

Interne notater

STATISTISK SENTRALBYRÅ

82/33

12. oktober 1982

TRE NOTATER OM GJENNOMFØRING AV UTVALGSUNDERSØKELSER I BYRÅET

	Side
ERLING SIRING: Utvalgsregistre og rutiner for trekking av person- og hus- holdningsutvalg	1
ROLF AABERGE: Metode for merking av uttrekte personar/familiar i 10 %- registeret	11
TOR HALDORSEN: Handtering av registerfeil, føring av frafall og avgang ..	25

UTVALGSREGISTRE OG RUTINER FOR TREKKING
AV PERSON- OG HUSHOLDNINGSUTVALG

av

Erling Siring

INNHOOLD

	<u>Side</u>
1. Innledning.....	2
2. Byråets utvalgsregistre.....	2
2.1. Oppretting og ajourføring.....	2
2.2. Muligheten for ajourføring 2 ganger i året.....	4
2.3. Kvaliteten av utvalgsregistrene.....	4
2.4. Organisering av utvalgsregistrene.....	5
2.5. Oppretting av delregistre.....	5
3. Rutine for trekking av utvalg.....	6
3.1. Beskrivelse av rutinen.....	6
3.2. Utfylling av parameterkort.....	7
3.3. Problemer med å trekke utvalg fra et register.....	7
3.4. Ajourføring av utvalg ved kobling med Person- registeret.....	8
4. Endring av folketallene i utvalgsområdene.....	9
5. Referanser.....	10

1. INNLEDNING

Utvalg til person- og husholdningsundersøkelser blir i Byrådet trukket fra to delregistre, som igjen blir trukket fra hvert sitt register. Disse to registrene er de såkalte utvalgsregistrene. Det er et utvalgsregister for Oslo, og et for resten av landet. Grunnen til at det er et eget register for Oslo, er at Oslo Folkeregisters magnetbånd, som danner grunnlaget for registeret, er bygd opp litt forskjellig fra registrene ellers i landet.

Hensikten med dette notatet er å beskrive utvalgsregistrene, og gi en oversikt over rutinene som blir brukt ved trekking av utvalg. Det er tidligere blitt laget et liknende notat (Sæbø (76)), men siden det har skjedd enkelte endringer i rutinene, har det vært viktig å få ajourført beskrivelsen av rutinene. I tillegg har det vist seg å være behov for å få beskrevet en del detaljer ved rutinene som ikke er beskrevet tidligere.

Sæbø (76) er forøvrig en av hovedkildene til dette notatet, og det er derfor en del overlapping mellom de to notatene.

I kapittel 2 beskrives hvordan registrene dannes og ajourføres. Det blir også beskrevet hvordan registrene er oppbygd og sortert. I kapittel 3 blir rutinen for trekking av utvalg beskrevet og hvordan parameterkortene skal utfylles. Videre blir muligheten for ajourføring av utvalg ved kobling med Personregisteret drøftet.

I kapittel 4 beskrives hvilke konsekvenser det har at folketallene i utvalgsområdene endrer seg, og at en til mindre undersøkelser bruker en tredjedel av hvert utvalgsområde.

2. BYRÅETS UTVALGSREGISTRE

2.1. Oppretting og ajourføring

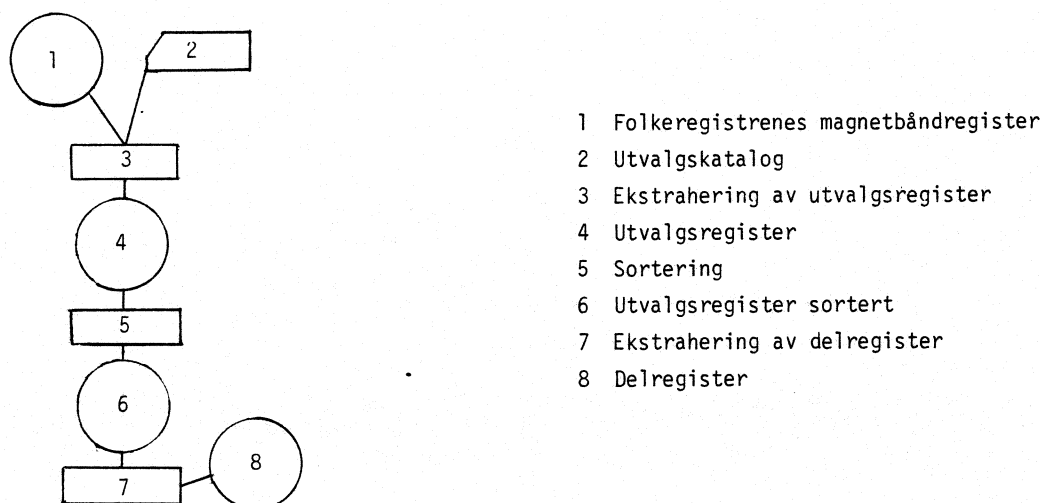
Grunnlaget for Byråets utvalgsregistre er folkeregistrenes magnetbåndregister. Dette inneholder alle familier og personer bosatt i Norge, og det blir ajourført etter oppgaver fra de lokale folkeregistre. Båndet som Byrådet bruker ved ajourføringen stammer fra Skattedirektoratet. Ajourføringen av utvalgsregistrene foregår ved at det rett og slett lages nye utvalgsregistre med ajourførte data fra folkeregistrene som grunnlag.

I Byråets utvalgsplan foregår trekking av personer og husholdninger i to trinn. I første trinn er det trukket områder som vi kaller utvalgsområder. Hvert av disse områdene består av en kommune eller i noen tilfeller av to eller flere mindre kommuner. I annet trinn trekkes personer eller husholdninger fra de uttrukne områdene.

Trekkingen på første trinn foretas ikke for hver ny utvalgsundersøkelse. En bruker nå utvalgsområder som ble trukket i 1974. I disse områdene har en bygd opp en fast intervjuerstab. Hvert utvalgsområde er delt opp i flere delområder for at en skal kunne fordele intervjuobjektene (IO-ene) på intervjuerne. I Stavanger, Kristiansand og Bærum er det 6 slike delområder, i Trondheim 9, i Bergen 18 og i Oslo 42, mens det i resten av de uttrukne utvalgsområdene er 3. Ved mindre undersøkelser bruker vi bare 1/3 av delområdene (ett sett), slik at vi i de fleste uttrukne utvalgsområdene utenom de største byene bruker én intervjuer. Oppdelingen av utvalgsområdene i delområder er foretatt etter valgkrets-grenser. I Oslo brukes sone- og rodegrenser ved denne oppdelingen. For en nærmere beskrivelse av utvalgsplanen henvises til Thomsen og Rideng (1974) og Thomsen (1977).

