mëtejshme të kësaj pune, prandaj këmbëngulet në mirëpërcaktimin e tyre.

Pra në konkluzion të kësaj çështje mund të themi se një formulim i mirë i qëllimit të vërtetimit është një themel i mirë i arritjes së një statistike cilësore.

2.1.2 Regjistrat dhe rëndësia e tyre në procesin e Vërtetimeve me Zgjedhje.

Regjistrat administrativë janë baza e sistemit statistikor të një vendi, si dhe burimi kryesor në shumë prej statistikave të një vendi. Një sistem statistikor nuk mund të kuptohet pa regjistra administrativë. Por, regjistrat administrativë përveçse mbulojnë një pjesë të informacionit statistikor shërbejnë si bazë për ndërtimin dhe mirëmbajtjen e regjistrave statistikorë të të dhënave. Ato dhe regjistrimet statistikore të popullatave (popullësi, ndërmarrje, ferma bujqësore etj) shërbejnë si bazë për ndërtimin e regjistrave statistikorë të të dhënave. Regjistrat statistikorë jo vetëm mbushin me informacion të bollshëm statistikor në fushën që mbulojnë por dhe shërbejnë si burim i ndërtimit të bazës së vërtetimeve me zgjedhje. Një regjistër statistikor i bollshëm në informacion dhe i azhurnuar në kohë shërben si një gur themeli për një vërtetim statistikor të mirë.

2.1.3 Përcaktimi i madhësisë së kampionit

Ekzistojnë dy lloje vërtetimesh: vërtetim me anë të kampioneve dhe regjistrim i përgjithshëm. Në rastin e vërtetimit me anë të kampionit, të dhënat grumbullohen vetëm për një pjesë (zakonisht një pjesë shumë e vogël) të njësive të popullatës ndërsa në rastin e regjistrimit të përgjithshëm, të dhënat grumbullohen për të gjitha njësitë e popullatës. Gjithashtu ekzistojnë dy lloje të marrjes së kampioneve: marrja e kampioneve jo-probabilitare dhe marrja e kampioneve probabilitare. Kampionimi jo-probabilitar siguron një mënyrë të shpejtë, të lehtë dhe të lirë për zgjedhjen e njësive nga popullata, por përdor një metodë subjektive të përzgjedhjes. Për të bërë konkluzione në lidhje me popullsinë nga një kampion jo-probabilitare, analisti i të dhënave duhet të supozojë se kampioni është përfaqësues i popullatës. Ky është shpesh një supozim i rrezikshëm duke pasur



REPUBLIKA E SHQIPËRISË
INSTITUTI I STATISTIKAVE

parasysh metodën subjektive të përzgjedhjes. Zgjedhja e kampioneve probabilitare është më komplekse, zgjat më shumë dhe zakonisht është më e kushtueshme sesa marrja e kampioneve jo-probabilitare. Sidoqoftë, për shkak se njësitë nga popullata janë zgjedhur në mënyrë të rastësishme dhe mund të llogaritet probabiliteti i zgjedhjes së secilës njësi, vlerësime të besueshme mund të prodhohen së bashku me vlerësimet e gabimit të marrjes së kampioneve dhe mund të bëhen konkluzione në lidhje me popullatën. Meqenëse marrja e kampioneve jo-probabilitare është zakonisht e papërshtatshme për një Institut Statistikash, ky udhëzues përqëndrohet në marrjen e kampioneve probabilitare.

2.1.4 Hartimi i Pyetësorit

Një pyetësor (ose formë) është një grup ose sekuencë pyetjesh të hartuara për të marrë informacion mbi një temë nga një i anketuar. Pyetësorët luajnë një rol kryesor në procesin e mbledhjes së të dhënave pasi ato kanë një ndikim të madh në cilësinë e të dhënave dhe ndikojnë në imazhin që Instituti i Statistikave i projekton publikut. Pyetësorët mund të jenë në formë letre ose të kompjuterizuar.

Problemet me të cilat përballen gjatë hartimit të pyetësorit përfshijnë: vendosjen e pyetjeve për të bërë, si t'i formuloni ato më mirë dhe si t'i rregulloni pyetjet për të dhënë informacionin e kërkuar. Qëllimi është të sigurohet informacion në mënyrë të tillë që të anketuarit e vrotimit të kuptojnë pyetjet dhe të mund të japin përgjigje të sakta me lehtësi, në një formë që është e përshtatshme për përpunimin dhe analizën e mëvonshme të të dhënave. Ndërsa ekzistojnë parime të përcaktuara mirë për hartimin e pyetësorëve, hartimi i një pyetësori të mirë mbetet një art që kërkon zgjuars, përvojë dhe testim. Nëse kërkesat e të dhënave nuk shndërrohen siç duhet në një instrument të strukturuar të mbledhjes së të dhënave me cilësi të lartë, një kampion "e mirë" mund të japë rezultate "të këqija".

2.1.5 Mbledhja e të Dhënave

Mbledhja e të dhënave është procesi i mbledhjes së informacionit të kërkuar për secilën njësi të zgjedhur në kampion. Metodatat themelore të mbledhjes së të dhënave janë: vetë-regjistrimi, ku i anketuari plotëson pyetësorin pa ndihmën e një intervistuesi, dhe të ndihmuar nga intervistuesi (qoftë përmes intervistave ballë për ballë ose atyre telefonike. Metodatat e tjera të mbledhjes së të dhënave përfshijnë: vëzhgimin e drejtpërdrejtë, raportimin elektronik të të dhënave nëpërmjet pyetësorit në faqen web dhe përdorimin e të dhënave administrative. Grumbullimi i të dhënave mund të jetë i bazuar në letër ose i ndihmuar nga kompjuteri. Me metoda të bazuara në letër,



REPUBLIKA E SHQIPËRISË
INSTITUTI I STATISTIKAVE

përgjigjet regjistrohen në pyetësorët e shtypur paraprakisht. Me metoda të ndihmuara nga kompjuteri, pyetësi shfaqet në ekranin e kompjuterit dhe përgjigjet hidhen drejtpërdrejt në kompjuter. Një nga përfitimet e metodës së asistuar nga kompjuteri është që: kapja e të dhënave - shndërrimi i përgjigjeve në një format të lexueshëm nga makineritë - ndodh gjatë mbledhjes, duke eliminuar kështu një aktivitet të përpunimit pas mbledhjes. Një përfitim tjetër është që: të dhënat e pavlefshme ose kontradiktore mund të identifikohen më lehtë sesa me një pyetësor letër.

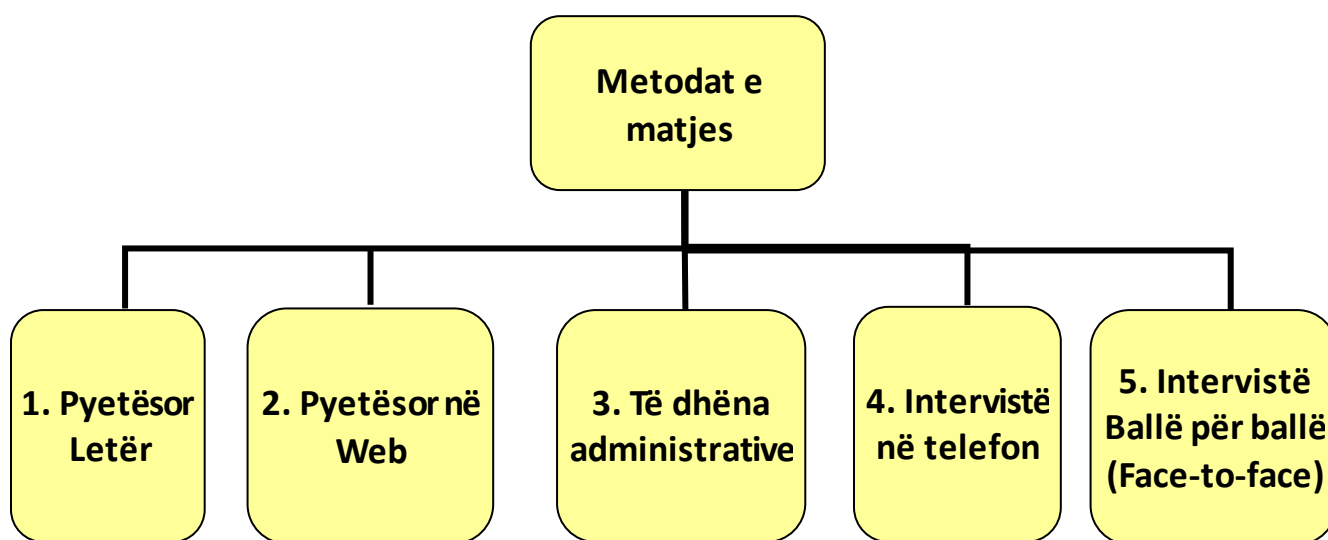


Figura 2: Metodat e matjes

2.1.6 Kapja e të dhënave dhe kodimi

Pasi janë mbledhur të dhënat, ato kodifikohen dhe, nëse nuk është përdorur metoda e mbledhjes e ndihmuar nga kompjuteri, bëhet kapja e tyre. Kodimi është procesi i caktimit të një vlere numerike ndaj përgjigjeve për të lehtësuar kapjen dhe përpunimin e të dhënave në përgjithësi. Disa pyetje kanë kategori të kodifikuara të përgjigjeve në pyetësor, të tjerët kodifikohen pas mbledhjes gjatë një procesi manual ose automatik. Kapja dhe kodimi i të dhënave janë aktivitete të shtrenjta dhe kërkojnë kohë që janë thelbësore për cilësinë e të dhënave pasi që çdo gabim i futur mund të ndikojë në rezultatet përfundimtare të vërtetimit. Prandaj, theksi duhet t'i kushtohet parandalimit të gabimit në fazat e hershme. Dy metoda të monitorimit dhe kontrollit të gabimeve janë sigurimi i cilësisë dhe kontrolli i cilësisë. Qëllimi i sigurimit të cilësisë është parashikimi i problemeve dhe parandalimi i tyre ndërsa qëllimi i kontrollit të cilësisë është të sigurohet që numri i gabimeve që ndodhin janë brenda kufijve të pranueshëm.



2.1.7 Editimi dhe Imputimi i të dhënave

Ky nën-proces aplikohet te të dhënat mikro të mbledhura dhe kontrollon çdo rekord në përpjekje për të identifikuar (dhe ku është e nevojshme të korrigjojë) problemet e mundshme, gabimet dhe mospërputhjet e tilla si outliers, njësitë pa përgjigje apo gabimet e kodimit.

Në këtë nënproces fillimisht merren të dhënat paraprake dhe analizohen për problemet e mësipërme. Mund të ekzekutohet në mënyrë iterative ky nënproces, përkundrejt rregullave të editimit të paracaktuara, zakonisht sipas një metode të caktuar. Rregullat e editimit mund të aplikohen në mënyrë automatike apo manuale për korrigjimin e të dhënave. Shqyrtimi, vlerësimi dhe redaktimi mund të aplikohet për të gjitha llojet e të dhënave si vrojtimit apo burimet administrative, para dhe pas integritetit. Në raste të caktuara, imputimi (nën-procesi 5.4) mund të përdoret si një formë e editimit.

Para integritetit të burimeve administrative është e nevojshme të pastrohen të gjitha burimet dhe të vendoset nëse ato përmbajnë gabime që mund të ndalojnë integritetin. Në hapin e parë, në bazë të kontrolleve të përcaktuara, gabimet zbulohen dhe printohen. Në hapin e dytë në bazë të përmbajtjes së udhëzimeve, zbatohen korrigjime manuale ose automatike. Nëse të dhënat janë me probleme (janë të pa përshtatshme), institucionit që ka dërguar informacionin i kërkohet ritransmetimi i të dhënave.

Për sa i përket nivelit të të dhënave, procedurat e redaktimit në përgjithësi mund të ndahen në të dhëna të redaktimit në nivel mikro dhe makro. Në përpunimin e të dhënave në nivelin mikro procedurat zbatohen në nivelin e njërive individuale, d.m.th. në nivelin e mikrodatave. Redaktimi i të dhënave në nivel mikro ndryshon në varësi të metodës së mbledhjes së të dhënave, terrenit apo vrojtimit nëpërmjet (telefonit, pyetësorit letër (hedhje e shpejtë e të dhënave nëpërmjet leximit optik).

Në vrojtimit që kryen në terren me CAPI, gjatë intervistës zbatohen kontrolle logjike. Pas intervistimit të dhënat integrohen në një skedar të përbashkët, ku nëse është e nevojshme, gabimet e zbuluara korrigjohen manualisht ose automatikisht.

Në vrojtimit me pyetësorë letër, pas hedhjes së shpejtë dhe kapjes së gabimeve nëpërmjet lexuesit optik, duhet të zbatohen kontrolle logjike. Gabimet e zbuluara korrigjohen manualisht ose automatikisht.

Redaktimi në nivel makro - Në një këndvështrim të ngushtë redaktimi në nivel makro nënkupton identifikimin dhe lokalizimin e gabimeve në të dhënat e agreguara tashmë. Në qoftë se me anë të



REPUBLIKA E SHQIPËRISË
INSTITUTI I STATISTIKAVE

redaktimit të të dhënave në nivel makro një gabim apo një vlerë është jashtë rangjeve të të dhënave, ajo duhet të kontrollohet dhe nëse është e nevojshme të korrigjohet në nivel mikro.

2.1.8 Vlerësime

Pasi të dhënat janë mbledhur, kapur, koduar, redaktuar (edituar) dhe imputuar, hapi tjetër është vlerësimi. Vlerësimi është mjeti me të cilin Instituti i Statistikave përfiton vlera për popullatën e interesit në mënyrë që të mund të nxjerrë përfundime në lidhje me atë popullatë bazuar në informacionin e mbledhur vetëm nga një kampion e popullsisë. Një vlerësim mund të jetë: një total, mesatare, raport, përqindje, etj.

Për një vrojtim me kampion, baza e vlerësimit është pesha e njësisë, e cila tregon numrin mesatar të njësive të popullsisë që përfaqëson. Një total i popullsisë mund të vlerësohet, për shembull, duke mbledhur vlerat e peshave të njësive në kampion. Pesha fillestare e modelimit (dizajnit) përcaktohet nga modeli i kampionit. Ndonjëherë, bëhen rregullime të kësaj peshe, për shembull, për të kompensuar njësitë që nuk i përgjigjen vrojtimit (d.m.th., total mos-përgjigje) ose duke marrë parasysh informacionin ndihmës. Rregullimet për mos-përgjigje mund të zbatohen gjithashtu për të dhënat e një regjistrimi të përgjithshëm (census).

Gabimi në kampionim ndodh në vrojtimit me mostra pasi vetëm një pjesë e popullatës është marrë në studim dhe njësitë e përzgjedhura nuk kanë saktësisht të njëjtat karakteristika si të gjitha njësitë e popullatës që ato përfaqësojnë. Një vlerësim i madhësisë së gabimit të kampionimit për çdo vlerësim duhet të jepet gjithmonë për t'ju treguar përdoruesve cilësinë e të dhënave.

2.1.9 Analiza e të dhënave

Në këtë fazë, statistikët prodhohen, shqyrtohen në detaje dhe përgatiten për shpërndarje. Kjo fazë përfshin nën-proceset dhe aktivitetet që mundësojnë analistët statistikorë për të kuptuar statistikën e prodhuara. Për rezultatet statistikore të prodhuara rregullisht, kjo fazë ndodh në çdo përsëritje. Faza e Analizës dhe nën-proceset janë të përgjithshme për të gjitha rezultatet statistikore, pavarësisht se si janë në lidhje me burimin e të dhënave.

Në këtë nën-proces statisticienët vërtetojnë cilësinë e rezultateve të prodhuara, në përputhje me kuadrin e cilësisë së përgjithshme dhe me pritshmëritë. Ky nën-proces përfshin aktivitetet e mëposhtme:

- Kontrollin se mbulimi i popullsisë dhe normat e përgjigjes janë siç kërkohet.



REPUBLIKA E SHQIPËRISË
INSTITUTI I STATISTIKAVE

- Krahasimin e statistikave me ciklet e mëparshme (nëse ka).
- Konfrontimin e statistikave me të dhënat e tjera (të brendshme dhe të jashtme).
- Hetimin e mospërputhjeve në statistika.
- Kryerjen e editimit në nivel makro.
- Vlerësimin e statistikave përkundrejt pritshmërive të çdo domaini.

2.1.10 Publikimi

Në këtë hap përgatiten produkte për të përmbushur nevojat e përdoruesve. Produktet mund të marrin forma të ndryshme, duke përfshirë botimet e shtypura, njoftimet për shtyp dhe faqet e internetit. Hapat tipike përfshijnë:

- Përgatitjen e përbërësve të produktit (tekst shpjegues, tabela, listat etj).
- Grumbullimin e komponentëve në produkte.
- Redaktimin e produkteve dhe kontrollin se ato i përmbushin standardet e publikimit.

2.1.11 Dokumentimi

Modeli GSBPM perdoret si një model dokumentimi, për të harmonizuar infrastrukturën e llogaritjes statistikore dhe për të siguruar një strukturë për procesin e vlerësimit dhe përmirësimit të cilësisë. Modeli GSBPM - Generic Statistical Business Process Model (GSBPM) përshkruan dhe përcakton një bashkësi nën procesesh që nevojiten për të prodhuar statistika zyrtare. Ai siguron një strukturë standard dhe një terminologji të harmonizuar për të ndihmuar prodhuesit statistikorë të modernizojnë proceset e prodhimit të statistikave si dhe të ndajnë metoda dhe komponentë. Ky model shërben për të përshkruar prodhimin e statistikave në një mënyrë më të përgjithshme dhe të orientuar drejt proceseve. Ai përdoret brenda dhe midis zyrave statistikore si një bazë e përbashkët pune për prodhimin e statistikave në mënyra të ndryshme, si: cilësia, efienca, standartizimi dhe orientimi drejt proceseve. Dokumentimi i prodhimit të proceseve është një element bazë për çdo institut statistikor. Ai është thelbësor për të njohur dhe dokumentuar metodat e prodhimit. Është e nevojshme për të patur një format të përbashkët për të përshkruar dhe analizuar secilën pjesë të procesit, duke identifikuar praktikatat më të mira, duke parë mangësitë në mjete etj dhe për të marrë masat për reduktimin e risqeve në të ardhmen.

2.2 Cikli i Jetës së një Vrojtimi

Hapat e vërtimit të paraqitura më lart nuk janë domosdoshmërisht sekuenciale: disa janë kryer paralelisht, të tjerët - për shembull, redaktimi (editimi) - përsëriten në periudha të ndryshme gjatë gjithë procesit të vërtimit. Çdo hap së pari duhet të planifikohet, pastaj të modelohet dhe



REPUBLIKA E SHQIPËRISË
INSTITUTI I STATISTIKAVE

zhvillohet, zbatohet dhe vlerësohet përfundimisht. Fazat në jetën e një vrojtimi përshkruhen më poshtë.

2.2.1 Planifikimi i Vrojtimit

Faza e parë e procesit të një vrojtimi është planifikimi. Por përpara se të zhvillohet ndonjë planifikim, një strukturë menaxhimi dhe planifikimi duhet të zgjidhet dhe zbatohet. Një strukturë e përdorur zakonisht është qasja e projektit ose ekipit të anketës, përmes të cilës ekipit ndërdisiplinor të vrojtimit i është dhënë përgjegjësia për planifikimin, hartimin, zbatimin dhe vlerësimin e vrojtimit dhe produkteve të tij të planifikuara. Ekipi ndërdisiplinor përbëhet nga anëtarë që kanë aftësi të ndryshme teknike; për shembull, një statistikien, një programues kompjuterash, një ekspert në fushën e studimit, një ekspert i mbledhjes së të dhënave, etj.

2.2.2 Modelimi dhe Zhvillimi

Pasi të krijoni një kornizë të gjerë metodologjike, është e mundur të kryeni punë të hollësishme në hapat e ndryshëm të një vrojtimit, në atë që përmendet, në fazën e projektimit dhe zhvillimit. Qëllimi i përgjithshëm i kësaj faze është gjetja e grupit të metodave dhe procedurave që arrijnë një ekuilibër të duhur të objektivave të cilësisë dhe kufizimeve të burimeve.

2.2.3 Zbatimi

Pasi të sigurohemi që të gjitha sistemet janë në funksion, vrojtimi mund të fillohet tani. Kjo është faza e zbatimit. Të gjitha format dhe manualët e kontrollit të vrojtimit janë shtypur, së bashku me pyetësonin (nëse përdoren pyetësonët e letrës). Intervistuesit janë trajnuar, mostra është zgjedhur dhe informacioni është mbledhur, të gjitha në mënyrën e vendosur gjatë fazës së zhvillimit. Pas këtyre aktiviteteve, fillon përpunimi i të dhënave. Aktivitetet e përpunimit përfshijnë kapjen e të dhënave, kodimin, redaktimin (editimin) dhe imputimin. Rezultati është një tërësi e strukturuar dhe e plotë e të dhënave nga e cila është e mundur të prodhohen tabelat e kërkuara dhe të analizohen rezultatet e vrojtimit. Këto rezultate kontrollohen më pas për konfidencialitet dhe shpërndahen. Në çdo hap, cilësia e të dhënave duhet të matet dhe monitorohet duke përdorur metoda të hartuara dhe të zhvilluara në fazën e mëparshme.

2.2.4 Vlerësimi i Vrojtimit

Vlerësimi i vrojtimit është një proces i vazhdueshëm gjatë gjithë vrojtimit. Çdo hap i vrojtimit duhet të vlerësohet në drejtim të efikasitetit, efektivitetit dhe kostos së tij, veçanërisht në rastin e vrojtimeve të përsëritura, në mënyrë që përmirësimet në hartimin dhe zbatimin e tyre të mund të bëhen me kalimin e



REPUBLIKA E SHQIPËRISË
INSTITUTI I STATISTIKAVE

kohës. Kjo përfshin vlerësime të metodave të përdorura, si dhe vlerësimeve të efektivitetit operacional dhe performancës së kostos. Këto vlerësime shërbejnë si një provë e përshtatshmërisë së praktikave teknike. Ato gjithashtu shërbejnë për të përmirësuar dhe udhëhequr zbatimin e koncepteve ose përbërësve specifikë të metodologjisë dhe operacioneve, brenda dhe përgjatë vrojtimeve.

3. Dizenjimi i Vrojtimit

3.1 Hyrje

Vendimi regjistrim apo vrojtimit me zgjedhje është kryesisht çështje ku duhen marrë në konsideratë edhe faktoret e mëposhtëm:

- burimet në dispozicion.
- përmasa e popullatës në studim.
- koha në dispozicion.
- specifikat e vrojtimit.

Duke qenë se vrojtimi me zgjedhje është një aktivitet shumë më i vogël se regjistrimi cilësia dhe kontrolli i tij janë shumë më lehtë të menaxhueshme dhe të monitorueshme. Pra, avantazhet e një vrojtimi me zgjedhje qëndrojnë në ofrimin në një kohë të shpejtë dhe me shumë më pak shpenzime të një informacioni statistikor me një cilësi të pëlqyeshme.