Det er to utvalgsregistre, ett for Oslo og ett for resten av landet. Som nevnt i innledningen er det eget register for Oslo p.g.a. at magnetbåndet til Oslo Folkeregister er bygd opp annerledes enn magnetbåndene til folkeregistrene ellers i landet. De to utvalgsregistrene vil i nærmeste framtid bli slått sammen til ett register. Utvalgsregistrene inneholder alle familier og personer bosatt i utvalgsområdene. Det er én record for hver person. Ekstrahering av utvalgsregisteret for landet utenom Oslo foregår etter en katalog av parameterkort. Denne katalogen angir hvilke kommuner og valgkretser som hvert utvalgsområde består av. Figur 1 er et skjematisk oppsett over prosedyren ved opprettingen av dette registeret. Den videre framgangsmåten ved uttrekking av et delregister er også vist her. Denne prosedyren blir beskrevet i avsnitt 2.5.

Figur 1. Rutiner for oppretting av utvalgsregister for landet utenom Oslo



Som nevnt tidligere foregår ajourføringen av utvalgsregistrene ved at det rett og slett blir opprettet nye registre. De uttrukne utvalgsområdene holdes fast ved ajourføringen av registeret, mens den videre oppdelingen av utvalgsområdene ofte justeres. Justeringer er nødvendige når valgkretsgrenser endres eller når det kommer nye valgkretser til. Ellers kan det hende at en foretar justeringer av oppdelingen i tre sett (heretter kalt tredelingen) for å få redusert utvalgsvariansen og/eller feltkostnadene.

Byrået får ajourførte bånd fra folkeregistrene i begynnelsen av hvert år. Det er meningen å ajourføre utvalgsregistrene hver år i mars/april. Ajourføringen begynner med at Underavdeling for intervjuundersøkelser får tilsendt tabeller fra 1. kontor som viser antall personer etter alder i de forskjellige valgkretsene. Disse tabellene blir sammenlignet med de tilsvarende tabellene fra året før for å se om folketallene har endret seg i de enkelte valgkretsene. Hvis det har skjedd større endringer i en valgkrets, kan dette ha sammenheng med at valgkretsgrensene er blitt endret. Hvis det har skjedd en endring i folketallet i en valgkrets på ca. 10% eller mer, skriver en til det lokale folkeregisteret og spør om det har skjedd noen endringer i valgkretsgrensene. Videre blir folkeregisteret bedt om å tegne inn eventuelle endringer på et kart som følger med brevet, og om å returnere kartet til Byrået.

Når en foretar justeringer av oppdelingen av utvalgsområdene i sett, tar en hensyn til reisetid og arbeidsbyrde for intervjuerne, og ønsket om at folketallene i de tre delene av et utvalgsområde skal være så like som mulig. Justeringer av "tredelingen" innebærer at katalogen over utvalgsområdene blir endret.

Ved ajourføringen av utvalgsregistrene ser en nærmere på "tredelingen" bare hvis det har skjedd større endringer i "tredjedelsområdene" fra året før. Hvis det skjer bare små årlige endringer i forholdet mellom innbyggertallene i "tredjedelsområdene" innen et utvalgsområde, kan slike små endringer over tid oppsummere seg til store endringer uten at det blir oppdaget. Med jevne mellomrom bør en derfor gå gjennom tredelingen av utvalgsområdene for å se om innbyggertallene varierer mye mellom de tre settene.

Arbeidet med justeringen av "tredelingen" forsinker ajourføringen av utvalgsregistrene med bortimot 2 måneder hvert år. Grunnen til at dette arbeidet tar så lang tid, er at enkelte av de lokale folkeregistrene er svært seint til å svare på henvendelsene fra Byrået.

6. kontor har foreslått for Intervjuavdelingen at en i stedet for å bruke valgkretser som grunnlag for oppdelingen av utvalgsområdene, går over til å bruke grunnkretser. Dette vil ha flere fordeler. Grunnkretsene er langt mer stabile enn valgkretsene. Når det en sjelden gang skjer en endring i grunnkretsene, består endringen i enten at én grunnkrets blir delt i to eller at to grunnkretser blir slått sammen til én. Endringer i grunnkretsgrensene kunne en få beskjed om fra 6. kontor, slik at en kunne slippe de tidkrevende henvendelsene til de lokale folkeregistrene. En annen fordel med grunnkretssystemet er at grunnkretsene er mindre enheter enn valgkretsene, slik at en ved bruk av grunnkretssystemet kunne få en mer tilfredsstillende oppdeling av utvalgsområdene. Nå er det f.eks. ofte slik at en valgkrets er felles for to tredjedelsområder.

Foreløpig er ikke opplysningen "grunnkrets" lagt inn i noe sentralt register som er aktuelt som ajourføringsgrunnlag for utvalgsregistrene. Det eksisterer imidlertid et adresseregister i Byrået med grunnkretsopplysninger. Ved å koble dette registeret med utvalgsregistrene skulle det være mulig å få lagt "grunnkrets" inn i utvalgsregistrene. Hvilke praktiske vanskeligheter dette kan by på, og om ajourholdet av adresseregisteret er godt nok til å kunne ta i bruk grunnkretssystemet i nærmeste framtid, har vi ikke helt oversikt over. Muligheten til å ta i bruk grunnkretssystemet bør i hvert fall vurderes.

Arbeidet med justeringen av oppdelingen av utvalgsområder i delområder blir foretatt av Databearbeidingsgruppa. Det er imidlertid Metodegruppa som har det faglige ansvaret for dette arbeidet. Før justeringer foretas skal Feltgruppa tas med på råd.

Når dette arbeidet er ferdig, blir den ajourførte katalogen over utvalgsområdene oversendt Driftskontoret som så oppretter nye utvalgsregistre ved hjelp av ferdiglagde datarutiner.

2.2. Muligheten for ajourføring 2 ganger i året

Som nevnt tidligere blir utvalgsregistrene ajourført én gang hvert år i mars/april. Byrået får imidlertid ajourførte bånd fra folkeregistrenes magnetbåndregister 2 ganger i året, slik at det er mulig å ajourføre utvalgsregistrene både i mars/april og i august/september. Det er nå avgjort at en fra 1982 av skal ajourføre registrene to ganger i året.

Ved ajourføringen i august/september trenger en ikke justere "tredelingen" siden endringer i valgkrets- og kommunegrenser bare blir foretatt ved årsskiftet. Kostnadene ved ajourføringen i august/september vil da i hovedsak utgjøres av kostnadene ved kjøringene på datamaskinen. Kostnadene ved ajourføring av utvalgsregistrene var i 1981 om lag 20 000 kr. I denne summen var inkludert maskinleie til Driftssentralen, kostnader ved bruk av Byråets egne båndstasjoner og operatørkostnader.

2.3. Kvaliteten av utvalgsregistrene

En faktor som hyppigheten av ajourføringene betyr mye for kvaliteten av utvalgsregistrene. Det er imidlertid klart at uansett hvor ofte en ajourfører utvalgsregistrene, kan aldri kvaliteten av disse bli bedre enn kvaliteten av det registeret som danner grunnlaget for ajourføringen, nemlig folkeregistrenes magnetbåndregister. Hvor godt sistnevnte register er, skal vi ikke gi noen generell vurdering av her. Vi skal derimot trekke fram en konkret svakhet ved registeret.

Registeret er ikke bygd opp som ett register, men sammensatt av mange småregistre (ett for hver kommune). Dette betyr at ikke hele registeret blir oppdatert på en gang. For å illustrere hva dette kan føre til skal vi anta at en person flytter fra kommune A til kommune B. Hvis kommune A er raskere enn kommune B til å sende flyttemelding til det sentrale magnetbåndregisteret, vil personen ikke finnes i registeret i det hele tatt mellom de to meldingstidspunktene. Hvis omvendt kommune A er seinere enn kommune B til å melde om flytting, vil personen være registrert bosatt i begge kommunene mellom de to meldingstidspunktene. På grunn av dette er det ca. 5 000 dubletter i registeret. I delregistrene som en trekker utvalgene fra (jfr. avsn. 2.5) blir det imidlertid foretatt dublett kontroll, slik at det ikke er noen dubletter i disse registrene. Av personer som mangler i folkeregistrenes magnetbåndregister, regner en med at det til enhver tid er ca. 10 000. Siden mange av disse er spebarn, representerer dette egentlig ikke noe stort problem.

Det ovenfor nevnte problemet kunne helt omgås hvis en tok i bruk Byråets sentrale personregister som grunnlag for ajourføringen av utvalgsregistrene, siden hele Personregisteret blir oppdatert på en gang. Personregisteret inneholder imidlertid ikke kjennetegnet "valgkrets" eller noe annet kjennetegn som kan danne grunnlaget for oppdelingen av utvalgsområdene, og er i dag derfor ikke egnet som grunnlag for ajourføringen.

I framtiden vil det sannsynligvis bli lagt inn kjennetegn i Personregisteret som kan kodes om til "grunnkrets". Hvis dette blir gjort og grunnkretssystemet viser seg å være velegnet som grunnlag for oppdelingen av utvalgsområdene, bør en vurdere om en skal bruke Personregisteret som ajourføringsgrunnlag for utvalgsregistrene. Foruten at en kunne unngå registerfeil som nevnt over, ville dette ha den fordel at en ville få muligheten til å ajourføre utvalgsregistrene langt oftere enn nå. Hvis en kunne få foretatt ajourføringene ved de ordinære kjøringene på Personregisteret, ville trolig også

kostnadene ved en ajourføring bli langt lavere enn nå.

2.4. Organisering av utvalgsregistrene

Utvalgsregistrene inneholder en record for hver person bosatt i utvalgsområdene. Når det gjelder en detaljert beskrivelse av recordenes innhold og redigering, henviser jeg til Sæbø (76). Recordenes innhold har imidlertid blitt endret noe. Felt nr. 18 heter nå "Deltakelse i tidligere undersøkelser; hovedpersoner". I utvalgsregistrene er dette feltet blankt. Opplysningen "Krets i rekkefølgenr." i registeret for landet utenom Oslo er kuttet ut. Ellers kan vi nevne at recordene bl.a. inneholder utvalgsområdenummer, kommunenummer, navn, adresse, familienummer, personkode, ekteskapelig status og statsborgerskap. Registeret for landet utenom Oslo inneholder også valg- og skolekrets, mens registeret for Oslo inneholder sone- og rodekode.

Redigeringen av recordene holdes fast ved alle sorteringer, ekstrahering av delregister og trekking av utvalg.

Registrene sorteres etter utvalgsområdenummer, familienummer og personkode. Når det gjelder nummereringen av utvalgsområdene og fordelingen av disse på sett, henviser jeg igjen til Sæbø (76).

Familienummeret er fødselsnummeret til familiens hovedperson. Sorteringen foregår etter alder på denne. Personkoden angir om en person er hovedperson, ektefelle eller barn. En familie i folke-registeret (familieenhet) består aldri av mer enn to generasjoner. En ugift mor med barn som bor hjemme hos foreldrene, utgjør f.eks. en egen familieenhet med eget familienummer. (En nøyaktig definisjon av familienummersystemet finnes i Gåsemyr (1970)). En husholdning kan ofte bestå av flere familier i registeret, og dette er et problem ved husholdningsundersøkelser. I praksis løses dette ved at vi trekker familier, men gir intervjuerne instruks om å ta med alle husholdningsmedlemmer.

Fødselsnummeret gir oss alder og kjønn. Det gir oss mulighet til å trekke utvalg av personer innen en viss aldersgruppe og av samme kjønn. Kombinerer vi dette med kode for ekteskapelig status, kan vi f.eks. trekke utvalg av gifte kvinner.

2.5. Oppretting av delregistre

Utvalgsregistre for landet utenom Oslo består av ca. 2 millioner records, mens vi i Oslo har med ca. 450 000. For å trekke et utvalg må en gjennomløpe hele registret. Dette blir dyrt med så store registre, og vi har derfor opprettet et mindre delregister (10%-register) for hvert register. Trekking av utvalg foregår fra disse.