Elementet e mëposhtëm janë elemente të cilët duhet të merren parasysh kur dizenjohet një vrojtimit me zgjedhje:

- *Koston*- Dihet që regjistrimet janë aktivitete shumë të kushtueshme.
- *Kohën*- Koha që duhet në grumbullimin, përpunimin dhe shpërndarjen e rezultatit të një vrojtimit me zgjedhje është dukshëm shumë më e shkurtër se koha që kërkon një regjistrim total.
- *Përmasën e popullatës në studim*. Në qoftë se popullata në studim është shumë e vogël, ndoshta regjistrimi mund të jetë shumë më i arsyeshëm. Kjo për arsyen e thjeshtë që për arritjen e një vlerësimi me gabim të vogël kampionimi dhe precizion të lartë mund të kërkohej një kampionin i madh, relativisht në krahasim me popullatën (sampling fraction i lartë). Në të tilla raste ndoshta me një kosto shumë të vogël shtesë mund të kemi informacion nga e gjithë popullata në vend të një cope të saj. Në të kundërt, nëse popullata në studim është e madhe kampionimet janë shumë më të arsyeshme. Për më tepër dhe më gjerë në këtë temë do të ndalemi në kapitullin e përcaktimit të përmasës së kampionit dhe alokimit të tij.



REPUBLIKA E SHQIPËRISË
INSTITUTI I STATISTIKAVE

- *Vlerësimet në zona të vogla (small domeins).* Në vazhdim të diskutimit të mësipërm, regjistrimet rezultojnë shumë të mira në vlerësime në zona të vogla dhe japin rezultate për çdo nëngrup sado i vogël të jetë, ndërkohë që vrojtimet me zgjedhje kanë limitin e tyre të detajimit të rezultatit në varësi të gabimit të lejuar të vrojtimin në çdo nëngrup.

- *Prevalencën e parametrut (shkalla e përhapjes së dukurisë në studim).* Supozoni një vrotim që kërkon të vlerësojë një pjesë të popullatës që gëzon në karakteristikë të caktuar. Në qoftë se kjo karakteristikë është e shpeshtë atëherë mjaftojnë kampione të vogla për vlerësim të mirë (me gabim të pranueshëm). Pra, vrojtimet rezultojnë më efikënte se regjistrimet. E kundërta ndodh kur karakteristika në studim është shumë e rrallë.

- *Karakteristika specifike të dukurisë në studim.* Për të ilustruar këtë ide po përmendim vrotimin e keqshqyerjes së fëmijëve të zhvilluar nga INSTAT në vitin 2010. Vrotim përveç të tjerash kërkonte të maste dhe numrin e fëmijëve 0-5 vjeç që vuajnë nga anemia për shkak të keqshqyerjes. Për këtë qëllim ekipi i intervistuesve kishte në përbërje dhe një infermier që bënte të mundur marrjen e gjakut për analiza kimike të specializuara. Për këto lloj vrotimsh specifike as mund të mendohet se mund të bëhet census, për shkak të kostove shumë të larta dhe ekipeve shumë të specializuara të kërkuara. Madje po për arsye të kostos jo të gjithë fëmijët 0-5 vjeç bënë analiza gjaku por vetëm një nën-kampion i tyre.

- *Të tjerë faktorë.* Regjistrimet (censuses) janë aktivitete madhore që nuk kanë vetëm qëllimin e fotografikimit të njësive të vrotimit (banesa, persona, individë, ndërmarrje ekonomike etj.) studimin e cilëve të tyre socio-ekonomike në një moment të caktuar kohor, furnizimin me të dhëna shumë të vlefshme socio-ekonomike në këtë moment por dhe qëllimin e krijimit të regjistrave statistikor si bazë për vrojtimet e mëvonshme statistikore.

Informacioni ndihmës

Informacioni ndihmës për variablat kryesor mund të përdoret si në kohën e dizinjimit të kampionit (për krijimin e stratave, procedurë e selektimit etj) ose gjatë përpunimit të të dhënave për krijimin e peshës. Shumë herë ky informacion ndihmës nuk disponohet nga baza e vrotimit por është i disponueshëm pas vrotimit nga burime të ndryshme statistikore.

Në përfundim të kësaj çështje po shtojmë se regjistrat statistikorë që përdoren sot si burim për krijimin e bazës së një vrotimi përmbajnë informacione sipas koncepteve të standartizuara, përkufizimeve, procedura dhe klasifikime të standartizuara që përdoren në të gjitha vrotimet e ngjashme dhe që njihen nga përdoruesit e statistikave. Kjo punë është një proces i vazhdueshëm përmirësimi jo vetëm në vendin tonë por në gjithë sistemin e statistikave të bashkimit evropian. Arritje e një rritje të krahasueshmërisë dhe harmonizimit të gjithë sistemit të statistikave BE nis



REPUBLIKA E SHQIPËRISË
INSTITUTI I STATISTIKAVE

nga një harmonizim dhe rritje e krahasueshmërisë së regjistrave statistikorë që këto statistika kanë në bazën e tyre.

Pra në konkluzion të kësaj çështjeje mund të themi se cilësia e mirë e bazës së zgjedhjes së vrojtimit që ka si burim regjistrin është një themel i mirë për arritjen e një statistike cilësore. Nëse, synojmë përmirësimin e cilësisë së statistikave në vend është vendi për të vënë theksin në përmirësimin e gjithë sistemit të regjistrave në Shqipëri.

Të tjera çështje në konsideratë për kampionin

- Dublikatat duhen përjashtuar gjithmonë nga listat sapo ato të detektohen
- Për disa arsye, siç mund të jenë psh, pamundësia apo kostoja e lartë për të arritur disa nën-popullata të vogla, rritja e precizionit të kampionit në këto nën-popullata përjashtohen nga frame dhe përdoret kështu metoda me prerje (cut-off).

Por bashkë me këtë përfitim të madh përdorimi i kësaj metode ngre një pikëpyetje tjetër. (Hipotezë – të supozuar si të vërtetë) – A mundet që vlerësimet e bëra nga popullata e mbetur reflektojnë të njëjtat veti si dhe ajo e përjashtuar? – Nëse n'a duhet të vlerësojmë totalin e popullatës (të gjithën bashkë me prerjen) kemi nevojë të ndërtojmë modele për të kompensuar për copën e përjashtuar.

Metoda të tilla përdoren gjerësisht në vrojtimit e bizneseve ku më parë është studiuar që kontributi në parametrat e interesuar nga ndërmarrjet shumë të vogla është i vogël. Gjithësesi ekziston mundësia e ndonjë difference në vlerësim. Nëse n'a duhet të vlerësojmë totalin e popullatës kemi nevojë të ndërtojmë modele për të kompensuar për copën e përjashtuar.

3.2 Baza e zgjedhjes

Me Frame listë ose bazë të zgjedhjes (kampionit) kuptojmë listën e njësive të popullatës në studim. Megjithëse e përdorim rëndom kështu në këtë përkufizim nuk jemi shumë të saktë. Popullata në studim ("target population") jo gjithmonë përputhet me frame listën. Kjo për arsyen e thjeshtë që regjistri që përdoret në krijimin e bazës së zgjedhjes nuk e lejon të ndërtohet një i tillë. Sa më shumë të përputhet kjo frame listë me popullatën në studim aq më e saktë do të jetë baza e vrojtimit e përdorur dhe aq më i saktë do të jetë në vazhdim gjithë procesi i vlerësimit të treguesit. Që një bazë zgjedhje të quhet me cilësi të mira ose i plotë ajo duhet të plotësojë sa më mirë nevojën për të krijuar një frame listë të saktë në përputhje me popullatën në studim. Psh: Nëse duam të studiojmë gjithë ndërmarrjet e huaja që operojnë në Shqipëri në periudhën x të kohës, për disa tregues ekonomikë menjëherë hedhim vështrimin nga regjistri ekonomik i ndërmarrjeve. Fillimisht ndalemi në definicionin dhe shtrojmë pyetjen a është qëllimi i vrojtimit i shprehur saktë. Ç'do të thotë ndërmarrje e huaj? Duke patur parasysh që ndërmarrje të tilla janë ato me mbi 90% të kapitalit të tillë themi që qëllimi i vrojtimit është i shprehur saktë.



REPUBLIKA E SHQIPËRISË
INSTITUTI I STATISTIKAVE

(natyrisht faktori periudhë referemi duhet të jete i mirëpërcaktuar). A është e mundur të identifikohen këto ndërmarrje nga ky regjister? Përgjigja është: Po, sepse në regjister ekziston variabli pronësia ku janë të mirëkodifikuara ndërmarrjet e huaja. Atëhere në këtë rast baza e zgjedhjes përputhet me popullatën në studim.

Cilësitë e një Baze të Zgjedhjes bazohen në disa kriterë

Një nga këto kriterë e përmendëm më lart dhe kishte të bënte me faktin se sa ajo përfaqëson popullatën në studim. Pra, sa regjistri statistikor na lejon të krijojmë një bazë që të përputhet me popullatën në studim. Kjo natyrisht është e lidhur me plotshmërinë në informacion (variabla të updatuar në kohë) dhe fleksibilitetin që ky regjister lejon të lidhet me regjistra të tjerë.

Saktësia- Saktësinë e bazës mund t'a masim me disa karakteristika. Së pari janë problemet e mbulimit.

Mbimbulimet - përmban njësi të cilat nuk duhet të ishin, ose njësi jashtë fushës së qëllimit të vërtetimit. Këto njësi kur kontaktohen merret informacion për to dhe identifikohet sasia e tyre në kampion. Ky informacion është i vlefshëm për ngritjen e hipotezave në fazën e vlerësimit të indikatorëve.

Nënmbulimet. - nuk përmban njësi që duhet të ishin. Këto njësi është shumë e vështirë të identifikohen.

Nuk ka një rrugë reale që të sygjerojë përfshirjen e gjithë diferencave që ekzistojnë midis popullatës në studim dhe frame listës së vërtetimit, por kur këto diferenca zbulohen ato duhet të përfshihen në bazë. (KISH,1965) Në përgjithësi ato shfaqen në rastet e krahasimeve të serive kohore të të dhënave dhe kur zbulohen dhe vihet re që kanë impakt në seri përfshihen në të.

Dublikatat janë problemi më i rëndomtë që i ndodh një baze.

Keqkodifikimet – Edhe për problemet e klasifikimeve të bazave merret informacion gjatë procesit të mbledhjes së të dhënave të vërtetimit. Sa më e saktë të jetë kjo përputhje aq më i saktë do të jetë vlerësimi që duam të bëjmë.

Periudha kohore - çdo vërtetim i referohet një periudhe kohore të caktuar. Cilësia e një baze të mirë matet gjithashtu dhe me faktin sa ai është në kohë me këtë periudhe studimi, ose për shkak të procedurave që duhen për përditësimin e tij sa e afërt është ajo.

Baza vërtetimit ka impakt direkt në proceset e mëtejshme të vërtetimit si në përcaktimin e dizajnit të kampionit, të selektimit të tij, në zgjedhjen e metodës së vërtetimit, dhe rrjedhimisht vlerësimin e rezultateve. Rrjedhimisht një bazë cilësore është garanci për një rezultat cilësor.

Në rastet kur e gjithë popullata në studim nuk mbulohet nga një bazë zgjedhje e vetme përdoren bazat e shumëfishta. Rekomandohet që sa më shumë të jetë e mundur të evitohet përdorimi i tyre sepse kanë pothuajse gjithmonë probleme të mbivendosjes. Pra framet e shumëfishta duhet të përdoren vetëm në rastet e zgjedhjes së vetme të mbulimit të gjithë popullatës në studim. Kontrolli dhe monitorimi i cilësisë së mbulimit të bazës së vërtetimit është i domosdoshëm.



REPUBLIKA E SHQIPËRISË
INSTITUTI I STATISTIKAVE

Përmes procedurave të lidhjes me data-baza të tjera dhe verifikimit të informacionit gjatë fazës së grumbullimit të të dhënave të vrojtimit bëhet gjithmonë monitorimi i cilësisë së mbulimit të regjistrave. Gjithashtu, ky informacion i mbledhur nga vrojtimit e lidhura me këtë bazë zakonisht shfrytëzohet për përditësimin e regjistrave statistikor dhe janë pjesë e sistemit të përditësimit të tyre.

Ka raste kur nuk ekzistojnë baza për të mbuluar gjithë popullatën në studim dhe n'a duhet të gjejmë zgjidhje të tjera:

Bazat e shumëfishta

Baza të shumëfishta shfaqen atëherë kur për të arritur gjithë popullatën në studim të duhet të merresh me shumë burime informacioni me struktura të ndryshme. Është shumë e zakonshme kur studiohen fenomene të reja ose shumë e zakonshme në vendet në zhvillim ku regjistrat administrative janë në zhvillim.

Bazat e vrojtimit me Shtresa (cluster)

Shumë shpesh ndodh dhe që nuk ka një frame të mirë për njësitë në vrojtimit dhe krijimi i një të tili të jetë shumë i kushtueshëm. Kështu që një zgjidhje është të rrëmosh për kombinimin e metodave të vrojtimit dhe të krijohen baza me shtresa për vrojtimit.

Konsideroni psh: një studim të nxënësve të shkollës. Edhe në qoftë se baza të popullatës ekzistojnë për të gjithë fëmijët që ndjekin shkollën, puna në terren për vrojtimit e këtyre fëmijëve do të qe më e thjeshtë dhe më pak e kushtueshme nëse shkollat dhe klasat ishin zgjedhur në vend të një zgjedhje të rastësishme të fëmijëve në të gjithë popullatën e tyre.

Në popullata të mëdha dhe në vrojtimit pranë familjeve është shumë e zakonshme përdorimi i bazave me shtresa si një kombinim i disa elementeve të një zone gjeografike. Arsye kryesore e përdorimit të tyre është se janë më pak të kushtueshme se bazat e njëjësive të vrojtimit.

3.3 Gabimet e kampionimit

Në çdo vrojtimit janë prezentë dy tipe gabimesh:

- Gabime të kampionimit (sampling errors).
- Gabime jo të kampionimit (nonsampling errors).

Gabimet e kampionimit janë prezente vetëm në vrojtimit me zgjedhje dhe jo në regjistrimet. Gabimet jo të kampionimit janë prezente në të dyja. Sa më shumë rritet përmasa e vrojtimit aq më shumë zvogëlohen gabimet e kampionit dhe deri në eliminim e tyre në regjistrimet dhe aq më shumë rriten gabimet jo të kampionimit. Pra me kalimin nga vrojtimit me zgjedhje (kampionim) në regjistrim kemi një lloj kalimi (shkëmbimi) ndërmjet gabimeve të kampionimit dhe gabimeve jo të kampionimit.

Kur vendosim të bëjmë vrojtimit me kampionim apo regjistrim marrim në konsideratë listën e mëposhtme të faktorëve:



REPUBLIKA E SHQIPËRISË
INSTITUTI I STATISTIKAVE

Gabime të kampionimit. Gabimet e kampionimit shoqërojnë vërtetimet me zgjedhje (me kampion). Ato zakonisht maten duke vlerësuar në ç'interval do të shtrihet vlerësimi i parametrit të kërkuar bazuar në gjithë kampionet me të njëjtën përmasë dhe duke përdorur të njëjtën metodë kampionimi (sample design), sa këto vlerësime ndryshojnë nga njëri-tjetri. Pra, definicioni është shumë i gjerë dhe praktikisht i pamundur të matet në këtë mënyrë sepse në përgjithësi ne bëjmë një kampion për të vlerësuar një parametër. Kështu që si vlerësues i gabimit të kampionit është intervali i besimit të parametrit të estimuar. Në kapitujt e mëvonshëm do të japim shembuj të llogaritjes së tij në metoda të ndryshme kampionimi.

Gabime jo të kampionimit. Regjistrimet vërtetë nuk kanë gabime të kampionit sepse vërtetojnë gjithë popullatën. Duket se rezultatet në regjistrim janë shumë më të sakta se rezultatet nga një vërtetim. Por, gjithë vërtetimet janë objekt i gabimeve jo të kampionimit (të grumbullimit të informacionit, plotshmërisë të matjes së tyre, kodifikimit dhe përpunimit të të dhënave), të cilat mund të çojnë në shtrembërim të rezultatit njësoj si gabimet e kampionit.

4. Metodatat e mbledhjes së të dhënave

Mbledhja e të dhënave është procesi i mbledhjes së informacionit të kërkuar për secilën njësi të zgjedhur në kampion. Gjatë mbledhjes së të dhënave, anëtarët e popullatës - qofshin ata individë ose organizata, ndërmarrje – janë të vendosur dhe kontaktuar dhe kërkohet pjesëmarrja e tyre në anketim. Më pas administrohet një pyetësor dhe regjistrohen përgjigjet. Ky proces është i kushtueshëm, kërkon kohë, kërkon burime të gjera dhe ka një ndikim të drejtpërdrejtë në cilësinë e të dhënave. Meqenëse është kontakti kryesor i publikut të gjerë me institutin e statistikave, ai kontribuon në imazhin e institucionit dhe ka një ndikim të gjerë në rëndësinë e këtij institucioni dhe në cilësinë e të dhënave të saj.

Në shumicën e rasteve të dhënat statistikore janë rezultat i vërtetimeve statistikore. Një vërtetim statistikor përfshin të gjitha aktivitetet e mbledhjes, përpunimit dhe shpërndarjes së të dhënave statistikore. Përmes viteve praktika e organizatave statistikore ka krijuar lloje të ndryshme të vërtetimeve statistikore. Metoda e mbledhjes së të dhënave input krijon dallime midis llojeve të ndryshme të burimeve statistikore:

- **Censuset:** Të dhënat mbledhen për të gjitha njësitë e popullsisë së synuar.
- **Vërtetimet:** Të dhënat mbledhen për një kampion rastësisht të zgjedhur të njërive të popullatës së synuar.
- **Burimet administrative:** Të dhënat që fillimisht janë mbledhur për ndonjë qëllim tjetër dhe përdoren nga Sistemi Kombëtar Statistikor (SKS) për nxjerrjen e treguesve statistikore.



REPUBLIKA E SHQIPËRISË
INSTITUTI I STATISTIKAVE

Pasi kemi përcaktuar madhësinë e kampionit, duhet të zgjedhim një nga metodat e mbledhjes së të dhënave. Çdo metodë ka avantazhet dhe disavantazhet e veta. Më poshtë do të jepen disa nga metodat e mbledhjes së të dhënave.

4.1. Intervistë personale

Një intervistë quhet personale kur anketuesi kërkon të perballët me pyetje ballë për ballë me të intervistuarin.

Avantazhet

- Aftësia për të gjetur popullatën e synuar.
- Intervistat e shkurtra tolerohen më lehtë. Veçanërisht në intervistat në shtëpi njerëzit janë me të gatshëm të flasim ballë për ballë me një intervistues se sa në telefon.

Disavantazhet

- Intervistat personale mund të kushtojnë më shumë se metodat e tjera të mbledhjes së të dhënave.

4.2 Vrojtimt nëpërmjet telefonit

Vrojtimt nëpërmjet telefonit janë një nga metodat e intervistimit.

Avantazhet

- Njerëzit mund të kontaktohen më shpejt me anë të telefonit sesa me metodat e tjera. Në qoftë se intervistuesit janë duke përdorur CATI, rezultatet mund të jenë në dispozicion disa minuta pas përfundimit të intervistës.
- Të intervistuarit mund të japin përgjigje me të plotë se mënyrat e tjera.
- Anketuesi mund të kërkojë shpjegime në lidhje me një përgjigje të të intervistuarit.

Disavantazhet

- Është e vështirë marrja e përgjigjeve për pyetësorët e gjatë.

4.3 Vrojtimt nëpërmjet web-it

Avantazhet

- Vrojtimt nëpërmjet web-it janë me pak të kushtushme.
- Këto lloj vrojtimsh mund të realizohen nëse ju keni emrat dhe adresat e email të të gjithë popullatës, por jo numrat e telefonit.



REPUBLIKA E SHQIPËRISË
INSTITUTI I STATISTIKAVE

- I intervistuari nëpërmjet email mund të përgjigjet në kohën e lirë ndryshe nga format e tjera kur anketuesi mund ta gjejë në një kohë të papërshtatshme.

Disavantazhet

- Koha! Ju duhet të prisni shumë javë për të pritur përgjigjen e pyetësorëve.
- Në popullatat të cilat kanë një nivel të ulët arsimor, shkalla e përgjigjeve nëpërmjet email zakonisht mund të jetë e ulët. Edhe në popullatën me edukim të mirë shkalla e përgjigjes shkon nga 30-90%. Nuk ka një rregull të përgjigjeve në bazë të niveleve arsimore apo interesave në një fushë.

4.4 Skanimi i Pyetësorëve

Skanimi i pyetësorëve është një metodë e mbledhjes së të dhënave që mund të përdoret në pyetësorët letër ku këta pyetësorë janë realizuar nëpërmjet: intervistimit ballë për ballë, email apo vrojtim nëpërmjet telefonit.

Avantazhet

- Skanimi mund të ulë koston e hedhjes së të dhënave.
- Procesi i skanimit është më i saktë se leximi i një pyetësori nga një person.

Disavantazhet

- Procesi i skanimit kërkon si hardware dhe software për leximin e pyetësorëve, dhe kjo gjë do të thotë kosto shtesë.

4.6 ë Intervista personale me ndihmën e kompjuterit (CAPI)

Aktualisht gjithnjë e më shumë zbatohen metodat e reja për grumbullimin e të dhënave nëpërmjet Intervistimeve Personale me Ndhmën e Kompjuterit (CAPI) për të zvogëluar kohën për kapjen e të dhënave dhe për të përmirësuar cilësinë e të dhënave.

Avantazhet

Kjo metodë e intervistimit është e lirë kur ne na nevojiten një numër i madh i të intervistuarëve sepse:

- Nuk është e nevojshme rekrutimi dhe pagimi i operatorëve .
- Pastrimi i të dhënave kërkon më pak kohë, pasi pjesa më e madhe e kushteve llogjike/aritmetike janë vendosur në program.
- Hapësira e marrjes së të dhënave dhe ruajtjes ;
- Koha e parashikuar për rikthimin e grumbullimit të të dhënave;
- Cilësia e të dhënave të grumbulluara.



Disavantazhet

- Kostoja

5. Dizenjimi i Pyetësorit

Procesi i hartimit të pyetësorit fillon me formulimin e objektivave dhe informacionit të vërtetimit, kërkesat dhe vazhdon me hapat e mëposhtëm:

- Të konsultohet me përdoruesit e të dhënave dhe të anketuarit;
- Rishikoni pyetësorët e mëparshëm;
- Hartoni pyetësorin;
- Shqyrtoni dhe rishikoni pyetësorin;
- Testoni dhe rishikoni pyetësorin;
- Finalizoni pyetësorin.

5.1 Specifikimi i nevojave

Specifikimi i nevojave për të dhëna fillon kur të dhënat nuk ekzistojnë, ose kur të dhënat ekzistuese nuk i përmbushin plotësisht nevojat e të gjithë përdoruesve për të dhëna, ose në bazë të kërkesave të reja apo të marrëveshjeve të bashkëpunimit. Nevojat për të dhëna lindin nga përdorues të ndryshëm: institucione qeveritare, Banka e Shqipërisë, institucione kombëtare dhe ndërkombëtare, si dhe publiku profesional apo i përgjithshëm. Në kuadër të përgatitjes, duhet të studiohen metoda të ndryshme, në mënyrë që aktiviteti statistikor të plotësojë të gjitha nevojat e kërkuara. Në këtë nën-proces mund të planifikohen burimet financiare dhe materiale për zbatimin e detyrave. Planifikimi i saktë dhe në kohë është me rëndësi kyçe për zbatimin me efektivitet dhe eficiencë.