Utvalgene som blir brukt ved de fleste av Byråets undersøkelser blir trukket slik at hver enhet i befolkningen får samme trekkesannsynlighet. Slike utvalg kalles selveiende. I Byråets utvalgsplan er utvalgsområdene trukket med ulike sannsynligheter. For at det endelige utvalg skal bli selveiende, må vi derfor trekke familier eller personer med ulik sannsynlighet i de forskjellige utvalgsområdene. Vi tar imidlertid hensyn til dette alt ved uttrekkingen av delregistrene, slik at disse består av et selveiende utvalg av befolkningen. I områder som er trukket med sannsynlighet 1 i første trinn (gjelder f.eks. Oslo), trekker vi ut 10%, mens andelen fra andre områder er større. Delregisteret vårt fra landet utenom Oslo inneholder fra 10% av befolkningen i de største byene til nær 100% i mindre utvalgsområder. I alt består det av ca. 360 000 records, mens delregisteret for Oslo har med 45 000. Vi må trekke ut nye delregistre ved hver ajourføring av hovedregistrene. Trekkingen foregår ved at familieenheter trekkes med faste mellomrom i utvalgsregistrene (systematisk trekking). Delregisteret for Oslo opprettes ved at hver 10. familie trekkes systematisk, mens trekkingen til delregisteret for landet utenom Oslo styres av en katalog som angir hvor stor del av familiene som skal trekkes fra hvert utvalgsområde.

Metodegruppa har ansvaret for at trekketetthetene som katalogen angir, er slik som de skal være. Trekketetthetene justeres ved hver ajourføring pga. merkerutinen som er beskrevet i notatdelen "Metode for merking av uttrekte personar/familiar i 10%-registeret". Hvordan trekketetthetene skal justeres er beskrevet i nevnte notatdel. 10%-registeret blir trukket ved hjelp av et ferdiglaget program på Driftskontoret (se figur 1).

Delregistrene er sortert etter utvalgsområdenummer, fødselsår, -måned, -dag og personnummer for hovedperson og personkode. Redigering og innhold av recordene er som i utvalgsregistrene, med unntak av at feltene nr. 18, nr. 20 og nr. 21 (jfr. Sæbø (1976)) som er blanke i utvalgsregistrene, har et innhold i delregistrene.

Feltene nr. 18 og 20 blir brukt til å merke personer som deltar i en undersøkelse, slik at de ikke skal bli trukket ut til flere undersøkelser etter hverandre. De som blir trukket ut, blir merket med avslutningstidspunkt for intervjuperioden i felt nr. 20. Ved familieutvalg blir i tillegg uttrukne hovedpersoner merket i felt nr. 18. For personer som ikke er merket, er felt nr. 18 og 20 blanke. I Aaberge (1982) er det mer om dette.

3. RUTINE FOR TREKKING AV UTVALG

3.1. Beskrivelse av rutinen

De fleste utvalgene som blir trukket i Byrået, blir ikke trukket fra utvalgsregisteret, men fra 10%-registeret. Siden 10%-registeret er et selvveiende utvalg på 10% av hele landets befolkning, får en selvveiende utvalg ved å gi alle personer i 10%-registeret samme sannsynlighet for å bli trukket ut til utvalgene.

Det er laget et dataprogram som tar seg av trekkingen. Programmet virker slik at det først velger et tilfeldig startpunkt, og så trekker ut 10-er med faste mellomrom i registeret. På grunn av at trekkingen er systematisk og registeret er sortert slik at personer fra samme familieenhet ligger etter hverandre i registeret, vil det ved trekking av personutvalg normalt bare bli trukket ut én person fra hver familieenhet.

Personer som er merket etter å ha deltatt i en tidligere undersøkelse gis en trekkesannsynlighet lik 0, dvs. slike personer blir forbilest under trekkingen.

Før et utvalg trekkes må trekkeettheten beregnes. Dette blir gjort på Underavdeling for intervjuundersøkelser på grunnlag av tabeller som oversendes fra Driftskontoret. Disse tabellene over såkalte basistall viser både antall enheter i alt i 10%-registeret (bruttotall) og antall umerkede enheter (netttotal) etter kjønn, alder og ekteskapeleg status. Ut fra en slik tabell kan en regne ut trekkeettheten enten trekkeenheten er familie, person, gifte menn mellom 30 og 40 år eller annet. Dette gjelder også hvis en ønsker å trekke fra bare ett av de tre settene. Når et utvalg trekkes, ajourføres disse tabellen automatisk (netttotalene endres) og skrives ut. Derved slipper en å gjennomløpe registeret en ekstra gang før hver trekking.

På Intervjuavdelingen blir tabellene over basistall som foreligger etter en trekking, sammenliknet med tabellene fra før trekkingen. Ved å sammenligne netttotalene før og etter en trekking kontrollerer en at det riktige antall trekkeenheter blir merket hvis utvalget skal merkes. På denne måten får en stadig sjekket at merkerutinen fungerer som den skal.

Etter at et utvalg er blitt trukket, blir det på Intervjuavdelingen også sjekket at antall trekkeenheter fra hver utvalgsområde er i samsvar med det antall trekkeenheter en har "bestilt" ifølge parameterkortet. Videre blir det sjekket at alders- og kjønnsfordelingen i utvalget likner på de tilsvarende fordelingene i hele populasjonen.

De ovenfornevnte arbeidsoppgaver som blir utført på Underavdeling for Intervjuundersøkelser, blir for tiden utført av Databearbeidingsgruppa. Metodegruppa har det faglige ansvaret for dette arbeidet (Jfr. Thomsen (1980)).

I Sæbø (1976) er deler av trekkerutinen beskrevet mer detaljert enn her. Videre finnes det et eksempel på en tabell over basistall og et skjematisk oppsett over trekkerutinen. I det nevnte notatet er også trekkerutinen for arbeidskraftundersøkelsene og rutiner for kontroll og klargjøring av utvalg beskrevet.

I Stenbekken (1977) er rutiner for trekking og oppretting av utvalg til forskjellige løpende undersøkelser beskrevet.

I Byrået trekkes også en del utvalg som skal oppfylle slike spesifikasjoner at den vanlige trekkerutinen ikke kan følges. I noen tilfeller trekkes utvalgene fra helt spesielle registre, og i andre tilfeller trekkes utvalgene direkte fra utvalgsregistrene. Det kan også forekomme at utvalg

trekkes fra "mini-registre som blir ekstrahert enten fra 10%-registrene eller fra utvalgsregistrene. Metodegruppa lager spesielle utvalgsplaner i slike tilfeller, og avgjør hvilket register det skal trekkes fra.

Når en ikke kan bruke den vanlige rutinen ved trekking av utvalg, kan en som oftest bruke et generelt trekkeprogram som finnes i Byrådet.

3.2. Utfylling av parameterkort

Ved trekking av utvalg fra 10%-registeret må to parameterkort fylles ut. Det ene parameterkortet fylles ut med 10 vilkårlige siffer som skal danne et vilkårlig startpunkt for trekkingen. Dette kortet fylles ut på Driftskontoret.

Det andre parameterkortet, som fylles ut på Intervjuavdelingen, angir hvilke enheter som skal ha mulighet for å komme med i utvalget og trekkesansynligheten for disse enhetene. Oppsettet er som følger:

Kolonne	Lengde	Feltbeskrivelse
1	1	Sett
2 - 5	4	Fødselsår fom./tom. (alderseleksjon)
6 - 10	5	Ekteskapelig status
11	1	Kjønn
12	1	Husholdning/person
13 - 14	2	Undersøkellesnr.
15	1	Puljenr.
16	1	Merking/ikke merking
17 -20	4	Trekkeetthet
21	1	Klyngestørrelse

Blankt i feltet "sett" medfører trekking fra alle tre sett. Kodene 1, 2, 3 for sett gir trekking av personer fra henholdsvis sett I, II og III. Kode 4 gir trekking fra settene I og II, kode 5 fra sett I og III og kode 6 fra sett II og III.

Ved hjelp av feltene "Fødselsår", "Ekt.status" og "Kjønn" kan en selektere ut spesielle grupper for alder, kjønn og ekteskapelig status.

Gyldige koder for husholdning/person er henholdsvis H og P. Står det H i feltet, trekkes familieenheter ved at hovedpersoner (personkode 1) trekkes. Dette feltet skal alltid fylles ut. Det samme gjelder for undersøkelsesnummer. Hvis utvalget skal merkes, settes 1 i feltet "Merking/ikke merking".

Lovlige koder for trekkeettheten er 0010-9999. Programmet kan trekke ut maksimum alle personer som oppfyller seleksjonskravene (kol. 17 - 20 = 0010), og minimum hver 999,9-ende person (kol. 17 - 20 = 9999), dvs. gjennomsnittlig hver 999,9-ende person.

Programmet kan trekke klynger. Hvis pos. 21=5, vil en familie som er trukket i henhold til pos. 17-20, "dra med seg" de fire neste familiene som tilfredstiller seleksjonskriteriene. I personundersøkelser vil i dette tilfellet den første personen i de fire påfølgende familier (dvs. hovedpersonen) som oppfyller seleksjonskravene, også bli trukket.

Vitsen med å trekke klyngeutvalg er å få samlet 10-ene bedre slik at en kan få redusert feltkostnadene. Skal en oppnå dette, må 10%-registeret omsorteres slik at personer som bor på samme sted, blir liggende etter hverandre i registeret.

I Sæbø (1976) er beskrevet detaljer omkring utfyllingen av parameterkortene som ikke er tatt med her.

3.3. Problemer ved å trekke utvalg fra et register

Problemene ved å trekke utvalg fra et register skyldes vesentlig to forhold:

- 1) De enheter som registeret inneholder, person- og familieenhet, faller ikke alltid sammen med de enheter vi ønsker å trekke.

2) Registeret er aldri helt à jour. Spesielt oppstår problemer pga. flyttinger og dødsfall.

Som nevnt i avsnitt 2.4 må en ved husholdningsundersøkelser trekke utvalg av familieenheter, og så la intervjuerne supplere dette med eventuelle andre husholdningsmedlemmer. En slik framgangsmåte gir husholdninger som inneholder flere familieenheter større trekkesannsynlighet enn andre. En husholdning som f.eks. består av to voksne ugifte personer og altså utgjøres av to familieenheter, har dobbelt så stor trekkesannsynlighet som en husholdning bestående av én familiekjerne.

I tillegg til at den ovenfor nevnte framgangsmåten gir feil trekkesannsynligheter ved husholdningsundersøkelser, skaper den også problemer i form av merarbeid både i felten og ved Intervjukontoret ved at utvalget øker.

For å begrense de nevnte ulempene har en innført følgende regel: En uttrukket familieenhet kan høyst trekke med seg 4 andre familieenheter til intervju. Er husholdningen større, skal bare den uttrukne familieenheten intervjues.

I praksis oppstår problemene med store husholdninger helst ved trekking av en pasient på pleie- eller sykehjem. Alle andre pasienter på vedkommende rom skal da intervjues hvis det totale antallet på rommet er 5 eller mindre. På rom med 6 eller flere pasienter skal bare den uttrukne intervjues.

Et problem som oppstår ved alle undersøkelser er at uttrukne enheter er flyttet eller døde. Dette kan skje fordi registeret aldri er helt à jour. Når det gjelder flyttede, er hovedregelen ved husholdningsundersøkelser at en intervjuer den husholdningen som eventuelt er flyttet inn i samme leilighet/hus. Ved personundersøkelser er hovedregelen at en skal intervjuer flyttede på det nye bostedet.

I undersøkelser med roterende utvalg som arbeidskraftundersøkelsene følges vanlige regler ved første gangs intervju. Folk som flytter i løpet av den perioden de er med i en slik undersøkelse, skal imidlertid følges opp og intervjues på ny adresse.

I Haldorsen (1982) og Thomsen (1977) er det mer om problemene ved å trekke utvalg fra et register, og om mulige tiltak for å håndtere disse problemene. Når det gjelder foreliggende instruksjoner for håndtering av registerfeil, henvises til Haldorsen (1982) og til Brevkurs for intervjuere (brev 3 og 4).

3.4. Ajourføring av utvalgene ved kobling med Personregisteret

Byråets personregister inneholder alle personer som er bosatt i Norge. Det blir ajourført en gang i måneden på grunnlag av de ordinære meldinger om fødsler, flyttinger, vielser osv. Disse meldingene sendes via de lokale folkeregistrene, som kontrollerer meldingene, påfører personnummer, ajourfører sine registre og videresender alle meldingene til Byrådet en gang i uka.

For en nærmere beskrivelse av Personregisteret henvises til Karlsen og Skaug (1968).

Som nevnt i avsnitt 2.3 er Personregisteret på det nåværende tidspunkt ikke egnet som grunnlag for ajourføring av utvalgsregistrene.

Det kan derimot brukes som grunnlag for ajourføring av uttrukne utvalg. Registeret ajourføres den 18. hver måned. Ved ajourføringene blir det alltid "hengt på" en masse jobber. Ved disse ajourføringene kan en også "henge på" utvalg og få fjernet døde og utvandrede fra disse og få ajourført adresser. En kan regne med at det tar ca. en måned fra en leverer inn et utvalg til Systemkontoret til en får det tilbake i ajourført tilstand.