5.2 Rishikimi i pyetësorëve të mëparshëm

Vërtetimet e tjera janë një burim i mirë informacioni kur zhvillonin një studim. Ekzaminimi i pyetjeve të përdorura nga vërtetimet të tjera mbi të njëjtën temë ose të ngjashme mund të jetë një pikënisje e dobishme kur merret parasysh si të hartoni një pyetje (d.m.th., formulimin e pyetjeve). Në disa situata, të njëjtat pyetje duhet të përdoren, për shembull kur përpiqeni të krahasoni rezultatet nga vërtetimet të ndryshme. Gjithashtu, dokumentacioni mbi cilësinë e të dhënave nga anketimet të tilla duhet të ekzaminohet për të vlerësuar efektivitetin e pyetësorit (p.sh. problemet me formulimin e pyetjeve, barrën e përgjigjes, nivelet e refuzimit, etj)



5.3 Hartimi i pyetësorit

Hapi tjetër është hartimi i një drafti të të gjithë pyetësorit. Për shkak se pyetësi ndikohet nga modeli i plotë i vërtetimit, së bashku me objektivat e anketës, duhet të merren parasysh edhe faktorët e mëposhtëm:

- Metoda e mbledhjes së të dhënave
- Karakteristikat e të anketuarve
- Ngarkesën e të intervistuarit
- Kompleksiteti i të dhënave që duhet të mbledhen
- Konfidencialiteti dhe ndjeshmëria e informacionit
- Përkthimi
- Krahueshmëria e rezultateve me vërtetimet e tjera
- Konsistenca

5.4 Shqyrtimi i pyetësorit

Është thelbësore që pyetësi të rishikohet së brendshmi përpara se të testohet. Ky rishikim duhet të identifikojë ndonjë problem të dukshëm me pyetësin siç janë gabimet në drejtshkrim ose gramatikore ose formulime të çuditshme. Është gjithashtu e dobishme në këtë fazë që pyetësi të rishikohet nga njerëz që nuk janë të përfshirë direkt me projektin. Rishikuesit mund të përfshijnë ekspertë në fushën e studimit, njerëz që kanë përvojë në hartimin e pyetësorëve, intervistuesit ose anëtarët e popullatës që do të anketohen. Ata shpesh mund të japin komente dhe sugjerime të dobishme që çojnë në rishikime në pyetjet dhe kategoritë e përgjigjeve.

5.5 Testimi

Është e rëndësishme të testoni të gjitha versionet (d.m.th., të gjitha versionet gjuhësore) të pyetësorit mbi të anketuarit 'tipikë' shumë kohë përpara se të fillojë mbledhja e të dhënave (d.m.th., tipike për popullatën e synuar, që mund të nënkuptojë të anketuarit e një moshe, seksi dhe niveli arsimor të caktuar). Mund të jetë gjithashtu e rëndësishme të testoni pyetësin në nëngrupe specifike të popullatës së synuar që mund të kenë probleme me pyetje të veçanta.

5.6 Finalizimi

Hartimi i pyetësorit është një proces përsëritës: përgjatë zhvillimit dhe testimit të pyetësorëve, ndryshimet bëhen vazhdimisht. Objektivat dhe kërkesat e informacionit janë formuluar dhe rivlerësuar, konsultohen përdoruesit e të dhënave dhe të anketuarit, pyetjet e propozuara hartohen



REPUBLIKA E SHQIPËRISË
INSTITUTI I STATISTIKAVE

dhe testohen, pyetjet shqyrtohen dhe rishikohen derisa të hartohet një pyetësor përfundimtar. Faza përfundimtare e procesit vjen kur të merret vendimi se nuk do të bëhen ndryshime të mëtejshme në pyetësor. Në këtë pikë, pyetësori është finalizuar dhe shtypur ose programuar, në varësi të metodës së kapjes së të dhënave që përdoret.

6. Dizenjimi i kampionëve probabilitarë

Gjatë kapitullit të mëparshëm pamë se në fazën e pranifikimit të vërtetimit vendosëm të bëjmë regjistrim apo vërtetim me zgjedhje dhe nëse vendimi është vërtetim me zgjedhje atëherë metodologjisti duhet të vendosë si duhet zgjedhur ky kampion. Mënyra se si selektojmë një nënbashkësi të të gjithë popullatës me qëllim grumbullimin e informacionit dhe përgjithësimin e tij për gjithë popullatën përbën atë që quhet dizajn i kampionit.

Kampionet mund t'i ndajmë në dy grupe të mëdha: Kandidate jo-probabilitare dhe kandidate probabilitare. Kampioni është probabilitar kur për çdo njësi të bazës së vërtetimit njihet probabiliteti nëse është në kampion apo jo.

Pra: $\forall u_i \in U$ njihet probabiliteti p_i që $u_i \in S$

Pra për çdo njësi njihet niveli i përfaqësimit që ai ka për të tërën. Pra këto kandidate janë të bazuara mbi parimin e zgjedhjes së rastit. Në këtë mënyrë kandidate probabilitare lejojnë përgjithësimin e rezultatit në nivel të popullatës në studim.

Nga ana tjetër kandidate jo probabilitare janë një metodë subjektive e selektimit të njësive të vërtetimit. Me qëllim përgjithësimin e rezultatit statisticienët pranojnë hipotezën që ky kampion është përfaqësues. Kjo hipotezë është shumë me risk në kushtet e një kampionimi jo-probabilitar.

6.1 Kandidate jo-probabilitare

Kandidate jo-probabilitare janë një metodë e zgjedhjes në mënyrë subjektive (jo të rastësishme) e njësive nga popullata. Këto kandidate nuk kërkojnë një bazë vërtetimi të të gjithë popullatës. Ato janë një mënyrë e shpejtë e marrjes së të dhënave dhe me shumë pak shpenzime. Problemi kryesor i gjithë këtyre metodave është se këto kandidate nuk janë aspak të qarta nëse mund t'a përgjithësojnë rezultatit për gjithë popullatën. Këto kandidate zakonisht ia lënë në dorë intervistuesit të bëjë kampionimin. Natyrisht është shumë normale të mendojmë se ata do të intervistojnë anëtarë që mund t'i kontaktojnë lehtë dhe janë të ngjashëm nga pikëpamja sociale me shoqërinë e tyre. Kjo mënyrë sjellje e tij mund të sjell një reduktim në variabilitetin e përgjigjeve të të intervistuarve, sjell një shmangie sistematike në mënyrën e rastësishme të



REPUBLIKA E SHQIPËRISË
INSTITUTI I STATISTIKAVE

zgjedhjes dhe rrjedhimisht një rezultat të shtrembëruar, jo përfaqësues. Më poshtë po përmendim disa tipe kampionesh jo-probabilitare si:

Kampionimet me risk të madh, pa plan.

Kampionet me risk të madh janë kampione pa plan të caktuar, pa rregull selektimi të kampionit. Një shembull mund të jetë ai i intervistimit të personave në rrugë, të plotësimit me e-mail të pyetësoreve dërguar tek miqtë e tu. Natyrisht janë kampione jo të përshtatshme për përgjithësime të rezultatit.

Kampionet vullnetare.

Edhe këto kampione nuk kanë rregull selektimi. Kujtoni për shembull vrojtimt televizive dhe statistikat që jepen për një mini-pyetësor drejtuar dëgjuesve të radios apo televizionit. Pra, pjesëmarrësit në këto lloj vrojtimsh marrin pjesë në mënyrë të vullnetshme.

Kampionimet me gjykim eksperti.

Këta lloj kampionesh janë të bazuar në ide të një eksperti që ka njohuri mbi përbërjen e popullatës. Gjykimi i tij përpiket t'a bëjë kampionin reprezentativ për popullatën në studim. Këto lloj kampionesh ndonjëherë janë më të rrezikshme se kampionet me risk të madh kur gjykimi i ekspertit rezulton të jetë i gabuar. Megjithatë janë kampione të përdorshme sidomos në testimin e pyetësorëve adresuar grupeve të veçanta të intervistuesve apo për të testuar pyetësor me specifika të veçanta.

Kampionimet me quota.

Kjo është një nga metodat më të përhapura të përdorura sot nga agjenci të ndryshme statistikore. Kjo metodë synon arritjen e një përmase të caktuar kampioni në nëngrupe të popullatës në studim, ndoshta bazuar në një shpërndarje proporcionale të kampionit në nëngrupe të popullatës në studim. Ka një ngjashmëri me kampionin me strata por dallimi i madh qëndron në mënyrën sesi selektohen njësitë brenda grupit. Brenda grupeve të caktuara ky selektim është i njëjtë si në rastin e një vrojtimi me risk të lartë, pa plan.

6.1.1 Avantazhet e kampioneve jo-probabilitare

- Janë shumë të shpejta dhe të përshtatshme
- Janë me shumë pak shpenzime
- Nuk janë të shpërndara në gjithë territorin, nuk kanë zakonisht kosto udhëtimi.
- Nuk kërkojnë një bazë zgjedhje.



REPUBLIKA E SHQIPËRISË
INSTITUTI I STATISTIKAVE

- Janë të vlefshme si studime paraprake ose vërtetim pilot për të analizuar zhvillimin e një vërtetimi shkencor me zgjedhje, testuar pyetësin, programin e hedhjes së të dhënave në kompjuter etj.

6.1.2 Disavantazhet e kampioneve jo-probabilitare.

- E kanë të pamundur të përcaktojnë se me ç'probabilitet një njësi në popullatën në studim do të jetë në kampion kështu që rrjedhimisht e kanë të pamundur të vlerësojnë një parameter për gjithë popullatën dhe të masin gabimin e këtij vlerësimi.

Megjithatë kampionet jo probabilitare përdoren sot nga kërkuesit si një metodë jo e kushtueshme, e shpejtë e marrjes së të dhënave por kurrësesi ato nuk mund të zëvendësojnë kampionet probabilitare për arsyet e përmendura. Këto kampione përdoren zakonisht në gjenerimin e ideve, në kampionet pilot që bëhen në fazën e testimit të gjithë hapave të një vërtetimi të madh probabilitar, testimin e pyetësorit, testimin e anketuesve, të vështirësive të terrenit, të programit të regjistrimit të të dhënave etj. Këto kampione nuk janë të rekomandueshme ose më saktë nuk janë shkencore dhe nuk duhet të jenë të përdorshme nga studiuesit që kërkojnë të nxjerrin rezultate për gjithë popullatën në studim.

Një rast i veçantë i përdorimit të tyre me shumë sukses në sistemin statistikor është kampioni në llogaritjen e indekseve të çmimeve. Këtu shporta e produkteve mbulon top listën prej 80% të listës së produkteve dhe çmimet për to grumbullohen në ndërmarrje të mëdha apo shpërndarje gjeografike të ndryshme.

6.2 Kampionet probabilitare

Siç edhe e kemi përmendur me lart, kampioni është probabilitar kur për çdo njësi të bazës së vërtetimit njihet probabiliteti nëse është në kampion apo jo.

Pra: $\forall u_i \in U$ njihet probabiliteti p_i që $u_i \in S$

Pra, për çdo njësi njihet niveli i përfaqësimit që ai ka për të tërën. Këto kampione janë të bazuara mbi parimin e zgjedhjes së rastit. Në këtë mënyrë kampionet probabilitare lejojnë përgjithësimin e rezultatit për gjithë popullatën në studim.

Shihet që ka dy apekte të cilat e bëjnë kampionin probabilitar.

- njësitë zgjidhen në mënyrë të rastësishme.
- të gjitha njësitë e frame listës kanë probabilitet më të madh se zero për të qenë të zgjedhur. Nuk është e nevojshme të kenë probabilitet të njëjtë. Ky probabilitet varion nga njësia në njësi në varësi të metodës së kampionimit.



REPUBLIKA E SHQIPËRISË
INSTITUTI I STATISTIKAVE

Ka shumë tipe të dizajnit të kampioneve probabilitetare por më të përdorshmet e hasura në praktike janë SRS-kampioni i rastit thjeshtë, SRS-sys-kampioni i rastit i thjeshtë sistematik, PPS-kampion proporcional me përmasën, cluster-sampling – kampion me clastera.

Në fund po përmendim dhe kampionimet e kombinuara – Këto janë kampione që fillojnë me metodë probabilitare kampionimi të zonës dhe vazhdojnë me jo-probabilitare të njësive brenda saj.

6.3 Teknika të selektimit dhe të vlerësimit të të dhënave

Në një vrojtim me zgjedhje, një kampion probabilitar merret nga një frame i popullatës në studim duke përdorur një dizenjim të mirëpërcaktuar. Zakonisht, një dizenjim konsiston në një kombinim i teknikave të ndryshme të selektimit të kampionit. Një dizenjim kompleks mund të përfshijë dhe hapa të shumtë kampionimi. Në një rast të thjeshtë, selektimi i elementëve bëhet direkt nga baza e vrojtimit.

Shembuj të mirë ne lidhje me këtë çështje janë Kish (1965), Cochran (1977); Lohr (1999), and Lehtonen & Pahkinen të cilat janë burimet primare të shkrimit të këtij seksioni.

Materiale shtesë të vlefshme mund të gjeni në linkun: <http://www.math.helsinki.fi/VLISS/> të cilat janë një shtesë e librave të Lehtonen dhe Pahkinen. Shumë nga teknikat e zakonshme të selektimit të kampionit janë tashmë të implementuara në softe të ndryshme statistikore; psh SAS brenda procedurës SURVEYSELECT.

Vetitë e teknikave të kampionimit variojnë në varësi të efieciencës dhe një sërë aspektesh të tjera praktike, si qëndrueshmëria e një pune selektimi, kërkesa për një aplikacion të mirëpërdorshëm. Por, gjithmonë studimet e dizenjimit të kampionit marrin parasysh dhe kufizimet në buxhet. Një aspekt shumë i rëndësishëm i përfitimit të një dizenjimi kampioni eficient është dhe roli i një informacioni ndihmës. Le të diskutojmë atëherë skemën standarte të selektimit SRS nga kjo pikëpamje:

6.3.1 Përdorimi i një informacioni ndihmës në kampionim dhe vlerësim.

Shumë e zakonshme dhe e vlefshme që informacionin ndihmës në lidhje me popullatën t'a përdorim në modelimin (dizenjimin) e kampionit.

Praktikisht, ka shumë mënyra nga mund të merret informacioni ndihmës. Psh: zakonisht frame-që përdoren kanë informacione ndihmëse, ose mund të merren nga burime administrative në nivel mikro. Por janë shumë të përdorshme dhe informacionet ndihmëse në nivel të agreguar. Ky



REPUBLIKA E SHQIPËRISË
INSTITUTI I STATISTIKAVE

informacion ndihmës mund të përdoret si në krijimin e një kampioni më eficient ashtu dhe në fazën e vlerësimit të të dhënave për rritjen e precizionit.

Për të qenë i rëndësishëm dhe i përdorshëm një informacion ndihmës duhet të jetë i lidhur me variacionin e variablit kryesor.

Në SRS, kampioni zgjidhet pa përdorimin e një informacioni ndihmës. Atëherë, SRS mund ta marrim si një bazë krahasimi për të vlerësuar përfitimin në eficiencën e vlerësimit nga përdorimi i informacionit ndihmës në dizajnin me komplekse.

Informacioni ndihmës nuk është se luan ndonjë rol në procedurë SYS-sistematike të selektimit (nëse frame nuk është i sortuar sipas ndonjë rregulli, - variabli ndihmës). Kështu që nëse frame lista nuk është e renditur sipas ndonjë rregulli procedura SYS dhe SRS e selektimit tentojnë të jenë të njëjta. Por, ndryshe qëndron puna nëse frame lista është e sortuar paraprakisht sipas një variabli i ndihmës. Në metodën e quajtur Statifikim implicit- (me stratifikim të brendshëm), informacioni ndihmës përdoret pikërisht për të sortuar frame paraprakisht. Në këtë rast metoda SYS rezulton më efiçiente se ajo SRS nëse ka ndonjë lidhje midis variablave të renditjes dhe variablit në studim.

PPS- Metoda me probabilitet proporcional me masën është një metodë ku informacioni ndihmës ka një rol kyç. Një variabël ndihmës sopozohet të përdoret si masë e madhësisë së një elementi të popullatës. Kështu që çdo elementi të popullatës i bashkëngjitet një probabilitet i selektimit që varion sipas madhësisë. Efienca e kampionit rritet shumë nëse ky variabël është shumë i lidhur me variablin kryesor në studim. Metoda PPS është një metodë që përdoret gjërësisht.

STR- Metoda me shtresëzim bazohet gjithashtu shumë në variabla ndihmës. Në këtë metodë frame i popullatës ndahet në nën-popullata pa mbivendsje të quajtura strata dhe kampionimi ekzekutohet pavarësisht për secilën nga këto strata. Në qoftë se këto strata janë homogjene në lidhje me një variabël në studim, nëse variacioni brenda stratës është i vogël dhe variacioni ndërmjet stratave është i lartë atëherë STR është më efiçient se SRS.

Në metodën CLU- me shtresa popullata mendohet gjithashtu e ndarë në nëngrupe që i korespondojnë një zone fizike të caktuar të quajtura shtresa. Si hap i parë merret një kampion i këtyre shtresave. Shtresat janë elementet e selektimit parësor, dhe nëse vërojnë gjithë elementet e shtresës kemi të bëjmë me atë që quhet one-stage sample, kampion njëhapësh. Nëse një kampion të elementëve nga shtresa e zgjedhur zgjidhen sërish në një hap të dytë atëherë kemi të bëjmë me atë që quhet two-stage sample – kampion dyhapësh. Nëse shtresat janë homogjene (që kështu ndodh zakonisht) CLU është më pak efiçient se SRS. Por ky efekt i shtresave mund



REPUBLIKA E SHQIPËRISË
INSTITUTI I STATISTIKAVE

të reduktohet disi me stratifikimin e popullatës në studim, dhe tentohet rritja e efficiencies. Edhe pse të tilla në krahasim me SRS janë gjerësisht të përdorshme në kampionet pranë familjeve sepse sjellin reduktim të kostove të vërtetimit.

Përdorimi i informacionit ndihmës në fazën e kampionimit është tipik në vërtetimet përshkruese (diskriptive) ku numri i variablove në studim është shumë i vogël. Mund të arrihet të fitosh në efficiencë të kampionit nëse variabli (variablat) në studim dhe variabli ndihmës janë shumë të lidhur.

Informacioni ndihmës për kampionin e zgjedhur mund të përdoret gjithashtu në fazën e vlerësimit. Përdorimi i këtij informacioni në këtë fazë do të thotë: dizenjimi i kampionit mund të thjeshtohet dhe në fazën e estimimit, informacioni ndihmës mund të shfrytëzohet për variablat në studim. Në vazhdim, kërkesat për informacion ndihmës në metodat standarte janë më të dobëta se në rastin e mësipërm, sepse të dhënat ndihmëse janë në nivel njësie dhe mund të inkorporohen dhe në nivel të agreguar në procedurat e vlerësimit. Metoda standarte mund të përmendim metodën e regresionit dhe nën-rastin Ratio, apo atë post-stratification. Këto metoda përdorin informacionin ndihmës si asistent në vlerësimin e të dhënave dhe si të tilla njihen me emrin – Model-assisted (vlerësim me model).

Në vlerësimet me model regresioni ose raporti, totali i popullatës është një variabël ndihmës i vazhdueshëm supozuar i njohur. Modeli që përdoret është i tipit të regresionit linear.

Ratio është një nënrast i këtij modeli me intercept zero. Mund të kemi përfitime nëse variabli në studim dhe variabli ndihmës janë të lidhur. Por, metoda mund të jetë joeficiente nëse ekziston një zhvendosje nga origjina në modelin e zgjedhur.

Post-stratification ngjan metodën e stratifikimit por ky stratifikim bëhet pas procesit të selektimit. Kampioni i zgjedhur ndahet në nëngrupe jo- të mbivendosur të quajtur post-strata. Post-stratifikimi zakonisht ndodh për ajustimin për mospërgjigjet. Në ngjashmëri me stratifikimin synohet homogjenizimi i grupeve në lidhje me një variabël në studim.

Si përfundim, efficienca në vlerësim mund të rritet nëse një informacion ndihmës përdoret në formën dhe kohën e duhur.

6.3.2 Parametrat, vlerësuesit dhe matësit e cilësisë

Le të jetë totali i popullatës, i një variabli në studim y , parametri kryesor për të cilën interesohemi.

$Total\ T = \sum_{k=1}^N Y_k$ Y_k - vlerë e panjohur për çdo element k të popullatës



REPUBLIKA E SHQIPËRISË
INSTITUTI I STATISTIKAVE

Shumë parametra që gjerësisht përdoren në teorinë e kampionimit, të tilla si mesatarja, përpjesa, raporti dhe koeficientet e regresionit, mund të shpreheshin si funksione të totalit. Si gjithmonë, për të patur një vlerësim të parametrave të popullatës, p.sh. $\hat{t}$ zgjedhim një kampion prej n elementësh nga popullata.

Një vlerësim për totalin mund të merret nga formula:

$$\hat{t} = N\bar{y}$$

$$\hat{t} = N \sum_{k=1}^n Y_{k/n}$$

Pra duke përdorur një vlerësim për mesataren nga kampioni i zgjedhur mund të gjenerojmë një vlerësim për totalin e popullatës me formulat e mësipërme. Kuptohet kjo në rastin e një zgjedhje të thjeshtë, formula të tjera ju referohen dizajneve më komplekse.

Në teorinë e kampionimit vlerësuesit kërkohet të plotësojnë disa veti teorike.

Këto janë disa:

- Të pazhvendosur

$$E(\hat{t}) = T$$

Pritja matematike e vlerësuesit është sa ai i vërteti.

Zhvendosja është përkufizuar si diferencë $Bias(\hat{t}) = E(\hat{t}) - T$.

- Konsistent

Konsistenca është një veti e vlerësimit që lidhet me konvergjencën e tij tek parametri i vërtetë kur madhësia e kampionit rritet dhe vlerësuesi konvergjon tek e vërteta kur $n=N$.

- Preçisioni

Precisioni mund të matet me variancën $Var(\hat{t})$, devijimin standart ose me intervalin e besimit të parametrave. Sa më e vogël varianca aq më preçis parametri. Një vlerësim preçis quhet vlerësim eficient. Themi gjithmonë vlerësim preçis apo eficient nëse kemi variancë brenda limiteve të pranueshme. Sa më e vogël varianca aq më preçis vlerësimi, ose aq më eficient është ai.

- Saktësia



REPUBLIKA E SHQIPËRISË
INSTITUTI I STATISTIKAVE

Është një kombinim i zhvendosjes me precizionin dhe në një vlerësim matet me:

Means square error: $MSE(t^{\wedge}) = Var(t^{\wedge}) + Bias^2(t^{\wedge})$.

Në teorinë e kampionimit, sfida e metodologjistëve është të desinojnë kampione dhe vlerësojnë parametra me vlerësime të pazhvendur ose të paktën konsistent ku varianca e tyre të jetë sa më e vogël e mundshme. Kjo për të patur vlerësime sa më të besueshme nga vrojtimet me zgjedhje.

Devijimi standart (s.e), koeficienti i variacionit (c.v) dhe efekti i desinjimit (deff) të një vlerësimi shërbejnë si masa të cilësisë së vlerësimit.

Design effect (deff) (Kish 1965) mat eficiencën e dizajnit të zgjedhur të kampionit duke e krahasuar atë me zgjedhjen e rastit të thjeshtë (SRS) dhe jepet me:

$$deff = \frac{Var(t^{\wedge})}{Var_{SRS}(t^{\wedge})}$$

Duke përdorur efektin e dizenjimit, përmasa efektive e kampionit (n_{eff}) mund të përcaktohet si:

$$n_{eff} = n / deff(t^{\wedge})$$

Ku n është përmasa aktuale e kampionit.

Duke komentuar këtë formulë mund të themi:

1. $deff < 1$. Dizajni aktual është më eficient se SRS. Përmasa efektive e kampionit është më e lartë se ai aktual
2. $deff = 1$ Eficienca e kampionit aktual është e njëjtë me SRS.
3. $deff > 1$ Dizajni aktual është më pak efficient se SRS. Në këtë rast përmasa efektive e kampionit është më e vogël se përmasa aktuale e kampionit.

Në praktikën e vrojtimeve me zgjedhje, synimi mbetet gjithmonë të arrihet pika e parë, dhe për këtë qëllim përdoret informacioni ndihmës që në fazën e kampionimit. Desinjimet me strata dhe PPS shpesh përdoren për këtë qëllim. Për më tepër, eficienta mund të përmirësohet në fazën e vlerësimit të të dhënave duke përfshirë në të informacionin ndihmës përmes modeleve dhe analizave të ndryshme. Në kampionimet me klastera (rasti3) megjithëse janë më pak eficientë se SRS janë shumë të përdorshme sepse homogjeniteti i klasterave realizohet përmes një variabli ndihmës.



6.4 Teknikat bazë të kampionimit.

Teknikat bazë të kampionimit përfshijnë zgjedhjen e rastit të thjeshtë (SRS), zgjedhjen e rastit me strata dhe atë me probabilitet proporcional me përmasën (PPS). Këto metoda përdoren në selektimin final të elementeve apo në selektimin e atyre që quhen njësitë primare të vrojtimit (psu). Por shpesh në praktikë përdoren teknika të klusterizimit apo ato multi-stage. Stratifikimi i popullatës në shtresa jo të mbivendosura duke përdorur variabla ndihmës është një teknikë që rrit efikasitetin e kampionimit. Kurse në teknikat e përdorimit të klasterave aspekte të tjera të jetës së vrojtimit dhe grubullimit të informacionit janë motivimi kryesor i përdorimit të informacionit ndihmës në fazën e vrojtimit. Kryesisht këto vrojtme kanë kosto më ulët.

6.4.1 Zgjedhja e rastit të thjeshtë

Simple random sampling (SRS) njihet si forma bazë e kampionit me zgjedhje. Përdoret në rastet kur mungon çdo informacion ndihmës. Shërben si vijë bazë për krahasimi të efikasitetit për gjithë metodat e tjera. Kjo zgjedhje siguron që çdo element zgjidhet me të njëjtin probabilitet selektimi. Pra futet në vrojtmet me probabilitet të njëjtë. Si një teknikë bazë ajo e gjen veten të përfshirë në metodat e tjera.

Në zgjedhjen e rastit të thjeshtë, zgjidhen n elemente. Çdo element k i popullatës zgjidhet me të njëjtin probabilitet selektimi $\pi = \pi_k = n/N$. Çdo element i popullatës ka një probabilitet pozitiv.

Në praktikë, SRS mund të përdorim më zëvendësim (WR) ose pa zëvendësim (WOR).

Tipi i anketimit WOR i referohet rastit kur elementet në kampion nuk mund të zëvendësohen. Në të dyja këto raste probabiliteti i selektimit të elementeve është i njëjtë. Ajo që ndryshon është varianca. Por WOR SRS është më e eficientë se WR –SRS sepse varianca në WOR është më e vogël. Kjo shpjegon dhe përdorimin më të gjerë të WOR teknike në vrojtmet me zgjedhje kundrejt asaj WR.

Nën zgjedhjen e rastit të thjeshtë vlerësimi i totalit mund të jepet me:

$$\hat{t} = N \sum_{k=1}^n Y_{k/n} = N\bar{y}$$

Ose duke përdorur inversin e probabilitetit të selektimit (peshën) mund të shkruajmë:

$$\hat{t} = N \sum_{k=1}^n Y_{k/n} = N\bar{y}$$



REPUBLIKA E SHQIPËRISË
INSTITUTI I STATISTIKAVE

$$\hat{t} = \sum_{k=1}^n \frac{N}{n} Y_k = \sum_{k=1}^n W_k Y_k$$

Ku: W_k është pesha.

Varianca e vlerësuesit $\hat{t}$ në zgjedhjen SRS-WOR jepet me formulën:

$$v^*(\hat{t}) = N^2 (1 - n/N) (1/n) s^2$$

Ku: s^2 është varianca e variablit në studim.

Termi $(1 - n/N)$ njihet si faktori i rregullimit për popullësi të fundme (fpc). Në rastet kur sampli është i vogël (në kampioninët pranë familjeve) është e vogël, ky faktor është afër 1 pra rëndësia praktike e tij është e vogël. Por ndryshon puna në kampionet pranë ndërmarrjeve kur shpesh herë kampionet janë të mëdha, dhe ky faktor është i rëndësishëm në llogaritjen e variancës.

Në zgjedhjen e rastit me zëvendësim i vetmi ndryshim në formulën e variancës $v^*(\hat{t})$ është faktori fpc që jepet me $(1-1/N)$. Ky ndryshim tregon dhe eficiencën më të madhe të kampionit SRS-WOR.

Efekti i dizenjimit të $\hat{t}$ në zgjedhjen me zëvendësim $deff(\hat{t}) = (1-1/N) / (1 - n/N) > 1$, duke supozuar që n është më e madhe se 1 dhe më e vogël se N .

Po të shikojmë me vëmendje vumë në formulën e $deff$ krahasimin me variancën e SRS-WOR si referencë të kampionimit SRS. Kjo është zgjedhja e duhur por shumë herë në softet statistikore vendoset në këtë rol SRS-WR.

Për të përmbledhur, nëse sampling fraction n/N është i vogël, fpc (faktori për rregullim të madhësisë) do të jetë afër 1. Dhe anasjelltas nëse n afrohë tek N atëherë $v^*(\hat{t})$ do të reduktohet shumë. Në kushte ideale dhe Census varianca është 0.

Shumë softet statistikore ndihmojnë në kampionime të tilla. Për shembull në SAS, procedura SURVEYSELECT mund të përdoret për të dyja variantet e SRS. Në jetën reale zgjedhja e rastit e thjeshtë i referohet SRS-WOR- zgjedhjes së rastit të thjeshtë pa zëvendësim.

Shembull i një skeme kampionimi SRS-WOR – Kampionimi BERNULI

Kampionimi Bernuli jep një shembull të një skeme selektimi SRS-WOR. Në këtë metodë madhësia e kampionit nuk është e fiksuar por është një madhësi e rastit me pritje matematike n , madhësinë e dëshiruar të kampionit. Pra, kjo veti çon në një madhësi të kampionit me



REPUBLIKA E SHQIPËRISË
INSTITUTI I STATISTIKAVE

pritsmërinë $N\pi$ dhe variance $N(1 - \pi)\pi$ ku π – është probabiliteti i përfshirjes në kampion. Rastësia e madhësisë së kampionit është relativisht e parëndësishme në kampione të mëdha.

Le ta përshkruajmë shkurtimisht një kampion Bernuli. Për të patur një të tillë duhet të ndjekim hapat e mëposhtëm:

Hapi 1. Për një kampion me përmasë n fiksojmë probabilitetin e selektimit $0 < \pi < 1$ të tillë që:
 $\pi = N/n$.

Hapi 2. Ndërtojmë variablat

PROB = π ,

IND=0 për të gjithë elementët e popullatës N .

UNI, merr vlerën e një shpërndarje uniforme në rangun $(0, 1)$ – e njëjtë me fraksionin e kampionit.

Gjithashtu dhe një variabël PI- nga një pseudo random generator, një variabël me numra të rastit të çfarëdoshëm nga $(0, 1)$

Hapi 3. Rregulli për të vendosur nëse elementi k i përket kampionit apo jo është $UNI < \pi$. Pra nëse ky kusht plotësohet elementi është në kampion IND=1, ndryshe elementi nuk është në kampion IND=0.

Hapi 4. Kushti aplikohet për të gjithë elementet e frame listës.

Lista e gjithë elementëve ku IND=1 është kampioni i përfshirur BERNULI. Madhësia e kampionit e përfshirur në këtë mënyrë do të jetë i barabartë ose afërsisht n . Një kampion i tillë mund të programohet lehtësisht në SAS – software.

Zgjedhja e rastit sistematike SYS.

Në teknikën e zgjedhjes së rastit sistematike e cila përsëri është një zgjedhje e rastit me probabilitet të njëjtë $\pi = n / N$ veprohet kështu:

Hapi 1. Përcaktoni intervalin e kampionit $q = N/n$, ku q supozohet të jetë numër i plotë. Në të kundërt ai rrumbulloset.

Hapi 2. Selektoni një numër të plotë a – të rastit ndërmjet 1 dhe q .

Hapi 3. Zgjidhni elementet $a, a + q, a + 2q, a + 3q, \dots, a + (n-1)q$ në kampion.



REPUBLIKA E SHQIPËRISË
INSTITUTI I STATISTIKAVE

Kjo procedurë rezulton në zgjedhjen e n elementeve në kampion me metodën SYS. Në praktikë ka shumë mënyra për të marrë një zgjedhje sistematike. Kjo që ne përshkruam është metoda me një random start. Por mund të marrim një kampion me n elemente edhe duke marrë m random start dhe hap mxq. Kjo teknikë është e përshtatshme sidomos në përdorimin e të ashtëquajturës metodë replikative të llogaritjes së variancës (Wolter 2007).

6.4.2 Metoda e kampionimit PPS – Probabilitet proporcional me përmasën.

Metoda e kampionimit PPS, probabiliteti i selektimit varet nga madhësia e elementit. Variabli nga i cili varet ky probabilitet (variabli ndihmës) është shumë i rëndësishëm sepse reduktimi në variancë në këtë metodë selektimi pritet të vijë nëse ky variabël është shumë i lidhur me variablin në studim. Pra, në këtë metodë supozohet se njihet variabli Z për çdo element k të popullatës. Psh, në studimet pranë ndërmarrjeve numri i të punësuarve është një variabël ndihmës shumë i mirë, në vrojtimit në shkolla totali i numrit të nxënësve është gjithashtu një madhësi ndihmëse shumë e mirë. Metoda PPS rezulton shumë efficient nëse një variabël ndihmës shumë i mirë është i disponueshëm. Probabiliteti i përfshirjes në kampion të elementit k të popullatës është:

$$\pi_k = np_k = nZ_k / T_z \quad \text{ku:}$$

$$T_z = \sum_{k=1}^N Z_k \quad \text{totali i madhësisë (ndihmëse).}$$

p_k quhet *single-draw selection probability*.

Pra në kampionimin PPS, probabiliteti i përfshirjes së elementëve është i ndryshëm dhe varet nga variabli Z i zgjedhur si variabël ndihmës. Për këtë arsye metoda PPS është metodë me probabilitet të ndryshëm selektimi.

Një kampion PPS mund të zgjidhet me zëvendësim ose pa zëvendësim. Llogaritjet e probabilitetit të selektimit π_k janë më lehtë të menaxhueshme në skemën WR- sepse popullata mbetet e pandryshueshme pas çdo selektimi. Në PPS-WOR popullata ndryshon pas çdo selektimi kështu që probabiliteti i përfshirjes duhet ri-llogaritur për elementet e mbetur.

Parimet bazë të vlerësimeve nën PPS – WOR për një vlerësues të pazhvendosur të totalit mund të japim shkurtimisht si:



REPUBLIKA E SHQIPËRISË
INSTITUTI I STATISTIKAVE

$$\hat{t} = \sum_{k=1}^n w_k y_k = \sum_{k=1}^n y_k / \pi_k,$$

Ku: $1/w_k = \pi_k$ pesha e kampionimit. Vlerësuesi quhet *Horvitz-Thompson (HT)*

Ky është një vlerësues shumë popullor dhe i pazhvendosur.

Një vlerësues të variancës së totalit të vlerësuar mund të jepet me formulën:

$$\hat{v}(\hat{t}) = \sum_{k=1}^n \sum_{l=1}^n (w_k w_l - w_{kl}) y_k y_l,$$

Ku: $1/w_{kl} = \pi_{kl}$. Formula e variancës e HT- vlerësim të totalit përmban probabilitetin e përfshirjes së dy elementeve π_{kl} që është probabiliteti që të jenë në kampion të dy elementet njëkohësisht (elementet k, l). Kështu që llogaritja e tij sidomos në kampione shumë të mëdha është jo-praktike. Një përafrim të llogaritjes së variancës mund të jepet me formulën e mëposhtme:

$$\hat{v}(\hat{t}) = N^2 (1/n) \sum_{k=1}^n (y_k / (N p_k) - \bar{y})^2 / (n-1),$$

e cila i korespondon PPS-WR në të cilën $1/w_{kl} = \pi_{kl}$ janë zero sepse selektimi i elementeve supozohen të pavarura nga njëra-tjetra.

Ka skema të ndryshme praktike të selektimit – PPS. Më të hasurat janë metoda komulative me zëvendësim ose pa zëvendësim, PPS-sistematike ku është i rëndësishëm sortimi, apo ajo POISON si një metodë me probabilitete të ndryshme selektimi e ngjashme me skemën Bernuli. Ndryshon nga ajo në faktin që është skemë me probabilitete të ndryshme selektimi, pra dhe kolona e probabilitetit të selektimit është gjithashtu nga një shpërndarje e rastit. Përveç faktit që ka përmasë të kampionit të rastit kjo teknikë konsiderohet shumë e përdorshme sepse njihen probabilitetet e selektimit të dy elementeve $\pi_{kl} = \pi_k \pi_l$ e rëndësishme kjo në llogaritjen e variancë. Në librin e Brewer & Hanif (1983) mund të gjendet një burim i mirë për teknikat më të përdorshme të metodës PPS të selektimit të kampionit. Gjithashtu, teknikat më të përdorshme PPS janë implementuar në SAS në procedurën SURVEYSELEC.



6.4.3 Kampionim me strata

Në metodën me Strata popullata ndahet në nëngrupe jo të mbivendosura me njëra tjetrën. Në secilën prej tyre mund të aplikohet skema e selektimit në mënyrë të pavarur (qoftë kjo SRS, SYS, apo PPS, etj). Në përgjithësi ka disa arsye për të cilat popullata stratifikohet:

1. Qëllimi kryesor për të cilën përdoret stratifikimi është përfitimi në efikasitet. Ndërtimi i një kampioni më efikas si rezultat i garantimit të homogjenitetit të stratës në lidhje me një variabël kryesor në studim.
2. Për arsye administrative, shumë popullata janë të ndara geografikisht. Këto ndarje mund të përdoren dhe në stratifikim. Psh stratë: mund të shërbejë Rajoni- urban/rural.
3. Stratifikimi lejon një përdorim fleksibël të informacionit ndihmës për kampionim dhe vlerësim të të dhënave. Psh. Mund të përdoret metoda PPS – bazuar në informacionin ndihmës në micro për selektimin e elementeve dhe vlerësimi ratio, regresion apo metoda kalibrimi në fazën e vlerësimit të të dhënave në varësi të informacionit ndihmës të disponueshëm.
4. Stratifikimi garanton gjithashtu prezencën në kampion të popullatave të vogla. Kjo nënkupton që probabiliteti i selektimit mund të variojë nga strata në stratë. Kjo variancë kontrollohet nëpërmjet teknikave që njihen me emrin teknika alokimi.

Në kampionin me strata vlerësimi i totalit jepet me formulën:

$$\hat{t} = \sum_{h=1}^H \hat{t}_h \quad \text{ku: totali në çdo stratë } h$$

$$\hat{t}_h = \sum_{k=1}^{n_h} y_{hk} / \pi_{hk} = \sum_{k=1}^{n_h} w_{hk} y_{hk}$$

Vlerësimi i variancës së këtij totali jepet me shumën e variancave në çdo stratë.

$$\hat{v}(\hat{t}) = \sum_{h=1}^H \hat{v}(\hat{t}_h)$$

Në rastin e alokimit proporcional në strata ku N_h / n_h është i njëjtë me N/n formula e vlerësimit të totalit transformohen lehtësisht në:

$$\hat{t} = \sum_{h=1}^H N_h / n_h \sum_{k=1}^{n_h} y_{kh}$$



REPUBLIKA E SHQIPËRISË
INSTITUTI I STATISTIKAVE

$$\hat{t} = \sum_{h=1}^H N_h / n_h \sum_{k=1}^{n_h} y_{kh} = N / n \sum_{h=1}^H \sum_{k=1}^{n_h} y_{kh}$$

Dhe reflektojnë të ashtëquajturën veti të vetëpeshimit të një kampioni të tillë.

6.4.4 Kampionimi me Clastera

Për të realizuar një kampionim me clastera një kampion prej clusterash selektohet fillimisht nga një frame prej clusterash. Si clastera mund të jenë psh: kompanitë (që përmbajnë si elemente të tij punonjësit), klasat (që kanë si elemente nxënësit), zonat gjeografike (që kanë si elemente familjet). Ky kampion mund të tërhiqet duke përdorur një nga metodat bazë të kampionimit si SRS, SYS, or PPS. Kjo metodë është gjithmonë e përdorshme me qëllim uljen e kostos së vërtetimit. Në Metodën një-hapëshe me clastera të gjithë elementët e clasterit intervistohen, kurse në atë dy hapëshe një kampion i elementëve të klasterit selektohen përsëri duke përdorur një nga teknikat bazë të selektimit.

Një nga avantazhet e këtij kampionimi është se nuk nevojitet frame i të gjithë elementëve të popullatës në studim por vetëm i njësive primare të selektimit.

Kështu që informacioni ndihmës në nivel klasteri është i domosdoshëm për qëllime të stratifikimit apo të selektimit të klasterave.

Psh: Në metodën dy hapëshe të dizenjimit të kampionit, shpesh përdoret metoda PPS e kampionit në hapin e parë. Një kampion i vetëpeshuar përfitohet në rastin kur elementet primare të zgjedhura kanë madhësi të njëjtë.

Metoda me klastera ul koston e vërtetimit (mendon si shembull popullsinë e shpërndarë shumë gjeografikisht) por ka dhe një kosto të uljes së efijencës së kampionit. Në përgjithësi llogaritet rritje e përmasës së kampionit për këtë arsye.

Nëse secila nga këto klastera do të ishte një pasqyrë e strukturës të popullatës në studim atëherë ne do të kishim përfitimin maksimal nga një metodë e tillë, pra devijimi standart i parametrut në studim nuk do të kalonte atë të metodës SRS, ose e shprehur ndryshe do të përfitonim kampionin me deff afër 1, të krahasueshëm me metoden SRS. Por zakonisht në praktikë hasim klastera homogjene duke rritur atë që quhet intra-clas homojenitet dhe rënie të efijencies së këtij kampioni në krahasim me SRS. Këto praktika janë të shumë diskutuara në tekstet e statistikës praktike psh Lehtonen & Pahkinen (2004, kapitullin 5, 7–9) japin shumë shembuj të botës reale të këtij fenomeni.



REPUBLIKA E SHQIPËRISË
INSTITUTI I STATISTIKAVE

Për kampionet me klastera efekti i dizenjimit të kampionit është mbi 1, pra janë kampione më pak eficiente se SRS, pra ju korespondon një madhësi efektive e kampionit më e vogël se ajo e zgjedhur $n. n_{eff} = n / deff$.

7. Përcaktimi i përmasës së kampionit dhe alokimi i tij.

Ky kapitull do të merret me metoda të përcaktimit të përmasës së kampionit në një vërtetim me kampion. Qëllimi është trajtimi teorik jo për të sjellë dëkë të re në teorinë e vërtetimeve me zgjedhje por për të kuptuar bazën teorike të aplikimit praktik të një ankete në Shqipëri. Ky kapitull merr përsipër të analizojë influencën e faktorëve në këtë proces, të ilustrojë si të llogaritet përmasa e një kampioni për një precizion të caktuar, si të alokohet një kampionin i stratifikuar dhe japë disa komente për mënyrën e alokimit të një kampioni. Gjithashtu do të jepen disa sygjerrime në bërjen e më të mirë të këtij procesi me qëllimin kryesor vlerësimin e një treguesi statistikor cilësor.

7.1 Faktorët që ndikojnë në përcaktimin e përmasës së kampionit.

Një nga pyetjet më të zakonshme kur një studiues vendos të bëjë një vërtetim me zgjedhje si metodë kërkimi të punës së tij është: Sa njësi më duhet të studioj për të patur një rezultat të besueshem? Për t'i dhënë përgjigje kësaj pyetje nuk është shumë e thjeshtë. Kjo është një nga përgjegjësitë e metodologjistikës të vërtetimeve me zgjedhje. Kjo pyetje është e lidhur me precizionin e vlerësuesit në një nivel përfaqësimi (domëin) të caktuar. Kur një metodologjist nis të japë përgjigje për këtë pyetje, ai ngre shumë pyetje të tjera si:

- Kush është treguesi kryesor në këtë studim?
- Ka patur studime të tjera të mëparshme për llogaritjen e këtij treguesi apo është hera e parë që kërkohet të llogaritet?
- A kemi informacion në regjistër për këtë tregues, për një periudhë të kohe? etj

Pse ai bën këtë listë me pyetje do t'a kuptojmë përmes këtij kapitulli. Disa nga faktorët që influencojnë në përmasën e një kampioni janë:

- Përmasa e popullatës në studim
- Variacioni i treguesit kryesor në studim në popullatë
- Precizioni i kërkuar për këtë tregues.
- Dizënjimi i kampionit dhe vlerësuesi që do të përdoret për të.
- Parashikimi për treguesin e përgjigjes.

Të gjithë këta faktor janë të lidhur me procesin e kampionimit. Formulatat për llogaritjen e përmasës së kampionit kërkojnë që të plotësohet kriteri i një precizioni të caktuar në një nivel specifik



REPUBLIKA E SHQIPËRISË
INSTITUTI I STATISTIKAVE

publikimi. Ato duhet të marrin në konsiderate faktorët e lartpërmendur. Gjithashtu, pavarësisht se në një studim mund të maten disa tregues llogaritet për caktimin e përmasës së kampionit mbështen gjithmonë në një tregues kryesor.

Përmasa e një kampioni nuk kupton vetëm grand totalin e përmasës së kampionit por gjithashtu përmasën në nivele të caktuara publikimi. Kështu pra pyetja e përmasës së kampionit dhe e precizionit të kërkuar mund të shtrihet në grupet e publikimit. Një vendim i qartë për kompozimin e këtyre grupeve, dhe precizionin e kërkuar të treguesit në to janë të domosdoshme në fillimin e kësaj pune.

Ka gjithashtu dhe faktorë të tjerë jo të lidhur me kampionim që nuk i përmendëm më lart por që ndikojnë në përcaktimin e përmasës së kampionit si:

- Buxheti
- Koha për realizimin e këtij aktiviteti.
- Kufizime të tjera të lidhura me terrenin.

Ka dy strategji të caktimit të përmasës së kampionit:

- Së pari caktimi i totalit të përmasës së kampionit dhe më pas alokimi i tij në grupe.
- Ose, përcaktimi i një kriteri precizioni të kërkuar në çdo grup publikimi dhe përmasë kampioni në secilin prej tyre, kështu që totali i n -i kampionit do të jetë një shumatore e këtyre mini-kampionimeve.

Disa vlerësime se si këta faktorë ndikojnë në përmasën e kampionit.

- Faktori kryesor i përmendur më lart është precizioni i kërkuar për treguesin kryesor.

Precizioni i kërkuar i një vlerësimi (supozoni $\hat{t}$ - për totalin) mund të shprehet në terma të devijimit standart $SE(t)$, të marzhit të gabimit $z \times SE(t)$ ose të intervalit të besimit, ose të koeficientit të relacionit $SE(t) / \hat{t}$.

Përcaktimi i përmasës së kampionit gjithmonë kërkon specifikimin e një precizioni të kërkuar të shprehur me një ose disa nga termat e sipërpërmendura. Në formulat e ilustrimit të mëposhtëm ne lehtësisht do të shohim që përmasa e një kampioni rritet nëse kërkohet një precizion më i lartë i vlerësimit, pra sa më preciz ta kërkojmë vlerësuesin aq më shumë njësi vrojtimi duhen pyetur. Pra, përmasa e kampionit është ngushtësisht e varur nga precizioni i kërkuar i vlerësuesit.