Ved opp-prioritering av dette arbeidet kunne en få det til å gå noe raskere. Likevel ville slike ajourføringer medføre at utvalgene måtte trekkes langt tidligere i forhold til undersøkelsesperioden enn hva som er nødvendig nå.

Hvis en etter å ha trukket et utvalg ajourfører de opplysningene som er mulig ved kobling med Personregisteret, må en ta en rekke avgjørelser for hvordan en skal bruke den nye informasjonen. En må f.eks. ta stilling til hva en skal gjøre med IO-er som er flyttet, og hva en ved husholdningsundersøkelser skal gjøre med husholdninger som er gått i oppløsning. For en nærmere utredning av disse problemene, henvises til Haldorsen (1982).

Nytten av å koble utvalgene med Personregisteret ligger først og fremst i det at en kan få "luket bort" døde fra utvalgene og at en kan spare intervjuerne for bomturer til fraflyttede hus og leiligheter. Ellers løser ikke ajourføring av utvalgene alle problemer i forbindelse med flytting. Personer som f.eks.

har flyttet inn i et utvalgsområde fra et ikke-utvalgsområde etter siste ajourføring av utvalgsregisteret, vil fortsatt ha trekkesannsynlighet 0. Hvis en velger i størst mulig grad å følge opp uttrukne IO-er som har flyttet, vil dette føre til en del merarbeid på Intervjukontoret ved at arbeidet med å fordele flyttere på intervjuere må foretas manuelt. Videre vil eventuell oppfølging av IO-er som har flyttet fra et utvalgsområde til et ikke-utvalgsområde øke feltkostnadene en del.

Når det gjelder kostnadene ved ajourføringer av utvalg, er ikke de særlig store siden Personregisteret likevel gjennomløpes en gang i måneden.

4. ENDRING AV FOLKETALLENE I UTVALGSOMRADENE

I dette avsnittet skal vi se på hvilke metodiske konsekvenser det har at folketallene i utvalgsområdene endrer seg og at en har delt opp utvalgsområdene i sett.

Som nevnt i avsnitt 2.1 er Byråets utvalgsplan en to-trinns plan, der en i første trinn har trukket utvalgsområder. Disse utvalgsområdene er blitt brukt siden 1975. Før en foretok trekkingen på første trinn, ble utvalgsområdene inndelt i strata. Deretter trakk en ett utvalgsområde fra hvert stratum med trekkesannsynligheter proporsjonale med antall personer i utvalgsområdene.

På annet trinn trekkes personer eller husholdninger fra utvalgsområdene på en slik måte at utvalgene blir selvveiende. Med et selvveiende utvalg menes at alle personer i landet har den samme sannsynlighet for å bli trukket ut når en tenker seg at trekkingen på første trinn ikke er foretatt. For at dette skal være oppfylt må produktet av trekkesannsynligheten på 1. trinn og på 2. trinn være konstant i alle utvalgsområder.

Når utvalgene er selvveiende og trekkesannsynlighetene på 1. trinn er proporsjonale med antall personer i utvalgsområdene, gjelder følgende:

- i) Antall personer som skal trekkes fra et utvalgsområde, er proporsjonalt med antall personer i det stratumet som utvalgsområdet tilhører, og er uavhengig av utfallet av trekkingen på 1. trinn. Dette medfører at hvert stratum blir helt riktig representert med hensyn til andel enheter i utvalgene.
- ii) Vanlige uveiede estimatorer er forventningsrette.
- iii) For mange undersøkelsesvariable vil utvalgsvariansen være nesten like så liten som ved optimal allokering.

Siden 1975 har en brukt de samme trekkesannsynlighetene på annet trinn når en trekker 10%-registeret. Det betyr at utvalgene fortsatt er selvveiende, som igjen sikrer at vanlig uveiede estimatorer er forventningsrette. Derimot har en ikke lenger noen garanti for at trekkesannsynlighetene på 1. trinn er proporsjonale med folketallene i utvalgsområdene, siden folketallene har endret seg siden trekkingen på 1. trinn. At ikke dette er oppfylt, medfører at i) ikke er oppfylt og at variansen i mange tilfeller blir noe større enn ved proporsjonal allokering.

Etter å ha sett nærmere på endringen av folketallene har det imidlertid vist seg at trekkesannsynlighetene på 1. trinn fortsatt er nær ved å være proporsjonale med folketallene i utvalgsområdene i de fleste strataene, slik at i), ii) og iii) er nær ved å være oppfylt ved større undersøkelser.

Noe som i denne forbindelse har større betydning enn at folketallene har endret seg siden trekkingen på 1. trinn, er at en ved mindre undersøkelser bruker ett sett av "tredjedelsområder". Ved oppdelingen av utvalgsområdene var det en forutsetning at folketallene i de tre delene av et utvalgsområde skulle være omtrent like.

Med dette oppfylt ville hvert stratum ha blitt helt riktig representert med hensyn til andel enheter i utvalgene.

Men ved oppdelingen av utvalgsområdene måtte en også ta hensyn til at arbeidsmengden og reisetid for intervjuerne ble jevnt fordelt. Dette siste hensynet har ført til at i noen utvalgsområder varierer folketallet mye fra tredjedelsområde til tredjedelsområde.

Et eksempel på et utvalgsområde der variasjonen er stor, er utvalgsområdet bestående av de tre øyene Øygarden, Fedje og Solund. Her utgjør, naturlig nok, hver øy et tredjedelsområde. På

Øygarden bor det ca. 3 ganger så mange mennesker som på Fedje. Siden trekkesannsynligheten på 2. trinn er den samme om en bruker det ene eller andre settet, betyr dette at andelen personer i utvalgene fra de tre øyene er tre ganger så stor når en bruker sett I av tredjedelsområder som når en bruker sett III.

I 6 av utvalgsområdene bor det 2-3 ganger så mange mennesker i et tredjedelsområde som i et annet. Dette har vært tilfellet helt siden utvalgsplanen var ny. For disse 6 utvalgsområdene vil det altså være slik at andelen personer i utvalgene fra hvert utvalgsområde vil variere med en faktor på 2-3 ettersom en bruker sett I, sett II eller sett III av tredjedelsområder. Dette fører til at variansen øker. Det bør understrekes at utvalgene likevel er selvveiende, slik at vanlige uveide estimatorer er forventningsrette.