- Përmasa e popullatës.

Nga praktika është parë që përmasa e popullatës ndikon në përmasën e kampionit deri në njëfarë maksimumi të përmasës së kampionit dhe më pas ai nuk influencon më, pra më pas përmasa e kampionit mbetet e njëjtë megjithëse merr në studim popullatë shumë herë më të madhe.



REPUBLIKA E SHQIPËRISË
INSTITUTI I STATISTIKAVE

- Dizenjimi i kampionit dhe vlerësuesi.

Kur caktohet përmasa e kampionit është Efekti i desinjimit $deff = \text{Var}(\text{desing})/\text{Var}(\text{SRS})$ që merret gjithmone parasysh. Nëse na duhet të arrijmë një precision të caktuar të vlerësuesit, një variancë të lejuar të tij, si dhe duke patur parash faktin që desinjimi i kampionit ndikon në variancën e këtij vlerësimi atëherë është e qartë që efekti i dizajnit duhet të merret në konsideratë nëse parashikojmë jo një kampion SRS. Në një SRS $deff = 1$. Pra $deff$ shumëzon gjithmonë faktorët e tjerë në formulën e kalkulimit të përmasës së kampionit.

- Treguesi i përgjigjes.

Duke mos harruar që në një vrojtim me zgjedhje hasim mospërgjigje, është e rëndësishme që këtë faktor t'a marrim në konsideratë në llogaritjen e përmasës së kampionit.

Pra, faktorët e lartpërmendur kërkojnë që të ketë një vlerësim për variancën e treguesit kryesor në studim dhe treguesin e përgjigjes. Pra, nëse ekziston një studim i mëparshëm mbi të njëjtin regjistër, apo nëse është bërë një studim me një desinjim të ngjashëm apo të njëjtë duhen marrë në konsideratë varianca e treguesit, $deff$ për këtë tregues, treguesi i përgjigjes apo precizioni i arritur i këtij treguesi nga një studim i mëparshëm. Të gjitha këto janë vlerësime që ndikojnë në përmasën e kampionit.

7.2 Disa nga formulat më të përdorshme në caktimin e përmasës së kampionit

Më poshtë do të jepen disa nga formulat më të përdorshme për caktimin e përmasës së kampionit që ka si qëllim llogaritjen me një precision të fiksuar të indikatorëve si: mesatare, total, dhe përpjesë.

1. Kur jepen precizioni për një vlerësim të mesatares $\hat{y}$, në rastin SRS.

Marzhi i gabimit dhe formula e devijimit standart të një vlerësimi të treguesit të marrë në studim, në rastin konkret vlerësim i mesatares në kampionim SRS pa zëvendësim, merren në konsideratë kur nisim të caktojmë përmasën e kampionit. Formula që bën këtë vlerësim është:

$$S(\hat{y}) = \sqrt{\left(1 - \frac{n}{N}\right)} \frac{\hat{S}}{\sqrt{n}}$$

Ku $\hat{S}$ është një vlerësim i devijimit standart të Y- variablilit kryesor në studim. Zakonisht ky informacion nuk ekziston në regjistër dhe bëhet një vlerësim i tij nga të dhënat e një vrojtimi të mëparshëm.



REPUBLIKA E SHQIPËRISË
INSTITUTI I STATISTIKAVE

Duke vendosur një marzh gabimi e , në terma të z , kemi:

$$e = z \sqrt{\left(1 - \frac{n}{N}\right) \frac{\hat{S}^2}{n}}$$

ku z -niveli i besimit në përqindje. Ai shpreh se sa përqind të devijimit standart të treguesit është e lejuar të gabojmë, është e lejuar të shmangemi nga pika e vlerësimit.

Duke zgjidhur ekuacionin e mësipërm për n kemi:

$$n = \frac{z^2 \hat{S}^2}{e^2 + \frac{z^2 \hat{S}^2}{N}}$$

Kështu pra, me qëllim caktimin e n , kërkohet:

- një marzh gabimi i lejuar, e ;
- një vlerë z që i takon një niveli besimi në përqindje.
- përmasa e popullatës në studim, N ;
- një vlerësim për devijimin standart të treguesit kryesor, $\hat{S}^2$

2. Kërkohet të vlerësohet përpjesa $\hat{p}$, me një precision të kërkuar, në rastin SRS.

Rasti është analog me të mësipërmin, por meqenese treguesi për të cilin interesohemi është përpjesa atëhere formula e llogaritjes së përmasës së kampionit rrjedh nga formula e vlerësimit të variancës së përpjesës së indikatorit në studim, e shprehur kjo në terma të marzhit të gabimit të lejuar.

Duke patur parasysh se në një popullatë të madhe, $\hat{p}$ vlerësimi i përpjesës ka një shpërndarje normale si dhe duke ditur formulën për vlerësimin e variancës të një variabli binar (0,1)

$$\hat{S}^2 = p(1-p) \quad (*)$$

Ne mund të fitojmë formulën e mëposhtme duke bërë zëvendësimin (*) në formulën e rastit 1. ku interesoheshim për mesataren.

Kështu pra:

$$n = \frac{z^2 p(1-p)}{e^2 + \frac{z^2 p(1-p)}{N}}$$



REPUBLIKA E SHQIPËRISË
INSTITUTI I STATISTIKAVE

Për zbatimin e kësaj formula n'a duhet të kemi një vlerë për parametrin p. Nëse ekziston në regjistër një vlerë për një kohë të mëparshme studimi mund të përdoret, nëse jo mund të përdoret një vlerësim nga një vrojtim i mëparshëm dhe e fundit në mungesë totale të një studimi paraprak sygjerohet që ky parameter të merret $\hat{p} = 0.5$ për të marrë vlerën maksimale të n –përmasës së kampionit.

Thëksojmë faktin se formulat e mësipërme e kanë marrë parasysh korrigjimin për popullatë të fundme me faktorin $(1-n/N)$.

3. Kërkohet të vlerësohet totali me një precision të kërkuar, rasti SSRS- SRS me strata.

Nisur nga precizioni i shprehur me koeficientin e variacionit:

$$CV = S(\hat{t}) / \hat{t} = \alpha$$

Duke patur në kusht që ky koeficient nuk duhet të kalojë një vlerë të caktuar α :

$$\frac{\sqrt{V(t)}}{t} \leq \alpha,$$

$$\text{Dhe duke zëvendësuar } V_{\min}(t) = \frac{N^2 (\sum_{h=1}^H (W_h S_h)^2)}{n} - N \sum_{h=1}^H W_h S_h^2, \text{ dhe zgjidhur për } n$$

arrijmë në formulën:

$$n = \frac{N^2 (\sum_{h=1}^H W_h S_h)^2}{\alpha^2 t^2 + N \sum_{h=1}^H W_h S_h^2}$$

e cila i referohet rastit të SSRS- Stratified Simple Random Sample.

Për aplikimin e kësaj formule nevojiten:

- $W_h = N_h / n_h$ peshat në nivel strate;
- Varianca e variablilit në studim (y) në popullatë ose në mungesë të saj një vlerësim, $\hat{S}_h^2$ në nivel strate.
- α -precizioni e shprehur, koeficienti i variacionit maksimal i lejuar i vlerësimit i shprehur në përqindje.

² Lorn, page



REPUBLIKA E SHQIPËRISË
INSTITUTI I STATISTIKAVE

- Një vlerë e totalit –Y nga regjistri, për një periudhë të mëparshme kohe ose në mungesë të saj një vlerësim $\hat{t}$ nga një vrojtım i mëparshëm.
- N-totali i njësive në studim.

Duhet patur gjithmonë parasysh se në të gjitha formulat e mësipërme si për mesataren, totalin, dhe përpjesën precizioni duhet shprehur në përqindje.

Për finalizimin e kalkulimit të përmasës së kampionit na duhet të modifikojmë (rritjen e kampionit) për efekt dizajni nëse kemi zgjedhur që kampioni ynë nuk do të jetë SRS:

$$n_2 = deff * n$$

Gjithashtu si hap i fundit duhen marrë parasysh mospërgjigjet.

$$n_3 = 1/r * n_2$$

Ku: r- Koeficienti i përgjigjes(response rate)

7.3 Alokimi i kampionit

Një metodë zgjedhje qoftë kjo SRS, SYS, apo PPS, të gjitha shkencore por shumë pak e përdorshme në formën e saj të thjeshtë. Efektiviteti i saj është i lartë nëse popullata në studim është homogjene në lidhje me variablin në studim (variancë të vogël). Por, nga praktika është vënë re se popullata në studim nuk është kaq homogjene. Psh: ndërmarrjet ekonomike nuk janë kaq homogjene në lidhje me xhiron që realizojnë. Kjo çon metodologjistin të mendojë të kalojë në kampionim me Strata – grupe dhe modifikojë metodën e selektimit fillimisht në krijimin e grupeve dhe më pas aplikimin e SRS në secilën prej tyre. Psh: po qe se marrim në shqyrtim vetëm ndërmarrjet e vogla të bar restoranteve mund të themi se homogjeniteti i shifrës së afarizmit rritet. Parimi bazë i stratifikimit është homogjeniteti i variablit në studim. Pra, themi se kemi bërë shtresëzim të mirë, ndarje në strata në mënyrën më eficiente të mundshme nëse kemi siguruar rritjen e homogjenitetit të variablit në studim brenda stratave, njësoj maksimizimin e heterogjenitetit ndërmjet stratave.

Në metodën me Strata popullata ndahet në nëngrupe jo të mbivendosura me njëra tjetrën. Në secilën prej tyre mund të aplikohet skema e selektimit në mënyrë të pavarur (qoftë kjo SRS, SYS, apo PPS etj.)

Pra, bashkë me stratifikimin ngrihet dhe problemi i alokimit të kampionit nëpër strata me qëllim rritjen e saktësisë së vlerësimit.



REPUBLIKA E SHQIPËRISË
INSTITUTI I STATISTIKAVE

Mënyrat e hasura të alokimit mund t'i ndajmë në dy kategori, proporcional me përmasën e popullatës N ose disproporcionale me të:

Metoda e alokimit proporcional është shumë e përhapur në të gjitha rastet kur nuk ekziston informacion për shpërndarjen e variablit në studim nëpër strata apo domeine. Në këto raste shpërndahet n në n_h në mënyrë:

$$n_h = n \frac{N_h}{N}$$

Pra alokimi N-proporcional nënkupton të njëjtën përpjesë nga secila stratë.

Kështu: $a_h = n_h/n = N_h/N$ faktori i alokimit është i njëjtë në çdo stratë.

Alokimi disproporcional.

Kjo metodë quhet disproporcionale sepse është e tillë në terma të njësisë së popullatës por ajo është proporcionale mbi një parameter tjetër. Si metoda të tilla mund të përmendim Y-proporcional, Alokimin Optimal, alokimin NEYMAN.

Alokimi Y-proporcional- është alokim proporcional në lidhje me një variabël ndihmës Y.

$$n_h = n \frac{Y_h}{Y}$$

Alokimi Optimal – është alokim proporcional me prodhimin e përmasës së popullatës dhe devijimit standart të një variabli ndihmës shumë të lidhur me variablin kryesor në studim. Ky alokim merr parasysh dhe koston për njësi të vërtetimit në stratë. Pra ky alokim ka nevojë dhe për një model të funksionit të koston.

Formula e përdorur është:

$$n_h = \left(\frac{\frac{N_h S_h}{\sqrt{c_h}}}{\sum_{l=1}^H \frac{N_l S_l}{c_l}} \right) n$$

ku c_h kostoja për njësi të vërtetimit në stratën h ($h = 1, 2, \dots, L$).

Alokimi Neyman- është një rast i veçantë i alokimit optimal që nuk merr parasysh funksionin e koston.



REPUBLIKA E SHQIPËRISË
INSTITUTI I STATISTIKAVE

$$n_h = n \frac{N_h S_h}{\sum_{l=1}^H N_l S_l}$$

Pra ky alokim është proporcional me $N_h S_h$ prodhimin e numrit të njësive në stratë me devijimin standart të një variabli ndihmës.

Përfundim

Në përfundim mund të themi se përcaktimi i përmasës së kampionit është një pikë thelbësore në arritjen e një vlerësimi statistikor me precision të caktuar. Ai kërkon të kemi një objektiv shumë të qartë të objektivit të anketës dhe precizionit të kërkuar të treguesit në fushat e publikimit. Pra, plani i tabulimit të vrojtimit duhet të përgatitet para fillimit të vrojtimit me zgjedhje.

Mund të themi gjithashtu se procesi i caktimit të përmasës së kampionit në praktikë nuk është thjesht zbatim i formulave teorike. Ai është një proces studimi dhe vendimmarrje në çdo hap për të caktuar çfarë të dhënash duhen përdorur dhe si, është vendimmarrje midis zgjedhjeve praktike dhe konfliktit që këto zgjedhje shkaktojnë me kufizimet e vendosura në lidhje me precizionin e kërkuar, buxhetin, koston e vrojtimit dhe kohën e nevojshme për publikimin e rezultatit të arritur.

Nuk ka zgjidhje perfekte për përcaktimin e saj megjithëse ekzistojnë formula precise të llogaritjes. Formulatat përpiqen të kontrollojnë gabimet e vrojtimit por shumë kufizime të tjera duhen marrë në konsideratë në këtë punë.

8. Vlerësimi i të dhënave

Me vlerësim të treguesit kuptojmë mënyrën me të cilin metodologjisti arrin të llogarisë parametrin (total, mesatare, perpjese) për popullatën në studim në bazë të informacionit të mbledhur nga vetëm një kampion i popullatës.

Parimi prapa një vlerësimin të tillë nga një vrojtimit probabilitar është që çdo njësi e kampionit të përfaqësojë jo vetëm vetveten, por edhe disa njësi të popullsisë të vrojtuar. Pra që në procesin e parë të dizenjimit të kampionit, çdo njësi ka probabilitetin e vet për të qënë në kampion ose me sakte ka numrin e vet të njësive që përfaqëson në kampion. Përcaktimi i kësaj peshe është një pjesë e rëndësishme e procesit të vlerësimit. Kjo peshë modifikohet në procesin përfundimtar të vlerësimit të të dhënave për të arritur në atë finale. Dy arsye më të zakonshme për të bërë këto



REPUBLIKA E SHQIPËRISË
INSTITUTI I STATISTIKAVE

rregullime të saj janë marrja në konsiderate e mospërgjegjeve dhe përdorimi i të dhënave ndihmëse.

Pasi peshat përfundimtare janë llogaritur, aplikohen në të dhënat e kampionit në mënyrë që të merren vlerësime për gjithë popullatën në studim. Parametrat e popullsisë si psh, mesataret dhe përqindjet janë vlerësuar në mënyrë tipike për një gamë të gjerë të karakteristikave të mbledhura nga njësitë e kampionit. Këto karakteristika, shpesh të referuara si variabla në teorinë statistikore, mund të jetë cilësore - për shembull, seksi apo statusi martesor, ose sasiore - për shembull, mosha apo të ardhurat. Në varësi të llojit të të dhënave, përshtaten dhe formulat për vlerësimin e parametrave të popullatës.

Një pjesë e shumë rëndësishme e vlerësimit është llogaritja e madhësisë së gabimit të vlerësimit si një vlerësim për cilësinë e llogaritjes së treguesit. Nëse kampioni është probabilitar kemi mundësi të llogarisim gabimin e kampionimit. Qëllimi i këtij kapitulli është të njihemi me konceptet bazë të llogaritjes së peshës, për të dhënë vlerësime për parametrat kryesore të totalit të popullatës në studim dhe të llogarisim gabimet e kampionimit për treguesit në kampionim probabilitar.

8.1 Peshimi

Peshimi i të dhënave është procesi i vendosjen së një variabli të ri në skedarin e njësive të vrojtimit i cili për çdo rekord të tij tregon se sa njësi të popullatës përfaqëson.

Hapi i parë për të bërë vlerësime është vendosja e një peshe për secilin njësi në një kampion.

Ky proces kalon nëpër disa faza. Në kampionet probabilitare ai nis peshën fillestare ose atë të desinjimit. Kjo peshë do të përputhej me atë finale nëse kushtet e kampionit do të ishin ideale. Pra, nëse frame lista e përdorur për kampionim do të që e azhornuar në kohë dhe nuk do të kishte probleme të variablave të kontaktit për të mundësuar vrojtimin, as të informacionit ndihmës, nëse në procesin e vrojtimit nuk do të ekzistonin mos-përgjigjet, nëse kampionimi do të mernte parasysh dhe grupet shumë të vogla të studimit të rezultateve etj. Në praktikë natyrisht që nuk ndodh kështu. Të gjitha sa më sipër ekzistojnë e kështu pesha fillestare mund të çojë në vlerësim jo të saktë ose vlerësim që bie ndesh me informacionin ndihmës që ekziston për popullatën. Në këtë mënyrë lind nevoja e axhustimit të peshës fillestare. Axhustimi i peshës bëhet për të kompensuar për:

- probabilitet jo të njëjta gjatë kampionimit.
- për të kompensuar për mos-përgjigjet dhe problemet e mbulimit të regjistrimit.
- për të axhustuar për informacionin ndihmës ekzistues të shpërndarjes së popullatës në studim.



REPUBLIKA E SHQIPËRISË
INSTITUTI I STATISTIKAVE

Pra, peshimi nis me krijimin e peshës fillestare ose të dizajnit. Mund ta hasni dhe me emrin pesha bazë. Nuk është gjë tjetër veçse inversi i probabilitetit π_i të selektimit të njësisë së vërtetimit.

Kjo peshë kompenson për probabilitete jo të njëjta të selektimit. Jepet me formulën:

$$w_{di} = 1/\pi_i.$$

ku π_i është probabiliteti me të cilin është përzgjedhur njësia i – për dizajn.

Pesha fillestare ose pesha e dizajnit të kampionit mund të mendohet si numri mesatar i njësive të popullatës që përfaqëson njësia në vërtetim i përcaktuar nga dezajni i kampionit dhe mënyra e selektimit të njësive. Peshat e dizajnit, w_d (ku d qëndron për dizajn) për një njësi që është në kampion është inversi i probabilitetit përfshirjes të saj, π . Vini re se për një dizajnim me shumë – faza apo me shumë – hapa, probabiliteti i një njësie të përzgjedhur është probabiliteti i kombinuar i zgjedhjes në çdo fazë apo hap. Kështu, për një kampion në dy – faza, ku probabiliteti i një njësie kampioni është π_1 në fazën e parë dhe π_2 në fazën e dytë, pesha fillestare ose e dizajnit për këtë njësi vërtetimi është:

$$w_d = \frac{1}{\pi_1} \times \frac{1}{\pi_2}$$

Të dhënat e vërtetimit jepen në përgjithësi në një skedar me një rekord për çdo njësi të saj. Kujtojnë se me marrjen e kampionimit me metoda probabilitare, çdo njësi ka një probabilitet të njohur, π , për të qenë i kampionuar.

8.1.1 Trajtimet e mospërgjigjeve nëpërmjet peshimit të të dhënave.

Të gjitha vërtetimet vuajnë nga mospërgjigjet e marra, gjë e cila ndodh kur i gjithë informacioni ose disa prej informacioneve të kërkuar për njësitë e kampionit janë të padisponueshëm për disa arsye. Mospërgjigjet klasifikohen në dy tipe, mospërgjigje e pjesëshme dhe mospërgjigje të total njësive. Mospërgjigja e pjesëshme ndodh kur informacioni është në dispozicion vetëm për disa variabla, të tilla si kur i intervistuari përgjigjet vetëm për një pjesë e pyetësorit. Në këtë rast, trajtimi më e zakonshëm është për t'i plotësuar vlerat që mungojnë. Me metoda të plotësimit (imputimit) të të dhënave të munguara do të merremi tek nën çështja – **Imputimi i të dhënave**. Ky kapitull fokusohet në peshimin dhe vlerësimet. Peshimi i të dhënave mund të shikohet dhe si metodë e trajtimit të mospërgjigjeve totale. Mospërgjigjeve totale e rastit ndodh kur të gjitha ose pothuajse të gjitha të dhënat për një njësi të kampionit nuk gjenden. Kjo mund të ndodhë kur një njësi e kampionit refuzon të marrë pjesë, nuk është bërë kontakt, njësia nuk mund lokalizohet apo informacioni i marrë është e papërdorshme. Në disa rrethana të veçanta ku vlerësimet



REPUBLIKA E SHQIPËRISË
INSTITUTI I STATISTIKAVE

proporcionale të mesatares pa rregullimet për mospërgjigjet në total janë të njëjta me ato të prodhuara nga përdorimi i axhustimeve për mospërgjigjet. Por, e përgjithshme është që, mospërgjigjet e njësive që nuk përgjigjen çon në nënvlerësim të totalit (për shembull, nënvlerësim të madhësisë së popullsisë ose ardhurat totale).

Mënyra më e zakonshme që të merresh me mospërgjigjet totale është ajo e cila rregullon peshat e dizajnit bazuar në supozimin se njësitë të cilat janë përgjigjur përfaqësojnë dhe të papërgjigjurat, përgjigjet dhe mos - përgjigjet. Kjo është llogjike nën supozimin që, për karakteristikat e matura të vrojtimit, njësitë që nuk përgjigjen janë si ato të cilët përgjigjen.

Pra hipoteza bazë e një trajtimi të tillë është fakti që njësitë e papërgjigjura janë po aq të rastësishme sa dhe ato të zgjedhurat dhe nuk ekziston ndonjë tendencë e njësive të papërgjigjur. Psh: të përgjigjurit janë ndërmarrje ekonomike me vlera të shifrës së afarizmit shumë të ndryshëm nga të pa përgjigjurit, ka një tendencë në mospërgjigje të ndërmarrjeve me xhiro të ulët. Gjithsesi nuk ka metodë që të bëjë një vlerësim të tillë, është zonë pa informacion. *Praktika më e mirë e mbrojtjes nga shtrembërimi për mospërgjigje është të bëhet maksimumi për minimizimin e tyre.* Por, gjithsesi ato gjithmonë ekzistojnë dhe natyrisht do të trajtohen në vlerësim.

8.1.2 Faktorët e rregullimit të peshave për Mospërgjigjet

Rregullimi i peshave të dizajnit për mospërgjigjet është një rishpërndarje e peshës tek të përgjigjurit. Kjo bëhet përmes caktimit të koeficientëve të mospërgjigjeve dhe shumëzimit të saj me peshën e dizajnit për të dhënë peshën e rregulluar për mospërgjigjet.

$$W_{nr} = W_d * K_r$$

Ajo që duhet të theksohet është se dhe në një regjistrim të plotë (census) gjithashtu mund të ketë një peshë të rregulluar për mospërgjigjet, ku pesha fillestare të dizajnit në këtë rast do të jetë e barabartë me një, $W_d = 1$.

Faktori rregullues për mospërgjigjet është zakonisht i përcaktuar si raporti shumës të të gjithë peshave në kampionin origjinal me shumën e peshave të të gjithë njësive që nuk përgjigjen. Në një kampion të vetë-peshuar, kjo është ekuivalente me raportin e numrit të njësive të kampionit origjinal me numrin e njësive që përgjigjen dhe kjo paraqitet më poshtë.



8.1.3 Përdorimi i Informacioneve ndihmëse për të rregulluar peshat

Pesha e dizajnit, e shumëzuar me faktorin rregullimit të mospërgjegjjeve mund të përdoret për të prodhuar peshat përfundimtare dhe vlerësime në studim për disa karakteristika të dëshiruara. Megjithatë, ndonjëherë informacion në lidhje me popullsinë e anketuar është i disponueshëm nga burime të tjera, për shembull nga një informacion i një regjistrimi të mëparshëm. Ky informacion mund të përfshihet në procesin e peshimit. Ka dy arsye kryesore për përdorimin e të dhënave ndihmëse në vlerësimin.

Për shembull, shumë vrojtimit sociale rregullojnë vlerësimet e studimit të tyre në mënyrë që të jenë në përputhje me vlerësimet (mosha, seksi shpërndarjet, etj) e regjistrimit më të fundit të popullsisë. Informacionet ndihmëse gjithashtu mund të merren dhe nga të dhënat administrative ose nga vrojtimit të tjera të cilat janë të konsideruara të jenë më të besueshme për shkak të madhësisë të madhe të kampionit ose sepse vlerësimet e saj të publikuara duhet të respektohen.

Arsyeja e dytë për përdorimin e informacionit ndihmës është që të përmirësojë saktësinë e vlerësimeve. Kujtojmë që një vlerësues me vlerë të vogël të një variance të kampionit thuhet që do të jetë e saktë. Në kapitullin 6 diskutohet rreth rëndësisë që ka përdorimi i të dhënave ndihmëse në ndërtimin e një kampioni në mënyrë që të përmirësohet efikasiteti statistikore e strategjisë të kampionimit. Në fazën e dizajnit, prapë se prapë, informacioni ndihmës duhet të jetë i disponueshëm për të gjithë njësitë në listën e zgjedhjes (frame). Në vlerësime, të dhënat ndihmëse mund të përdoren për të rritur saktësinë e vlerësimeve deri sa variablat ndihmëse të jenë mbledhur për njësitë e anketës dhe se çfarë total popullsie ose vlerësime janë të mundshme për këto variabla ndihmëse nga një burim tjetër i besueshëm.

Për shembull nëse mosha e një personi nuk është e disponueshme një listë zgjedhje (frame), sjell që ky variabël nuk do të përdoret për të shtresëzuar popullsinë. Sidoqoftë, nëse ky informacion mbledhet nga një anketë, atëherë vlerësimet e anketës mund të rregullohen që të përputhen me shpërndarjen e moshës të Censusit. Nëse mosha është e korreluar me një variabël tjetër të mbledhur nga vrojtimit (për shembull, variablat shëndetësore), përdorimi i të dhënave ndihmëse mund të përmirësojë saktësinë e vlerësimeve. Kjo do të paraqitet më e gjerë në seksionet e mëposhtme.

Efikasiteti i përfituar në vlerësime nga përdorimi i të dhënave ndihmëse varet nga mënyra se si variablat e anketës janë të korreluara me të dhënat ndihmëse të mundshme. Saktësia e informacionit ndihmës është shumë e rëndësishme. Jo vetëm të dhënat duhet të jenë të besueshme, është e rëndësishme dhe burimi i të dhënave të jashtme të fokusohet tek e nejtja



REPUBLIKA E SHQIPËRISË
INSTITUTI I STATISTIKAVE

popullsi në vëzhgim dhe të bazohet në disa koncepte krahasueshmërie, përkufizime, periudhë reference, etj.

Post – stratifikimi përdoret për të rregulluar peshat e një ankete duke përdorur variabla të cilat janë të përshtatshme për shtresëzimin por të cilat nuk mund të përdoren në fazën e zgjedhjes të kampionit sepse të dhënat nuk janë të disponueshme, ose për më shumë përditësime, për një informacion të besueshëm shtresëzimi për popullsinë i cili bëhet i disponueshëm pas përzgjedhjes të kampionit. Post – shtresëzimi përdoret kur të dhënat ndihmëse janë në formën e numrave, për shembull numri i meshkujve dhe femrave në një popullatë. Është shumë efektive në zvogëlimin e variancës kur mesatarja e popullsisë të variablave të interesit janë aq të ndryshme përgjatë shtresëzimeve. Duhet vënë në dukje se kurdo që është e mundur, është e preferueshme që të shtresëzohet që në fazën e ndërtimit të kampionit se sa në atë të post – shtresëzohet.

Shembulli i mëposhtëm tregon se si të përdoret post – shtresëzimi për të përmirësuar vlerësimet e numrit të duhanpirësve në një kompani.

8.1.4 Vlerësimet e raporteve

Një metodë e përdorur shpesh në vrojtimit për një informacion ndihmës që përdoret për të përmirësuar vlerësimet e vrojtimit janë dhe *vlerësimet për raportin*. Kur të dhënat janë në formën e numrave, llogaritja e raportit korrespondon tek post – shtresëzimi. Në përgjithësi, me vlerësimin e raportit peshat e të dhënave në një grup të klasifikuar janë rregulluar nga një faktor multiplikativ. Ky faktor është raporti i vlerësimit nga të dhënat ndihmëse me vlerësimet për të njëjtë variabël, për grupin klasifikues. Për shembull, nëse objektivi i një studimi është të vlerësojë sipërfaqe në akra të grurit në rajonet e caktuara, atëherë njohja e akra të totale në secilin rajon mund të jetë një variabël i dobishëm ndihmës. Nëse akra të grurit janë të lidhur shumë me tokën totale brenda rajonit, ky informacion ndihmës mund të përmirësojë vlerësimet e sipërfaqes në akra të grurit. Përdorimi i vlerësimeve të raportit, faktori i rregullimit për secilin rajon do të jetë sipërfaqja totale e tokës në rajonin pjetuar me vlerësimet e kampionit të zonës së përgjithshme të tokës në rajon (dmth., faktori i rregullues do të jetë i formës $Y / \hat{Y}$).

8.1.5 Peshat e rregulluara më komplekse: kalibrimi dhe regresion i shumëfishtë

Vlerësimi i raportit është i bazuar në supozimin se ekziston një lidhje e thjeshtë dhe multiplikative mes karakteristikave të vrojtimit dhe një variabël ndihmës (p.sh., 2.7, si shumë burra do të thotë 2,7 si shumë duhanpirës meshkuj). Megjithatë, ky supozim nuk mund të mbajë të vërtetën në të gjitha situatat.



REPUBLIKA E SHQIPËRISË
INSTITUTI I STATISTIKAVE

Së pari, duke përdorur një ndryshore të vetme ndihmëse mund të jetë e pamjaftueshme për të siguruar vlerësime të mira.

Së dyti, marrëdhënia në mes të variablave që duhet të vlerësohen dhe variablave ndihmëse mund të jetë më e ndërlikuar se një multiplikative e thjeshtë. Në këtë rast, vlerësimi i regresionit mund të përdoret. Vlerësimi i regresionit është një qasje më komplekse që lejon analistët të marrin në konsideratë modelet më të sofistikuar, duke përfshirë modele me më shumë se një variabël ndihmës.

Një tjetër problem me vlerësimin e raportit është ajo se, mund të jetë e nevojshme për të siguruar se kampioni i ponderuar arrin një marrveshje me totalin e popullsisë të njohur për më shumë se një karakteristikë. Kur peshat e kampionit total arrijnë të bien dakord me totalin referues për më shumë se një karakteristikë, përdoret një metodë e quajtur kalibrimit. Kjo situatë lind, për shembull, kur post-stratat janë formuar duke përdorur më shumë se një variabël dhe vetëm popullsia marxhinale për secilin variabël janë të njohur. Për shembull, në qoftë se popullsia do të ishte në dispozicion sipas grup-moshave dhe seksit, por numri i burrave dhe i grave brenda secilit grup moshe do të ishin të panjohur, në të cilin rast, metoda e post – stratifikimit përshkruhet më parë nuk mund të aplikohet duke përdorur të dy karakteristikat. (Vlerësimi i raportit dhe post-shtresëzimi janë duke përdorur një ndryshore kalibrimit.)

Një zgjatje e metodës së raportit, të referuara si rang i vlerësimit të raportit lejon agjensinë statistikore në llogaritjen e peshave të rregulluara në mënyrë që vlerësimet janë sa më shumë afër numrave mergjinale të referencës për secilën karakteristikë. Një përshkrim i kësaj metode mund të gjenden në Deming dhe Stephan (1940), dhe Arora Brackstone (1977), Bankier (1978), Brackstone dhe Rao (1978), Binder (1988).

Procedurat e vlerësimit të përgjithësuar janë zhvilluar edhe për të prodhuar vlerësime që bien dakord me referencat në total për shumë karakteristika në të njëjtën kohë. Në këto teknika të përgjithshme, procesi për të siguruar që totalin referues janë të lidhura, i është referuar kalibrimit dhe rregullimeve të peshës të prodhuar quhen faktorë kalibrimit.

Ashtu si me post-stratifikimin dhe raportin e vlerësimit, këto teknika rezultojnë në rregullimet për dizinimin fillestar të peshave. Peshat përfundimtare të përdorura për të llogaritur vlerësime bëhen produkt i peshave fillestare të rregulluara për mospërgjigjet dhe faktorët e kalibrimit. Procedurat e vlerësimit janë përgjithësuar përtej fushëveprimit të këtij dokument. Lexuesi do të gjejë teorinë e vlerësimit të regresionit të përgjithësuar në Särndal, Swensson dhe Wretman (1992), Deville dhe Särndal (1992), dhe Hidiroglou dhe Särndal (1998). Shih gjithashtu Kapitullin 11 – Analiza e të dhënave të Anketës për një hyrje në regresionin linear.



8.2 Prodhimi e vlerësimeve të thjeshta (Totalët, mesataret, përqindje)

Shembulli do të ilustrojë se si mund të llogarisim peshën dhe se si të bëjmë rregullimet për mospërgjigjet dhe të dhënat ndihmëse. Ky seksion shpjegon se si të gjenerojmë vlerësime duke përdorur vlerësuesin final të peshave.

Ne e kemi parë këtë çështje tek – Dezinjimi i kampionit probabilitar, shpërndarja e kampionit e një vlerësuesi mund të përcaktohet. Me kampionin pa probabilitet, derisa probabiliteti i kampioneve të ndryshme të përzgjedhura janë të panjohura, shpërndarja e kampionit nuk mund të llogaritet. Aftësia për të matur gabimet e kampionit është një pjesë e rëndësishme e vlerësimit të një ankete dhe gjithashtu një nga arsyet kryesore për të përformuar kampion me probabilitet.

8.2.1 Vlerësues për tipe të dhënash të ndryshme.

Statistikat përshkruese të thjeshta të tilla si totalët, mesataret dhe përqindjet janë prodhuar për pothuajse çdo anketë. Zakonisht, statistikat dhe analizat më të ndërlikuara janë gjithashtu të kërkuara. Në shumicën e studimeve, të dhënat janë mbledhur për një gamë më të gjerë të variablave të cilat mund të jenë cilësore (referuar edhe si kategorike) ose sasiore. Për disa variablave kualitative, të tilla si gjinia ose statusi martesor, vetëm pak kategori janë të mundshme. Për pyetjet opsionale, përgjigjet e të anketuarve janë mbledhur shpesh duke përdorur një shkallë marrëveshjeje të tilla si plotësisht dakord, dakord, as dakord as kundër, pajtohem, pajtohem plotësisht. Vini re se me të dhënave kategorike, çdo njësi përshtatet në një dhe vetëm me një kategori.

Nëse njësia e matjes tregon sasi të tilla si metra apo vite, të dhënat janë quajtur të dhëna sasiore. Të dhënat sasiore janë zakonisht përgjigjet në pyetjet e tipit sa? apo c'sasi? Disa shembuj janë mosha, numri i fëmijëve, numri i orëve të punës, shpenzimet dhe të ardhurat, dhe presion i gjakut. Llojet e ndryshme të vlerësuesve janë të përshtatshme për këto lloje të ndryshme të variablave. Përqindjet dhe totale janë prodhuar në mënyrë tipike për variablat cilësorë, ndërsa mesataret dhe totalët e vlerësuarara për variablat kuantitative. Në këtë seksion, procedurat për prodhimin e vlerësimeve do të paraqitet veçmas për të dhënat cilësore dhe sasiore.

Përveç llojin e të dhënave, një tjetër konsideratë gjatë vlerësimit është dhe natyra e popullsisë për të cilën janë për të bërë vlerësimet. Vlerësimet mund të prodhohen për popullsinë e anketës në tërësi ose për të nëngrupe te veçanta apo fusha të popullsisë (p.sh., krahinat). Kur klasifikimi origjinal i njësive të kampionit ka ndryshuar midis kohës së përzgjedhjes së kampionit dhe vlerësimit, klasifikim i ri duhet të përdoret për vlerësimin e fushave. Përgjigjet në pyetjet e mëposhtme do t'ju ndihmojnë të përcaktojë se sa janë llogaritur vlerësimet e anketës:

- Çfarë lloj statistike është e nevojshme? Një total, një mesatare, një përqindje?



REPUBLIKA E SHQIPËRISË
INSTITUTI I STATISTIKAVE

- Cfare lloj te dhenash po perdoret? Cilësore apo sasiore?
- Cilat janë peshat finale?
- Cilat janë fushat e interesit?

Proceduar për të vlerësuar totalet, mesataret dhe përqindjet për të gjithë popullatën në vërtetim dhe fushat që përdorin pesha janë të përshkruara më poshtë për të dhënat sasiore dhe cilësore. Vlerësimet e mëposhtme mund të aplikohen në shumë kampioni probabilitare – (SRS, SYS) ose më komplekse.

i. Vlerësimet e totaleve të popullsisë

Për të dhëna cilësore dhe sasiore, vlerësimet për numrin total të njësive në një anketë të popullsisë llogaritet duke shtuar peshat finale të rregulluara dhe njësitë që përgjigjen :

$$\hat{N} = \sum_{i \in S_r} w_i$$

Ku i është njësia e i -të që përgjigjet në kampoin dhe w_i është pesha finale e rregulluar dhe S_r është grupi i njësive që përgjigjen.

Për të dhënat sasiore, vlerësimi i vlerave totale (si për shembull shpenzimet totale) është produkti i peshave finale w_i dhe vlerës, y_i , për secilën njësi që përgjigjet kjo e gjitha si shumatore e të gjithë njësive që përgjigjen:

$$\hat{Y} = \sum_{i \in S_r} w_i y_i$$

ii. Vlerësimi i mesatareve të popullsisë

Për të dhënat sasiore, vlerësimi i një vlerë mesatare e popullatës është marrë duke shtuar se bashku produktit të vlerës së kampionit dhe peshat për çdo njësi përgjigjeje, dhe duke pjesëtuar këtë shumë me shumën e peshave. Me fjalë të tjera, vlerësimi i mesatares në popullsinë është vlerësim i përgjithshëm i vlera për të dhënat sasiore pjesëtuar me vlerësimin e numrit të përgjithshëm të njësive të popullsisë.

$$\begin{aligned}\hat{\bar{Y}} &= \frac{\sum_{i \in S_r} w_i y_i}{\sum_{i \in S_r} w_i} \\ &= \frac{\hat{Y}}{\hat{N}}\end{aligned}$$



REPUBLIKA E SHQIPËRISË
INSTITUTI I STATISTIKAVE

Vini re se për SRS dhe SYS me shkallë përgjigjeje 100% dhe pa kryer rregullim peshe, vlerësuesi reduktohet në formulën e mëposhtme:

$$\hat{\bar{Y}} = \frac{\sum_{i \in S_r} y_i}{n}$$

iii. Vlerësimi i përqindjeve të popullsisë.

Për të dhëna cilësore, vlerësimi i përqindjen e njësive në popullsinë e anketës që ka një të dhënë karakteristike është marrë duke shtuar sebashku peshat për njësitë që kanë atë karakteristike, dhe duke pjesëtuar këtë total me shumën e peshave për të gjithë të anketuarit. Me fjalë të tjera, vlerësimi i përqindjeve të popullsisë është llogaritja i numrit total të njësive që posedojnë karakteristiken e dhënë pjesëtuar me vlerësimi i numrit të përgjithshëm të njësive të popullsisë.

$$\begin{aligned}\hat{P} &= \frac{\sum_{i \in S_r \cap C} w_i}{\sum_{i \in S_r} w_i} \\ &= \frac{\hat{N}_C}{\hat{N}}\end{aligned}$$

Ku C- janë njësitë e mbledhura për n karakteristikë të dhënë.

iv. Vlerësimet e grupimeve të popullsisë

Vlerësime mund të kërkohen për fusha si mosha, llojet e banesave, madhësitë e personave, ose llojin e të ardhurave.

- madhësia e popullsisë për fushën e interesit për të dy llojet e të dhënave, sasiore dhe cilësore është e vlerësuar si më poshtë:

$$\hat{N}_{domain} = \sum_{i \in S_r \cap domain} w_i$$

- vlerësimi i një fushe për totalet e të dhënave sasiore është:

$$\hat{Y}_{domain} = \sum_{i \in S_r \cap domain} w_i y_i$$

- vlerësimi i një fushe për mesataret e të dhënave sasiore është :



REPUBLIKA E SHQIPËRISË
INSTITUTI I STATISTIKAVE

$$\begin{aligned}\hat{\bar{Y}}_{domain} &= \frac{\sum_{i \in S_r \cap domain} w_i y_i}{\sum_{i \in S_r \cap domain} w_i} \\ &= \frac{\hat{Y}_{domain}}{\hat{N}_{domain}}\end{aligned}$$

- vlerësimi i një fushe për përqindjet për të dhënat sasiore dhe cilësore është :

$$\begin{aligned}\hat{P}_{domain} &= \frac{\sum_{i \in S_r \cap domain \cap C} w_i}{\sum_{i \in S_r \cap domain} w_i} \\ &= \frac{\hat{N}_{domain \cap C}}{\hat{N}_{domain}}\end{aligned}$$

8.3 Çështjet e vlerësimit në grupe të vogla (domains)

Kur është e mundur, dizenjimi i kampionit duhet të marrë në konsideratë fushat e interesit përmes shtresimit. Kjo do të sigurojë madhësinë adekuate të kampionit dhe saktësi. Megjithatë, kufizimet mbi madhësinë e kampionit dhe kritereve të tjera të dizajnit (të tilla si informacion i kornizës së vrojtimit) mund të thotë se vetëm një numër modest i shtresave mund të formohet, kështu që për disa fusha, veçanërisht ato të vogla, madhësia e kampionit mund të jetë i pamjaftueshëm.

Madhësia e papërshtatshme e kampionit në një fushë mund të përbëjë një problem në vlerësim. Teknika të ndryshme janë në dispozicion për të gjeneruar vlerësime në raste të tilla. Këto përfshijnë vlerësimin sintetik, vlerësimin e përbërë, etj,... Këto metoda zakonisht kërkojnë informacione lidhur me to nga një burim tjetër, ose përdorimin e modeleve të mira. Aplikimi i këtyre teknikave mund të bëhen komplekse dhe është përtej fushëveprimit të këtij manuali. Më shumë informacion mbi këtë temë të avancuar mund të gjendet në Särndal, Swensson dhe Wretman (1992), Ghosh dhe Rao (1994), Singh, Gambino dhe Mantel (1994).



8.3.1 Outliers (Njësi të veçuara)

Barnett dhe Lewis (1995) e përcaktojnë outlier-in si një vëzhgim apo nëngrup i vëzhgimeve që duket të jetë në kundërshtim me pjesën e mbetur të dataset. Ka një shumëllojshmëri të metodave në dispozicion për të reduktuar ndikimin e outlierave në vlerësimet e anketës. Thjesht duke i injoruar ata, mund të rezultojë në një humbje në përpikëri; duke u dhënë atyre një peshë të barabartë me një ose zero mund të paragjykojmë rezultatin (i njëanshëm). Përndryshe, informacione ndihmëse dhe post-stratifikimi mund të përdoret për të siguruar që outlierat nuk kontribuojnë një sasi të madhe të paarsyeshme të vlerësimeve. Trajtimi i outlierave do të mbulohen më tej në Kapitullin 10 - Përpunimi. Më shumë informacion mbi këtë temë të avancuara mund të gjenden në Kish (1965), Hidioglou dhe Srinath (1981).

8.3.2 Llogaritja e gabimit të kampionimit të Vlerësimeve të Anketës

Vlerësimet e prodhuar nga një studim janë subjekt të gabimit. Në hyrje në Dizenjimin e një Ankete, u deklarua se ka dy lloje themelore të gabimit – gabime kampionimi dhe gabime jo kampionimi. Gabime të jo kampionimit shpesh shkaktojnë paragjykim (njëanshmëri) dhe janë të vështira për t'u matur; gabime të kampionimit shkaktojnë ndryshueshmëri - ato masin shkallën në të cilën një vlerësim nga mostrat e mundshme të ndryshme, të së njëjtës madhësi dhe dizajni, duke përdorur të njëjtën vlerësues, ndryshojnë nga njëri-tjetri.

Kapitulli – Dizenjimi i Kampionit shpjegoi rëndësinë e të paturit një vlerësim të variancës së kampionit në fazën e dizenjimit, në mënyrë për të krahasuar efikasitetin statistikor të dizenjimit të kampioneve të ndryshme.

Kapitulli - Përcaktimi i Madhësisë dhe Alokimi të Kampionit tregon se si një vlerësim i variancës së kampionit është përdorur për të të përcaktuar madhësinë e kampionit së kërkuar që të plotësojë një nivel të caktuar të saktësisë.

Qëllimi i këtij kapitulli është për të ilustruar se si matet varianca e kampioneve dhe rëndësinë e inkorporimit të saktë të dizajnit të kampionit. Ky kapitull paraqet vetëm vlerësuesit e kampionit për një mesatare ose totale të vlerësuar, për një SRS ose SRS të shtresëzuar, duke supozuar se nuk ka ndryshime në peshat origjinale të dizenjimit.

Vlerësimi i variancës për një pjesë të vlerësuar nga një SRS dhe dizajn më kompleks (duke përdorur dizajn efektin) janë paraqitur në Kapitullin - Përcaktimi i Madhësisë dhe Alokimi të Kampionit.



REPUBLIKA E SHQIPËRISË
INSTITUTI I STATISTIKAVE

Çdo strategji kampionimi ka formulat e saj, të veçanta, për vlerësimin e variancës së kampioneve, të komplikuar për kohën. Në mënyrë që të vlerësojmë saktë variancën e kampioneve për një anketë me të dhënat komplekse (p.sh., për një dizajn më kompleks dhe / ose në prani të rregullimeve të peshave), duhet të konsultoheni me një statisticien vrojtimesh, i familjarizuar me këtë lloj problemi. Formulatat e paraqitura në këtë kapitull mund të gjenden në çdo libër hyrës në teorinë e kampionimit. për shembull, Cochran, 1977 ose Lohr, 1999.

8.4 Llogaritja e variancës.

Matematikisht, varianca e kampionit së një vlerësimi është devijimi mesatar në katror në lidhje me vlerën mesatare të vlerësuesit, në të gjitha mostrat e mundshme. Tek seksioni hyrës në Dizajnin e Kampionit ka të listuar faktorët që ndikojnë në madhësinë variancës së kampionit:

- Ndryshueshmëria e karakteristikës së interesuar në popullatë;
- Madhësia e popullsisë;
- Shkalla e përgjigjes;
- Dizenjimi i kampionit dhe metoda e vlerësimit.

Efekti i këtyre faktorëve në variancën e kampionit është ilustruar më poshtë duke përdorur vlerësues për variancën e kampionit, të mesatares së vlerësuar të popullsisë nga një zgjedhje SRSWOR, me shkallë përgjigjeje 100%.

Vlera e çdo variabli, y_i , zakonisht varion nga një njësi në tjetrën brenda popullatës. Varianca e popullatës, σ^2 , e të gjitha njësive, y_i , në popullatë është përcaktuar si vijon:

$$\sigma^2 = \frac{(N-1)}{N} S^2$$

Ku,

$$S^2 = \sum_{i \in U} \frac{(y_i - \bar{Y})^2}{N-1}$$

dhe U është bashkësia e njësive në popullatë dhe ka N njësi në popullatë.