5. REFERANSER

- Edholm, A.: "Notat om merking av uttrukne personer/familier i Intervjuavdelingens utvalsregister". Stensil AEd/AnA, 13/6-79, Statistisk Sentralbyrå.
- Gåsemyr, S.: "Innføring av familienummersystemet-nye meldingsrutiner fra 1/11-1970" SG/LHT, 25/8-1970, Rundskriv nr. 24, Statistisk Sentralbyrå.
- Haldorsen, T.: "Handtering av registerfeil, føring av frafall og avgang". 1982.
- Karlsen, K. og Skaug, H.: "Statistisk Sentralbyrås sentrale registre". Artikler fra Statistisk Sentralbyrå nr. 22. (Statistisk Sentralbyrå, 1968).
- Stenbekken, E.: "Notat angående rutiner ved trekking og oppretting av ulike løpende undersøkelser ved Gruppe for metoder". Stensil EyS/KWa. 28/6-77. Statistisk Sentralbyrå.
- Sæbø, H.V.: "Utvalsregistre og rutiner for trekking og klargjøring av utvalg ved Underavdelingen for intervjuundersøkelser". Arbeidsnotater, IO 76/22, Statistisk Sentralbyrå.
- Thomsen, I. og Rideng, A.: "Oversikt over arbeidet med ny utvalgsplan". Stensil ITh/ARi/GHu, 21/5-74, Statistisk Sentralbyrå.
- Thomsen, I.: "Prinsipper og metoder for Statistisk Sentralbyrås utvalsundersøkelser". Samfunnsøk. studier nr. 33, Statistisk Sentralbyrå -77.
- Thomsen, I.: "Metodegruppas ansvar ved planlegging og trekking av utvalg". Stensil ITh/VER, 11/4-80, Statistisk Sentralbyrå.
- AAberge, R.: "Metode for merking av uttrekte personar/familiar i 10%-registeret". 1982.

METODE FOR MERKING AV UTTREKTE PERSONAR/FAMILIAR
I 10%-REGISTERET

Av

Rolf Aaberge

INNHold

	Side
1. Innleiing	12
2. Ny merkemetode	13
3. Bevaring og fjerning av merker	14
4. Vurdering av gammel og ny merkemetode på under- søkingane frå 1976 til 1981	16
Appendiks. Formell grunngiving av trekkesannsyna for umerka personar til 10%-registeret	24

1. Innleiing

Sidan det er uheldig å trekke ut same person eller familie til deltaking i fleire etterfølgjande undersøkingar, innførde Intervjuavdelinga ein metode for merking som skulle løyse eller iallfall redusere dette problemet. Metoden virkar slik at IO som er merka i 10%-registeret blir ute- stengde for deltaking i intervjuundersøkingar i minst 1 år framover. Regelen er at alle uttrekte IO skal merkast. Ei viktig innvending mot denne metoden er at vi strengt tatt berre kan tillate merking når vi trekker generelle person- eller familieutval. Slike utval førekjem svært sjeldan. I praksis har ein ofte innført merking viss den aktuelle populasjonen berre skiljer seg frå den generelle populasjonen ved at dei yngste og dei eldste aldersgruppene fell utafør. Eksempel på slike undersøkingar er AKU (15-74 år) og Tidsnyttingsundersøkinga (16-74 år). Men merking av IO i undersøkingar som AKU og Tidsnytting, fører til at vi ender opp med eit skeivt register. Dei yngste og dei eldste aldersgruppene blir derfor overrepresentert blant einingane med trekkesannsyn i registeret, og dermed også i seinare generelle person- eller familieutval. I tillegg til undersøkingar av typen AKU blir det gjennomført mange undersøkingar der utvala blir trekte frå populasjonar med stor aldersavgrensing eller eventuelt geografisk, kjønnsmessig, osv. avgrensing. Dei trekte einingane til slike undersøkingar blir ikkje merkte i 10%-registeret. Dette har ført til at enkelte intervjuarar har opplevd å måtte oppsøke same person to ganger i løpet av kort tid. Det er fire ulike situasjonar som kan føre til at ein bestemt person blir trekt ut for å delta i to undersøkingar like etter kvarandre.

- (i) Personen har delteke i ei undersøking som ikkje blir merka og blir like etterpå trekt ut for deltaking i ei ny undersøking.
- (ii) Etter å ha vore IO i ei personundersøking, blir personen trekt ut for å delta i ei hushaldsundersøking. Dette kan skje viss IO ikkje er "hovedperson".
- (iii) Personen bur på same adresse som det uttrekte hushaldet (familiekjernen) og skal derfor vere med i undersøkinga. Slike personar blir ikkje merka i registeret og kan derfor bli trekt ut for deltaking i ei ny undersøking like etterpå.
- (iv) IO blei trekt ut for å delta i ei ny undersøking på same tid som IO blei intervjuet i AKU. Dette kan skje fordi IO i AKU blir intervjuet

fire ganger og eit AKU-utval blir brukt i to år. I løpet av undersøkningsperioden vil det vere ein del IO som både fell ut og kjem inn att i 10%-registeret, slik at dei vil vere umerka før dei blir intervjuja siste gang i AKU.

Vi har ovafor peika på to svært uheldige sider ved den eksisterande merkemetoden:

- (1) Merking av ikkje-generelle utval fører til skeivt register
- (2) Same person eller familie blir trekt ut til fleire etterfølgjande undersøkingar.

Vi skal no leggje fram forslag til ein ny merkemetode og diskutere denne i forhold til dei skisserte problema ((1) og (2)) og vidare vurdere aktualiteten av denne metoden på intervjuundersøkingane frå åra 1976, 1977, . . . , 1981.

2. Ny merkemetode

Vårt mål er å kunne innføre ein merkemetode som tillet merking av flest moglege (eventuelt alle) utval og samtidig sikrer at 10%-registeret ikkje blir skeivt. Følgjande merkemetode vil langt på veg tilfredsstillе desse krava:

- (I) Generelle person- og familieutval blir merkte som i dag
- (II) Når det gjeld dei spesielle utvala (dei aller fleste) må vi ta utgangspunkt i definisjonen av den aktuelle populasjonen. For å kunne merke eit utval frå den aktuelle populasjonen (f.eks. Aleneforeldres tidsbruk (1981)) må vi merke relativt like mange einingar frå resten av den generelle populasjonen i 10%-registeret.