Një Vlerësues i paanshëm i mesatares së popullatës për një zgjedhje SRSWOR të madhësisë n me shkallë përgjigjeje 100% është:



REPUBLIKA E SHQIPËRISË
INSTITUTI I STATISTIKAVE

$$\hat{\bar{Y}} = \sum_{i \in S_r} \frac{y_i}{n}$$

ku S_r është bashkësia e të përgjigjurave në kampion dhe gjenden n njësi në kampion.

Vlerësimi $\hat{\bar{Y}}$, ndryshon nga kampioni në kampion.

Një shprehje për variancën e kampionit të $\hat{\bar{Y}}$, për një zgjedhje SRSWOR të madhësisë n është:

$$Var(\hat{\bar{Y}}) = (1 - \frac{n}{N}) \frac{S^2}{n}$$

Është e qartë nga ekuacioni i mësipërm se një vlerësim për një karakteristikë - me dallime të mëdha nga njëra njësi tek tjetra - me ndryshueshmëri të lartë në popullatë, ka variancë më të lartë kampionimi sesa një vlerësim për një karakteristikë, ku ndryshueshmëria e popullatës është e ulët. Zakonisht, S^2 është i panjohur dhe duhet të vlerësohet.

Gjithashtu është e qartë se madhësia e popullsisë ka një ndikim në variancën e kampionit: termi $f = n / N$ i referohet pjesës që ze kampionit në popullatë dhe termi $(1-f) = 1-n / N$ si faktor korrektues në popullsi të fundme (fpc).

Me rritjen e madhësisë së kampionit, n , varianca e kampionimit ulët, në rastin e censusit (ku $n=N$) nuk ka variancë kampionimi. Kur fraksioni i kampionit në popullatë është i vogël (psh. madhësia e kampionit është e vogël krahasuar me popullsinë) fpc mund të injorohet. (Cochran, 1977, thekson se ky faktor mund të injorohet kur ai nuk e kalon 5%, dhe për qëllime të mëtejshme edhe në qoftë se është i lartë deri në 10%.) Vini re gjithashtu se mospërgjigjet rrisin variancën e kampionit, duke zvogëluar madhësinë e kampionit në mënyrë efektive.

Siç u shpjegua në seksionin - Dizajnimi i Kampionit, disa strategji të kampionimit janë më efektive se disa të tjera. Për shembull, stratifikimi (ose shtresimi) dhe përdorimi i vlerësuesve të raportit (ratio estimator) mund të përmirësojë saktësinë e një vlerësimi.

Në diskutimin e mëposhtëm të variancës së kampionit, supozohet se vlerësuesi është i paanshëm. Megjithatë, ka disa raste kur një vlerësues i njëanshëm (paragjykues) është i preferuar (për shembull, kur saktësi e tij është shumë më mirë se çdo preventivues vlerësues tjetër i mundshëm,



REPUBLIKA E SHQIPËRISË
INSTITUTI I STATISTIKAVE

i paanshmëm). Nëse paragjykimi (njëanshmëria) është i pranishëm për çfarëdo arsye - për shkak të një vlerësuesi të njëanshëm ose që rrjedh nga një gabim jo kampionimi – atëherë formulat e variancës së kampioneve të paraqitura në seksionet e mëposhtme llogarisin gabimin e mesatares në katror (MSE), i cili është një matës i variancës së kampionit dhe bias (njëanshmërisë). Kjo mund të çojë në probleme me intervalet e besimit, shih seksionin - Analiza e të Dhënave të Vrojtimit, për më shumë detaje.

Llogaritja e Variancës së vërtetë të kampionimit

Shembulli i mëposhtëm ilustron se si llogaritet Varianca e vërtetë e kampionimit e Mesatares së vlerësuar të shpenzimeve për video për një zgjedhje SRSWOR, e madhësisë $n=2$.

Shembull

Shpenzimet për produktet e videove për një popullsi prej 4 personash është dhënë më poshtë. Për një zgjedhje SRSWOR të madhësisë $n = 2$ cila është varianca e vërtetë e kampionimit e mesatares së estimuar të shpenzimeve?

Table 8: Video Expenditure per Household	
Household	Video Expenditure (\$)
1	10
2	20
3	30
4	40

Fillimisht, vini re se vlera e parametrin *mesatarja e shpenzimeve për Video e popullsisë* është:

$$\bar{y} = \sum_{i \in S} \frac{y_i}{n}$$

$$Var(\hat{\bar{Y}}) = (1 - \frac{n}{N}) \frac{S^2}{n}$$

$$Var(\hat{Y}) = Var(N \times \hat{\bar{Y}}) = N^2 Var(\hat{\bar{Y}}) = N^2 (1 - \frac{n}{N}) \frac{\hat{S}^2}{n}$$

Matjet e tjera të gabimit



REPUBLIKA E SHQIPËRISË
INSTITUTI I STATISTIKAVE

Para se të ilustrojmë se si të vlerësojmë variancën e kampioneve, matje të tjera të zakonshme të gabimit në kampionim do të jenë paraqitur, të tilla si:

- Gabimi standard;
- Koeficienti i variacionit;
- Marzhi i gabimit;
- Intervali i besueshmërisë.

Të gjitha këto janë shprehje të lidhura, është e mundur për të shkuar nga njëra tek tjetra duke përdorur operatione të thjeshta matematikore.

Gabimi Standart dhe Koeficienti i Variacionit

Gabimi standard i një vlerësuesi është rrënja katrore e variancës të tij të kampionimit. Kjo matje është më e lehtë për t'u interpretuar pasi ajo siguron një tregues të gabimit të kampionimit duke përdorur të njëjtën shkallë si vlerësim, ndërsa varianca është e bazuar në dallimet në katror.

Megjithatë, edhe gabimi standart mund të jetë i vështirë për t'u interpretuar në aspektin e "Sa është madhësia që një gabim standart të jetë i pranueshëm?" Çfarë është e madhe, varet nga madhësia e vlerësimit. Për shembull, një gabim standart i barabartë me 100 do të konsiderohet i madh për matjen e peshës mesatare të njerëzve, por nuk do të konsiderohet i madh për vlerësimin e të ardhurave mesatare vjetore.

Kjo është më e dobishme në shumë situata për të vlerësuar madhësinë e gabimit standart në lidhje me matjen e vlerësimit të karakteristikës. Koeficienti i variacionit ofron një masë të tillë. Ai është raporti i gabimit standart të vlerësimit të anketës me vlerën mesatare të vet vlerësimit, në të gjitha mostrat e mundshme. Koeficienti i variacionit llogaritet zakonisht si vlerësimin i gabimit standart të vlerësimit të anketës, për vet vlerësimin. Kjo masë relative e gabimit të kampionimit është shprehur zakonisht si një përqindje (10% në vend të 0.1). Kjo është shumë e dobishme në krahasimin e saktësisë e vlerësimeve të kampionit, ku madhësia e tyre ose shkalla ndryshon nga njëri-tjetri. Megjithatë, kjo nuk është aq e dobishme për vlerësuesit e karakteristikave, vlerat e vërteta të të cilëve mund të jetë zero ose negative, duke përfshirë vlerësimet e ndryshimeve (p.sh., ndryshim në vlerë mesatare të të ardhurat nga viti i kaluar).

Marzhi i gabimit dhe Intervalit të besueshmërisë

Është e zakonshme të shohësh rezultatet e një vrojtimit të raportuar në një gazetë si më poshtë: Sipas një studimi të fundit, 15% e banorëve të Otavës ndjekin shërbime fetare çdo javë. Rezultatet, bazuar në një kampion prej 1.345 banorë, konsiderohen të sakta brenda plus ose minus tre pikë përqindje 19 herë nga 20.



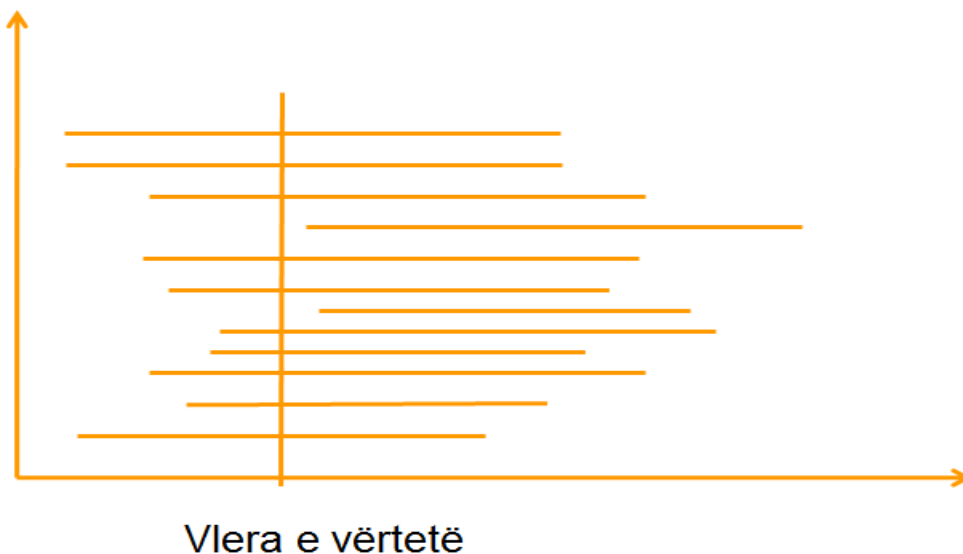
REPUBLIKA E SHQIPËRISË
INSTITUTI I STATISTIKAVE

Çfarë do të thotë kjo deklaratë? Kjo tregon se vlera e vërtetë, përqindja e vërtetë e njerëzve që ndjekin shërbime fetare çdo javë, ka të ngjarë të jetë brenda tre pikëve përqindje të vlerësimit (15%). Në shembullin e mësipërm, marzhi i gabimit është plus ose minus tre pikë përqindje ose thjesht tre pikë përqindje ndërsa intervali i besimit është diapazoni nga 12% deri në 18%. Kufijtë e gabimit përfshijnë gjithmonë një deklaratë besimi, nivelin e sigurisë që ndjehet në lidhje me këtë interval. Në këtë shembull, deklarata e besimit është 19 herë nga 20. Kjo do të thotë që, nëse vërtetimi përsëritej shumë herë, atëherë 19 herë nga 20 (ose 95% të kohës), intervali i besimit do të mbulonte vlerën e vërtetë të popullsisë.

Teoria që qëndron në ndërtimin e intervalit të besimit mund të përshkruhet si më poshtë. Supozoni se për një kampion të madh ekziston një vlerësim i mesatares së popullsisë, $\hat{Y}$, dhe një vlerësim i gabimit të tij standard, $SE(\hat{Y})$. Më pas, për shkak të Teoremës Qendrore të Limitit dhe Shpërndarjes Normale (figura mëposhtë), shanset janë:

- .10 që gabimi absolut $|\hat{Y} - Y|$ tejkalon $1.65 \times SE(\hat{Y})$ (kjo korrespondon me një interval besimi 90%);
- .05 që gabimi absolut $|\hat{Y} - Y|$ tejkalon $1.96 \times SE(\hat{Y})$ (kjo korrespondon me një interval besimi 95%);
- .01 që gabimi absolut $|\hat{Y} - Y|$ tejkalon $2.58 \times SE(\hat{Y})$ (kjo korrespondon me një interval besimi 99%).

Interval i Besimit





REPUBLIKA E SHQIPËRISË
INSTITUTI I STATISTIKAVE

Figura 3: Marzhi i gabimit dhe intervali i besueshmërisë

Këto formula aplikohen për të gjithë vlerësuesit e shpërndarë normalisht. Duhet të theksohet se gabimet standarde janë të dobishme jo vetëm për llogaritjen e intervalit të besimit, por edhe për analizën e të dhënave inferenciale, siç është testimi i hipotezës përgjatë analizës së të dhënave të vrojtimit.

Efekti i Dizenjimit

Efekti i dizenjimit krahason variancën e vlerësuesve nga një kampion i dizenjuar me ata nga një zgjedhje SRS. Është raporti i variancës së kampionimit të një vlerësuesi nën një dizajn të caktuar me variancën e kampionimit, të një vlerësuesi nën një zgjedhje SRS të së njëjtës madhësi.

Siç është përmendur në Dizenjimi i kampionit, ky matës është përdorur shpesh për të krahasuar efikasitetin e vlerësuesve nga dizajne të ndryshme. Nëse raporti është më pak se një kjo tregon se dizajni i kampionit është më efikas se zgjedhja SRS, më i madh se një tregon se dizajni i kampionit është më pak efikas se zgjedhja SRS.

Për shembullin e paraqitur më herët,

$$deff = \frac{Var_{STR}(\hat{Y})}{Var_{SRS}(\hat{Y})} = \frac{44,000}{112,200} = 0.39$$

tregohet se shtresëzimi përmirësoi në masë të madhe saktësinë e mesatares së vlerësuar të anketës.

Efektet e dizenjimit mund të përdoret gjithashtu për të siguruar vlerësime të përafërt të variancës për dizenjime komplekse. Nëse një vlerësim i efektit të dizenjimit është në dispozicion nga një hulumtim i mëparshëm, që përdorte dizenjim kampioni të njëjtën, ajo mund të përdoret për të llogaritur madhësinë e kërkuar të kampionit të një studimi (kjo do të diskutohet tek tema Përcaktimi i Madhësisë dhe Alokimi të Kampionit). Për më shumë informacion mbi efektet e dizenjimit, shih Kish (1965).

Llogaritja e Variancës së Kampionimit duke përdorur Kampionimin me përsëritje

Variancat e statistikave të thjeshta të tilla si mesataret dhe totalet mund të vlerësohen duke përdorur procedurat e deklaruara më parë për dizajnin e thjeshtë. Megjithatë, për dizenjime kampioni më komplekse apo vlerësues kompleks (p.sh., funksionet e totaleve) përcaktimi i formulës së variancës së kampionimit mund të jetë i vështirë.

Ekzistojnë metoda alternative (përveç aplikimit të efektit të dizenjimit) ku varianca e kampionimit mund të vlerësohet për një grup të përzgjedhur të kampionit të dizenjuar dhe



REPUBLIKA E SHQIPËRISË
INSTITUTI I STATISTIKAVE

procedurave të vlerësimit. Një metodë e tillë është kampionimi i përsëritur. Në marrjen e kampioneve që përsëriten, në vend të zgjedhjes të një kampioni me madhësi n , k mostrat të pavarura të madhësisë n/k janë zgjedhur. Për secilin prej këtyre kampioneve k , është prodhuar një vlerësim i karakteristikës së interesit. Ndryshueshmëria midis vlerësimeve të kampionit k është përdorur pastaj për të vlerësuar variancën e kampionimit. Vlerësimi, t , i karakteristikës së interesit është dhënë me mesataren e vlerësimeve të prodhuara për çdo kampion:

$$t = \sum_{j \in K} \frac{t_j}{k}$$

ku K është bashkësia e kampioneve, k është numri i kampioneve dhe t_j është vlerësim për mostrën j^{th} .

Varianca e vlerësuar e kampionimit t , $\hat{Var}(t)$, jepet nga shprehja:

$$\hat{Var}(t) = \sum_{j \in K} \frac{(t_j - t)^2}{k(k-1)}$$

Vlerësimi i variancës së kampionimit për $\hat{\bar{Y}}$, duke përdorur Kampionimin me Përsëritje, SRS. Në Shembullin që vijon dy mostra të madhësisë $n = 5$ janë zgjedhur. Rezultatet janë përshkruar në tabelën e mëposhtme:

Table 1: Calculationg the estimated sampling variance of $\bar{Y}$ using Repliated Sampling

Sample 1	Age of Sample Units	Sample 2	Age of sample Units
1	21	1	26
2	27	2	32
3	34	3	37
4	38	4	40
5	42	5	47
Average Age	32.4		36.4

1. Mesatarja e vlerësuar e Moshës së popullatës është:

$$\begin{aligned} \hat{\bar{Y}} &= \sum_{j \in K} \frac{\bar{Y}_j}{k} \\ &= \frac{32.4 + 36.4}{2} \\ &= 34.4 \end{aligned}$$



REPUBLIKA E SHQIPËRISË
INSTITUTI I STATISTIKAVE

2. Varianca e vlerësuar e kampionimit të Mesatares së moshës, e dhënë nga metoda e kampionimit me përsëritje është:

$$\begin{aligned}\hat{Var}(\hat{\bar{Y}}) &= \sum_{j \in K} \frac{(\hat{Y}_j - \hat{\bar{Y}})^2}{k(k-1)} \\ &= \frac{(32.4 - 34.4)^2 + (36.4 - 34.4)^2}{2} \\ &= 4\end{aligned}$$

3. Gabimi standard i vlerësuar, $\hat{SE}(\hat{\bar{Y}})$, është 2

Atëherë si përfundim mund të theksojmë:

Lista e mëposhtme jep një përmbledhje të pikave të rëndësishme për t'u marrë parasysh në vlerësimin e të dhënave të anketës:

- Vlerësimi duhet të marrë parasysh dizajnin e kampionit. Për këtë qëllim, pesha e dizajnit duhet të jetë e përfshirë në procesin e vlerësimit.
- Peshat fillestare të dizajnit duhet të rregullohen për shkak të mospërgjigjeve.
- Informacioni ndihmës, nëse ka cilësinë e duhur dhe lidhet me variablat kryesore të anketës, duhet të përdoret, kur është e mundur, për të përmirësuar qëndrueshmëri dhe saktësi të vlerësimeve.
- Dizajni i kampionit dhe alokimi i kampionit duhet të përdoret për të përmbushur kërkesat e fushave të interesit. Nëse kjo nuk është e mundur në fazën e dizajnit, metodat të veçanta vlerësimi duhet të konsiderohen në fazën e vlerësimit.
- Outlier-at mund të çojnë në ndryshueshmëri të madhe të kampionimit në vlerësime. Zbulimi i outlier-ave dhe trajtimi duhet të konsiderohen si pjese e rëndësishme e përmirësimit të vlerësimit.
- Vlerësimet e anketës duhet të përfshijnë një vlerësim të gabimit të kampionimit të tyre, në formën e variancës së kampionimit, gabimit standard, koeficientit të variacionit, marzhit të gabimit ose intervalit të besueshmërisë.

9. Organizimi i vrojtimeve me zgjedhje

9.1 Procedurat e mbledhjes së të dhënave

Zyra qëndrore



REPUBLIKA E SHQIPËRISË
INSTITUTI I STATISTIKAVE

Zyra qendrore zakonisht harton dhe zhvillon manuale për intervistuesit, mbikëqyrësit dhe personelin tjetër për mbledhjen e të dhënave. Këto procedura përfshijnë: intervistimin, renditjen e njësive të marra në kampion dhe azhurnimin e kornizës së vrojtimit (frame-t), gjurmimin e të intervistuarve, përcjelljen e të pa-përgjigjurve dhe redaktimin dhe kodimin që do të kryhen në terren.

Procedurat e trajnimit

Kjo përfshin zhvillimin e ushtrimeve, skriptet e intervistave të stimuluar, dhe materialet audio-vizive, si dhe udhëzuesit e trajnimit.

Orari i mbledhjes së të dhënave

Është krijuar një orar për mbledhjen e të dhënave, ku detajohen piketat e studimit dhe periudhat e raportimit në mënyrë që data e synuar të arrihet. Programi përfshin normat e zgjidhura të synuara për secilën periudhë raportuese si dhe nivelet e përgjigjes së synuar.

Sistemet e raportimit dhe kontrollit

Për të siguruar që të gjitha operacionet e mbledhjes së të dhënave janë nën kontroll, janë zhvilluar format e kontrollit për të gjithë kampionin dhe nga intervistuesi (i quajtur detyra e intervistuesit) dhe janë krijuar dhe janë ngritur procedura të rregullta raportimi për të azhurnuar statusin e secilës njësi të vrojtuar.

Zyra Rajonale

Përgjegjësitë kryesore të zyrave rajonale janë mbledhja e të dhënave, kapja e të dhënave dhe marrëdhëniet me publikun. Kapja e të dhënave shpesh bëhet nga zyrat rajonale pasi është më e lehtë për të kontrolluar, menaxhuar dhe për të kryer transmetimin në zyrat qendrore të dosjeve elektronike sesa të pyetësorëve në letër.

Mbikëqyrësi

Në varësi të madhësisë së vrojtimit, mund të ketë më shumë se një mbikëqyrës, secili përgjegjës për një ekip intervistuesish. Raporti i mbikëqyrësve ndaj intervistuesve ndryshon në varësi të:

- kompleksitetit të vrojtimit;
- përvoja e mbikëqyrësve dhe stafit intervistues;
- vendndodhja e kampionit;
- vështirësi në udhëtime;
- lehtësia e përdorimit të sistemit të informacionit të menaxhimit.



REPUBLIKA E SHQIPËRISË
INSTITUTI I STATISTIKAVE

Në bashkëpunim me menaxherin rajonal, mbikëqyrësi punëson dhe trajnon intervistuesit dhe blen dhe shpërndan pajisje dhe pajisje. Mbikëqyrësi është përgjegjës për administrimin ditor të mbledhjes së të dhënave, përfshirë mbikëqyrjen e intervistuesve dhe monitorimin e afërt të progresit dhe cilësisë. Një tjetër detyrë e rëndësishme e mbikëqyrësit është përcjellja e refuzimeve

Intervistuesi

Intervistuesi bën mbledhjen e të dhënave aktuale dhe raporton progresin dhe problemet rregullisht te mbikëqyrësi.

9.2 Përgatitja e procedurave të mbledhjes së të dhënave

Para se të fillojë mbledhja e të dhënave duhet të kryhen shumë detyra, për shembull:

1. Shkrimi i Manualeve. Çelësi për mbledhjen e suksesshme të të dhënave janë intervistuesit cilësor dhe personeli tjetër. Intervistuesit e mirë dhe personeli i anketës duhet të kenë aftësi dhe cilësi personale për të qenë efektiv. Konsistenca dhe cilësia e punës së tyre përcakton cilësinë e rezultateve të anketës. Manualët përcaktojnë procedurat standarde, si dhe udhëzimet për trajtimin e problemeve të parashikueshme.
2. Stafi i punësimit dhe trajnimit. Ata që janë të punësuar kanë aftësitë dhe cilësitë personale të kërkuara dhe që ata kanë trajnimin dhe mjetet e duhura. Instituti i statistikave duhet të mbajë një listë të intervistuesve të trajnuar në mënyrë që punësimi të mund të bëhet duke përdorur këtë listë. Nëse nevojat e vërtetimit janë të mëdha, mund të jetë e nevojshme të kërkon personel shtesë. Mund të bëhen njoftime ose reklamat të vendosura në gazetatat lokale ose në radio duke ftuar aplikantë të mundshëm ose mund të rekrutohen personel të përshtatshëm.
3. Hartimi i procedurave të listës: Kërkohet listimi kur marrim mostrat nga një zonë e kornizës (frame-t).

9.3 Atributet që duhet të plotësojnë anketuesit

Intervistuesi është pjesë kryesore e mbledhjes së të dhënave. Cilësia e të dhënave varet drejtëpërdrejt nga cilësia e punës së anketuesit.

- Anketuesit duhet të jenë të kualifikuar e motivuar për punën që kryejnë
- Anketuesit duhet të trajnohen në mënyrë që të jenë në gjendje të kuptojnë kompleksitetin e natyrës së punës së tyre dhe konceptet e vërtetimit
- Anketuesit duhet të kenë aftësi të mira komunikimi dhe në përputhje me metodën e mbledhjes së të dhënave (PAPI, CAPI, CATI)
- Anketuesit duhet të kenë njohuri për qëllimet e vërtetimit dhe pyetësozin



REPUBLIKA E SHQIPËRISË
INSTITUTI I STATISTIKAVE

- Anketuesit duhet të lexojnë pyetjet në mënyrë të qartë, të saktë dhe me ritmin e duhur, pa kaluar asnjë pyetje apo grupuar ato.
- Anketuesit duhet të jenë dëgjues kompetentë të aftë për të regjistruar me saktësi përgjigjet.
- Anketuesit duhet të jenë të organizuar, në përdorimin e kohës dhe të jenë efikasë.
- Anketuesit duhet të jenë në gjendje të ruajnë një ton neutral të zërit gjatë gjithë intervistës.
- Anketuesit duhet të jenë të aftë të bashkëveprojnë me njerëz me formim të ndryshëm.

9.4 Monitorimi

Komponentet thelbësore për krijimin dhe ruajtjen e cilësisë së vërtetimit përfshijnë:

- Trajnimin e përshtatshëm dhe përgatitjen e të intervistuesve
- Mbikëqyrjen gjithëpërfshirëse të punës në terren, duke përfshirë monitorimin e performancës dhe sigurimin e cilësisë.

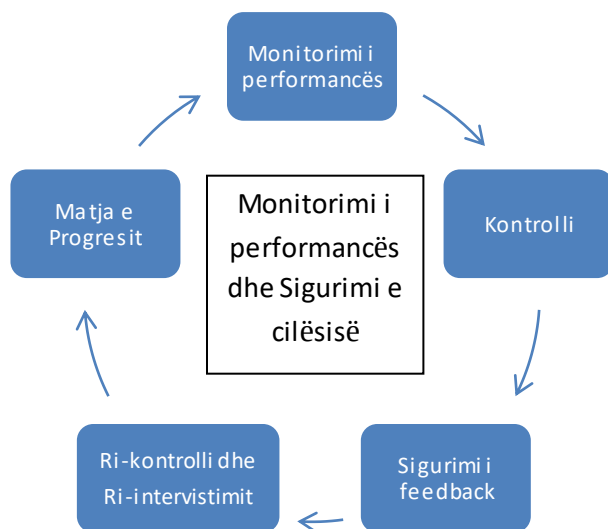


Figura 4: Fazat e monitorimit të vërtetimeve

10. Përpunimi i të dhënave

Kjo fazë përshkruan pastrimin e të dhënave të regjistruara dhe përgatitjen e tyre për analizë. Ajo është e përbërë nga nën-procese që kontrollojnë, pastrojnë dhe transformojnë të dhënat e mbledhura dhe mund të përsëritet disa herë.



10.1 Klasifikimi dhe kodifikimi

Ky nën-proces klasifikon dhe kodon të dhënat hyrëse. Për shembull, kodimi automatik ose manual ku përgjigjeve tekst u caktohen kode numerike, sipas një skeme klasifikimi të paracaktuar.

10.2 Editimi dhe imputimi i të dhënave

Kur të dhënat janë konsideruar të pasakta ose të munguara, vlera të reja mund të imputohen ose të editohen në këtë nën-proces. Termat editim dhe imputim përfshijnë një shumëllojshmëri metodash për tu zbatuar shpesh duke përdorur një qasje të bazuar në rregulla.

Hapat specifike zakonisht përfshijnë:

- Përcaktimin se çfarë e dhënë duhet të shtohet dhe ndryshohet.
- Përzgjedhjen e metodës së imputimit.
- Shtimin apo ndryshimin e të dhënave.
- Shtimin e vlerave të reja në dataset dhe shënjimi i tyre me një variabël shtesë.
- Prodhimi i indikatorëve të cilësisë në lidhje me imputimin dhe editimin.

Me metoda të përshtatshme statistikore ne vlerësojmë të dhënat që ne nuk ishim në gjendje për ti marrë në fazën e mbledhjes së të dhënave (të dhënat që mungojnë). Ka shumë metoda imputimi, por në thelb ato mund të ndahen në dy grupe:

- Metodat deterministe: Vlera e parashikuar është llogaritur me anë të një procedure analitike duke përdorur funksionin e duhur determinist.
- Metodat stokastike: Procedura e llogaritjes së vlerësimeve për vlerat që mungojnë është bazuar në procedurën që përdor një lloj mekanizmi probabilitar.

10.3 Metoda imputimi të të dhënave

Mospërgjigjet janë pjesë e trajtimit të të dhënave statistikore. Mospërgjigjet ndahen në dy grupe:

- mospërgjigje të plota, mospërgjigje të njësive së vrojtimit
- dhe mospërgjigje të pjesshme.

Mospërgjigjet e plota mund të ndahen në 3 kategori të mëdha:

- 1- Të përgjigjura
- 2- Refuzime ose pa kontakte
- 3- Jashtë qëllimit të vrojtimit



REPUBLIKA E SHQIPËRISË
INSTITUTI I STATISTIKAVE

Kryesisht mospërgjigjet e tilla, kategoria 2 dhe 3 e mësipërme trajtohen përmes procesit të ripeshimit të të dhënave trajtuar më sipër në kapitullin trajtimet e mospërgjigjeve nëpërmjet peshimit të të dhënave.

Mospërgjigjet e pjesshme trajtohen përmes metodave të imputimit të të dhënave me qëllim krijimin e të dhënave konsistente.

- Metoda e imputimit me mesatare.
- ‘Cold-deck’
- ‘Hot-deck’
- Imputimi me raport.
- Imputimi sipas një regresioni.
- Metoda të avancuara të trajtimit të të dhënave të munguara, Imputime.

10.4 Derivimi i variablave të reja

Ky nën-proces derivon (vlera për) variablat dhe njësitë statistikore që nuk janë përftuar në mënyrë eksplicite në fazën e mbledhjes së të dhënave, por që janë të nevojshme për dërgimin e outputit të kërkuar. Ai derivon variabla të reja duke aplikuar formula aritmetike të një ose më shumë variablave që janë tashmë të pranishme në setin e të dhënave.

Kjo mund të ketë nevojë të jetë përsëritëse, pasi disa variabla të derivuara në vetvete janë bazuar në variablat e tjera të derivuara. Për këtë arsye është e rëndësishme të sigurohet se variablat janë derivuar saktë. Njësi të reja statistikore mund të derivohen duke agreguar ose ndarë të dhënat nga njësitë mbledhëse ose me metoda të tjera të ndryshme të vlerësimit.

10.5 Të dhënat e agreguara

Ky nën-proces krijon të dhëna agregate dhe të dhëna totale për njësinë e observuar nga të dhënat mikro. Në procesin e llogaritjes së statistikave të dhënat janë përpunuar me programin e përgatitur, në mënyrë që llogaritjet e nevojshme, ri-kalkulimet dhe transformimet të zbatohen dhe rezultatet e agreguara statistikore të quajtur statistika, të përgatiten. Statistikat janë futur në një bazë të dhënash (makro) dhe/ose në një skedar të standardit të përdorur për nxjerrjen e rezultateve. Në rastin kur një kampion i rastit përdoret së bashku me llogaritjen e statistikave duhet të llogariten edhe vlerësimet e gabimit standard të këtyre statistikave.

10.6 Përgatitja e skedarëve finale

Ky nën-proces bashkon rezultatet e nën-proceseve të tjera në këtë fazë dhe rezultatet në një skedar të dhënash (zakonisht të të dhënave makro) i cili përdoret si input në fazën 6 (Analiza).



REPUBLIKA E SHQIPËRISË
INSTITUTI I STATISTIKAVE

Ndonjëherë ky mund të jetë ndërmjetës, jo një skedar final, veçanërisht për proceset statistikore ku ka kërkesa për të prodhuar vlerësimet paraprake dhe ato përfundimtare.

11. Shpërndarja e të dhënave

Pasi të ketë analizuar të dhënat, ekipi i Menaxhimit të Vrojtimit tani duhet t'i raportojë rezultatet tek përdoruesit për t'i ndihmuar ata të marrin përgjigjen e pyetjeve që ngritën në këtë studim. Në të njëjtën kohë, ekipi do të vlerësojë të dhënat në lidhje me objektivat e anketës.

11.1 Raporti kryesor i anketës

Trupi i raportit fillon me një prezantim ku jepet një përshkrim i shkurtër i objektivave të vrojtimit, organizatave të ndryshme të përfshira në hartimin dhe ekzekutimin e tij, përdoruesit kryesorë të pritshëm dhe një përmbledhje të metodologjisë së miratuar.

Pjesa më e madhe e raportit kryesor të vrojtimit përbëhet nga tabelat e para-planifikuara që rrjedhin drejtpërdrejt nga baza e të dhënave të përgjigjeve të pyetësorëve. Në mënyrë që këto tabela të jenë më të dobishme për përdoruesit, ekipi përgatit një raport të shkruar-tekst që përmbledh rezultatet më të rëndësishme me diskutimin e kuptimit të tyre në drejtim të objektivave origjinal të anketës, dhe komentimin e pikave të forta dhe të dobëta të të dhënave. Teksti i raportit përfshin një numër grafikësh dhe tabelash për të ilustruar pikat kryesore. Ky raport gjithashtu përfshin komente për pika të ndryshme nga të dhënat në bazën e të dhënave kryesore; për shembull, ai përmend probleme të veçanta të hasura gjatë punës në terren, ose variabla që kërkojnë norma jashtëzakonisht të larta imputimi. Gjithashtu nuk duhet të harrojmë pa përmendur raportet e tjera të planifikuara që do të publikohen në një datë të mëvonshme (Seksioni 11.3).

Seksioni i fundit kryesor i raportit nxjerr përfundime në lidhje me situatën e përgjithshme në referuar objektit të vrojtimit dhe përfshin rekomandime për një studim të mëtejshëm të disa pikave kryesore që nuk janë përfshirë plotësisht në këtë raund të parë të vrojtimit. Tabelat në shtojcat që shoqërojnë raportin duhet të paraqesin shpërndarjen e frekuencës së variablave, si dhe shumat, mesataret dhe proporcionet e vlerësuara përkatëse për të gjitha variablat e gjeneruara direkt nga pyetësi, ose që rrjedhin prej tyre gjatë përpunimit. Vlerësimet shoqërohen nga vlerësimet e gabimeve e standarde të tyre.



11.2 Raporti i Cilësisë së të Dhënave

Ekipi i Menaxhimit të Vrojtimit është gjithashtu në procesin e përpilimit të një raporti gjithëpërfshirës për cilësinë e të dhënave nga raundi i parë i vrojtimit, për të ndihmuar në zhvillimin e vrojtimit për vitet në vijim. Raporti do të përfshijë masa të ndryshueshmërisë së kampionimit, siç janë koeficientët e variacionit ose efektet e dizajnit. Ajo gjithashtu do të përfshijë shkallët e mungesave, nivelet e përgjigjes (si në total - nivel njësie ashtu edhe në nivel variablash) dhe shumë nga masat dhe indekset e krijuara nga hapat e kontrollit të cilësisë të vendosura për kodifikimin, kapjen e të dhënave dhe fazat e tjera të vrojtimit. Ai gjithashtu do të përshkruaj procesin e redaktimit dhe imputimit dhe do të diskutojë normat e imputimit dhe çdo problem tjetër që gjenden në të dhënat gjatë këtij procesi. Më në fund, ai do të komentojë mbi këshillueshmërinë e krahasimit të rezultateve të Vrojtimit me burime të tjera të disponueshme të të dhënave, duke theksuar krahasueshmërinë ose ndryshimet në koncepte dhe mjetet e matjes të përdorura dhe nëse ndonjë kalibrim në burim më të besueshëm është përdorur për disa nga të dhënat. Më poshtë është një listë e disa prej vlerave të shqyrtuara në raport:

- Koeficientët e variacionit (në nivel vendi, domain-e ose sipas qarkut);
- Efektet e dizajnit (p.sh sipas qarkut);
- Shkalla e mungesave (në nivel urbane / rurale);
- Shkalla e mos-përgjigjes (sipas qarkut dhe urbane / rurale brenda qarkut);
- Shkalla e mos-përgjigjes sipas llojit të mos-përgjigjes;
- Të dhënat e përdorura për poststratifikimin;
- Nivelet e dështuara të redaktimit sipas pyetjes;
- Shkalla e imputimit sipas pyetjes;
- Shkalla e gabimit të kodimit;
- Shkalla e gabimit në kapjen e të dhënave;
- Numri mesatar i kontakteve për të përfunduar një rast përgjigjeje;
- Numri mesatar i kontakteve për të përfunduar një rast mos-përgjigjes;
- Koha mesatare e intervistës për rastet e përgjigjes.

11.3 Raporte të tjera

Ekipi i Menaxhimit të Vrojtimit ka të ngjarë të hartojë disa raporte shtesë duke përfshirë:

- i. Raporti i Vlerësimit të Anketës. Ai do të sigurojë rekomandimet e ekipit për përmirësime në procesin e vrojtimit në mënyrë që raundi tjetër i tij të funksionojë edhe më mirë dhe me efikasitet.



REPUBLIKA E SHQIPËRISË
INSTITUTI I STATISTIKAVE

- ii. Raportet e detajuara të analizës. Ekipi i Menaxhimit të Vrojtimit planifikon të prodhojë një seri raportesh analitike që do të jenë shkruar në bashkëpunim me agjensi të ndryshme të përdoruesve.
- iii. Raporte teknike. Këto raporte do të japin detaje mbi metodologjinë e anketës, planin e marrjes së kampioneve, metoda e vlerësimit, procedurat e mbledhjes dhe përpunimit të të dhënave, etj.
- iv. Raportet e cilësisë sipas versionit standart SIMS të rekomanduar nga Eurostat, në të dy format e tij (ESMS - orientuar nga përdoruesi i statistikave dhe ESQRS - orientuar nga prodhuesi i Statistikave).

11.4 Kontrolli i konfidencialitetit dhe zbulimit

Sa i përket rezultateve përfundimtare të vrojtimit, kontrolli i zbulimit është një çështje e rëndësishme. Meqenëse nuk do të ketë microdata file me përdorim publik, kontrolli i zbulimit përqendrohet në tabelat e publikuara dhe masat e tjera përmbledhëse. (Sidoqoftë, Instituti i Statistikave ka të ngjarë të lejojë hyrjen e mikrodateve e disa analistëve të autorizuar pasi të jenë betuar sikur të ishin punonjës të institutit. Atëherë ata do t'i nënshtroheshin të gjitha dënimeve për shkelje të konfidencialitetit.)

Në lidhje me të dhënat e publikuara në raportin kryesor të anketës, ekipi po studion metodat e kufizuara në dispozicion për të dhënat tabelare. Duke qënë se, një nga qëllimet kryesore të raundit të parë të vrojtimit është marrja e informacionit të përmirësuar, për rregullimin e mirë të tij në vazhdim, strategjia e përgjithshme e miratuar është të jetë shumë e konservative. Kështu, ekipi është i gatshëm të shtypë shumë informacion nëse është e nevojshme, ndërsa bën vlerësime të hollësishme të cilësisë së të dhënave dhe rrezikut të dhënies së informacionit për zbatimin e mëvonshëm të rregullave të zbulimit që do të maksimizojnë sasinë e informacionit të publikueshëm në të ardhmen.

Për këtë arsye, ekipi ka adoptuar një rregull paraprak prerjeje bazuar në [Ligjin e Statistikave Zyrtare](#), Neni 31- Konfidencialiteti, në të cilin thuhet: “Për mbrojtjen e identifikimit përfshihen në total të paktën tri njësi dhe pjesa e një njësie nuk duhet të kalojë 85 për qind të totalit”, përpara se të dhënat të publikohen.

11.5 Përmbledhje

Shpërndarja dhe vlerësimi i të dhënave janë hapa shumë të rëndësishëm në një vrojtim. Qëllimi është të komunikojë informacionin tek përdoruesit në mënyrë që ata të jenë në gjendje të kuptojnë rezultatet e anketës dhe të marrin vendime. Vlerësimi i të dhënave duhet të vlerësojë



REPUBLIKA E SHQIPËRISË
INSTITUTI I STATISTIKAVE

rezultatet përfundimtare në dritën e objektivave original të anketës. Ai duhet të tregojë pikat e forta dhe të dobëta të vrojtimit në mënyrë që përdoruesit të mund të vlerësojnë shkallën në të cilën gabimet në të dhëna kufizojnë përdorimin e tyre.

Metodat e shpërndarjes përfshijnë raporte të shkruara me tabela dhe grafikë dhe / ose një skedar të mikrodatës me përdorim publik. Sidoqoftë, përpara se të shpërndahen rezultatet e vrojtimit (ose të dhënat), ato duhet të testohen plotësisht për të siguruar respektimin e konfidencialitetit të të anketuarve.

Referencat

- Andersson, C. & Nordberg, L. (1994). A Method for variance estimation of non-linear functions of totals in surveys ! theory and a software implementation. Journal of Official Statistics, Vol. 10, No. 4, 395-406.
- Andersson C. (2002). CLAN97 v. 3.1. Supplement to A User's Guide to CLAN97. (Unpublished).
- Biemer P. & Lyberg L. (2003). Introduction to Survey Quality. New York: John Wiley & Sons.
- Brewer K. & Hanif F. (1983). Sampling with Unequal Probabilities. New York: SpringerVerlag.
- Cochran, W.G. (1977). Sampling Techniques, 3rd ed. New York: John Wiley & Sons.



REPUBLIKA E SHQIPËRISË
INSTITUTI I STATISTIKAVE

- Deville J-C., & Särndal C-E. (1992). Calibration Estimators in Survey Sampling. Journal of the American Statistical Association, Vol. 87, No. 418, 376!382.
- Deville J-C., Särndal C-E., & Sautory O. (1993). Generalized Raking Procedures in Survey Sampling. Journal of the American Statistical Association, Vol. 88, No. 423, 1013-1020.
- European Union (1998). Council regulation (EC) No 577/98 of 9 March 1998 on the organisation of a labour force sample survey in the Community.
- Groves R.M., Dillman D.A., Eltinge J.L., & Little R.J.A. (eds.). (2002). Survey Nonresponse. New York: John Wiley & Sons.
- Groves R.M., Fowler F.J., Couper M.P., Lepkowski J.M., Singer E. & Tourangeau R. (2004). Survey Methodology. Hoboken: John Wiley & Sons.
- Hansen M.H., Hurvitz W.N. & Madow W.G. (1953). Sample survey methods and theory. New York: John Wiley & Sons.
- Kish L. (1965). Survey sampling. New York: John Wiley & Sons.
- Lehtonen R. & Pahkinen E. (2004). Practical Methods for Design and Analysis of Complex Surveys. 2nd ed. Chichester: John Wiley & Sons.
- Lehtonen R., Särndal C.-E. & Veijanen A. (2005). Does the model matter? Comparing modelassisted and model-dependent estimators of class frequencies for domains. Statistics in Transition, 7, 649–673.
- Lessler J. & Kalsbeek W. (eds.). 1992. Nonsampling Error in Surveys. New York: John Wiley & Sons.
- Lohr S. (1999). Sampling: Design and analysis. Pacific Grove: Duxbury Press.
- Lyberg L., Biemer P., Collins M., de Leeuw E. D., Dippo C., Schwarz N. & Trewin D. (eds.). (1997). Survey Measurement and Process Quality. New York: John Wiley & Sons.



REPUBLIKA E SHQIPËRIË
INSTITUTI I STATISTIKAVE

- Lynn P., Beerten R., Laiho J. & Martin J. (2003). Towards standardisation of survey outcome categories and response rate calculations. *Research in Official Statistics*, Vol. 5 1/2002, 61-84.
- Oh J.L. & Scheuren F. (1983). Weighting Adjustment for Unit Nonresponse, in W.G. Madow, I. Olkin & D.B. Rubin (eds.) *Incomplete Data in Sample Surveys*, Vol 2., 143-184. New York: Academic Press.
- Rubin D. (1987). *Multiple Imputation for Nonresponse in Surveys*. New York: John Wiley & Sons.
- Sautory O. (1993). La macro CALMAR. Redressement d'un échantillon par calage sur marges. I.N.S.E.E., Série des documents de travail nE F 9310. Paris: I.N.S.E.E.
- Sautory, O. & Le Guennec, J. (2005). La macro CALMAR2. Redressement d'un échantillon par calage sur marges. I.N.S.E.E. Paris: I.N.S.E.E.
- Schafer J. L. (1997). *Analysis of Incomplete Multivariate Data*. London: Chapman & Hall.
- Smith T.M.F. (1990). Post-stratification. *The Statistician*, 40, 315!323.
- Sundgren B. (1999). Information systems architecture for national and international statistical offices. Guidelines and recommendations. Conference of European Statisticians. *Statistical Standards and Studies*. No 51. Geneva: United Nations.
- Särndal C-E. & Lundström S. (2005). *Estimation in Surveys with Nonresponse*. Chichester: John Wiley & Sons.
- Särndal C-E., Swensson B. & Wretman J. 1992. *Model Assisted Survey Sampling*. New York: Springer-Verlag.
- United Nations (1950). *The Preparation of Sampling Survey Reports*. Statistical Papers Ser. C No. 1 (Revised). New York.



REPUBLIKA E SHQIPËRIË
INSTITUTI I STATISTIKAVE

- United Nations (1964). Recommendation for the Preparation of Sample Survey Reports (Provisional Issue). Statistical Papers Ser. C No. 1 Rev.2. New York.
- Wolter K. (2007). Introduction to Variance Estimation, Second Edition. New York: Springer.
- EUREEDIT (The Development and Evaluation of New Methods for Editing and Imputation): [http://www.cs.york.ac.uk/euredit/](http://www.cs.york.ac.uk/eureedit/)
- Eurostat Quality site: <http://ec.europa.eu/eurostat/quality>
- The European Statistics Code of Practice ESS Quality standards: Standard Quality Report Standard Quality Indicators How to Make a Quality Report Handbook on improving quality by analysis of process variables DESAP – self-assessment for survey managers
- International Monetary Fund. Data quality assessment framework. 2003, http://dsbb.imf.org/vgn/images/pdfs/dqrs_Genframework.pdf
- Organisation for Economic Cooperation and Development. OECD PISA database, 2001: <http://pisaweb.acer.edu.au/oecd/>
- Organisation for Economic Cooperation and Development. Quality framework and guidelines for OECD statistics. 2003. <http://www.oecd.org/dataoecd/26/42/21688835.pdf>
- Statistics Finland: Quality Guidelines for Official Statistics: http://stat.fi/tk/tt/laatuatilastoissa/alku_en.html
- VLISS-virtual laboratory in survey sampling: <http://www.math.helsinki.fi/VLISS/>.
- Statistics Canada. 2000d. “Policy on Informing Users of Data Quality and Methodology.” Statistics Canada
- Policy Manual. Section 2.3. Last updated March 4, 2009. http://icn-rci.statcan.ca/10/10c/10c_010_e.htm.
- Statistics Canada. 2002c. Statistics Canada’s Quality Assurance Framework - 2002. Statistics Canada Catalogue no. 12-586 http://method/Branch/Committees/MethodsAndStandards/QualityAssuranceFramework_e.htm.
- Statistics Canada. 2003d. Statistics Canada Policy Manual. Last updated May 27, 2009. http://icn-rci.statcan.ca/10/10c/10c_000_e.htm.
- Statistics Canada. 2000. “Policy on Informing Users of Data Quality and Methodology.” Statistics Canada Policy Manual. Section 2.3. Last updated March 4, 2009. http://icn-rci.statcan.ca/10/10c/10c_010_e.htm.



REPUBLIKA E SHQIPËRISË
INSTITUTI I STATISTIKAVE