For Aleneforeldres tidsbruk (1981) vil (II) medføre at vi måtte merke til saman 40 000 einingar i 10%-registeret. I praksis kunne vi trekke 40 000 einingar frå den generelle populasjonen og dermed forvente å få ca. 700 personar frå den spesielle populasjonen (einslege over 18 år med barn under 15 år). Poenget er altså at vi alltid må ta utgangspunkt i den generelle populasjonen. Når vi så skal trekke eit utval frå ein avgrensa populasjon, må vi finne ut kor stor del av 10%-registeret som blir representert av den avgrensa populasjonen. Viss denne delen er p og utvalet frå den avgrensa populasjonen skal vere på n einingar, må vi merke i alt n/p einingar. Av kostnadmessige årsaker kan det bli aktuelt å erstatte p med p' , der p' er den delen som den spesielle populasjonen utgjer av den generelle populasjonen. Vi har alltid at $p' \approx p$.

Denne merkemethoden vil løyse problem (1). Men som den oppvakte lesar allereie har notert seg, så vil vi kunne få vanskar med å utslette problem (2) fullstendig. Ein av grunnane til dette er at nokre fåe utval med stor eller moderat n og liten p (som f.eks. Aleneforeldres tidsbruk 1981) kan medføre at alle einingane i 10%-registeret blir merka. Følgjeleg må vi framleis rekne med å utelate enkelte undersøkingar frå merking. Den nye merkemethoden vil likevel føre til at langt fleire utval enn i dag vil bli merka, sidan hovudårsaka ((i)) til problem (2) vil bli delvis fjerna. Delproblema (ii) og (iii) kan vi ikkje gjere noko med, fordi merking av desse personane ville medføre at registeret blir skeivt. På den andre sida veit vi at desse to situasjonane vil førekomme forholdsvis sjeldan.

3. Bevaring og fjerning av merker

10%-registeret inneheld ca. 400 000 personar. Kvart år blir det trekt eit nytt 10%-register. Dette registeret blir kopla med 10%-registeret frå året før. Viss ein person i det gamle registeret også er med i det nye registeret, blir merket overført til det nye registeret. Det er kun i 10%-registeret at det blir merka; merka blir ikkje overførde til utvalsregisteret. Dette vil seie at ein merka person vil halde fram med å vere merka så lenge denne personen er med i 10%-registeret, og at fjerning av merker skjer ved utskiftingar i 10%-registeret.

I utvalsområder som er trekt med sannsyn lik 1 eller nær 1 (dei store byane), er det store utskiftingar i 10%-registeret frå år til år. Merka personar frå desse utvalsområda blir kvitt merka forholdsvis raskt. I utvalsområder som er trekt med småe sannsyn er det lite utskiftingar frå år til år. For slike områder har derfor merka ein tendens til å bli ståande. I utvalsområdet Bjarkøy og Ibestad er alle innbyggjarane med i 10%-registeret. Den delen som er merka i dette området blir derfor større frå til år. For tida er ca. 50 % av alle over 15 år merka i dette området. Konsekvensen av denne merkemethoden er at registeret er blitt skeivt. Småe utvalsområder blir representert med for fåe personar og store med for mange personar i registeret. Dette forholdet har også blitt avspegla i dei utvala som vi har trekt frå 10%-registeret. Problemet er blitt alvorlegare med tida. I siste halvdel av 1981 er f.eks. alle uttrekte utval berre representert med om lag halvparten av dei personane som skulle ha vore med frå Bjarkøy og Ibestad .

Frå 10%-registeret kan vi trekke to typar einingar; person eller familiekjerne. Familiekjernen blir trekt via hovudpersonen. Hovudpersonane utgjer derfor ei spesiell gruppe av

totalpopulasjonen. For å unngå at 10%-registeret etter ei tids bruk vil inneholde for få umerka hovudpersonar må vi innføre to typar merking. Hovudmerkinga blir den same som i dag (feltnr. 20 på recorden) og skal gjelde alle personar som er trekt ut for å delta i ei undersøking, enten trekkeininga er person eller familie. I tillegg vil vi krevje spesiell merking av uttrekte hovudpersonar ved trekking av familieutval. Til dette skal vi nytte feltnr. 18 (posisjon 99-100 for landet utanom Oslo og 110-111 for Oslo). Det er denne merkinga som skal gjelde når vi trekker familieutval. Det vil seie at ein hovudperson kan bli trekt ut sjølv om han har delteke i ei personundersøking (merke i feltnr. 20). Det motsette kan ikkje skje. Alle personar som har delteke i ei familieundersøking (også hovudpersonane) skal ha tildelt merke etter feltnr. 20 på recorden.

Ved den nye metoden skal personane (både for person- og familieutval) merkast med avslutningstidspunkt (månad og år) for intervjuperioden. For personutval skal merket stå i feltnr. 20. For familieutval skal alle dei uttrekte personane merkast i feltnr. 20 og i tillegg skal alle uttrekte hovudpersonar merkast i feltnr. 18. For utval til AKU og Forbruk må vi merke med tidspunkt for siste intervjuperiode.

Merka personar skal alltid vere med i 10%-registeret. Etter at personane har vore merka i to år, skal merka fjernast. Fjerning av merker (både etter feltnr. 18 og 20) skal berre skje i samanheng med ajourføringa av utvalsregisteret; to ganger om året i motsetnad til ein gong som i dag.

Fjerninga av merker vil bli parameterstyrt. Vi kan derfor forkorte eller forlenge merkeperioden for ei undersøking, dersom spesielle forhold skulle tilseie at dette er nødvendig.

I samband med ajourføring av utvalsregisteret, blir det i dag trekt nytt 10%-register. Vi opprettar 10%-registeret ved å trekke familiekjerner (hovudpersonar) frå utvalsregisteret. Vi skal framleis gjere dette, men framgangsmåten blir noko endra. 10%-registeret skal først rensast for uønskte merker (minst 2 år gamle) og deretter koplast med det ajourførte utvalsregisteret. På den måten får vi opplysningar om talet på aktuelle merka og umerka personar i kvart utvalsområde. Desse opplysningane skal vi bruke til å fastsetje trekketettheitene i dei ulike utvalsområda.

La M = talet på personar i utvalsområdet (summen av merka og umerka personar) etter ajourføringa, m = talet på merka personar i utvalsområdet etter ajourføringa og la q_1 = den opphavlege trekketettheiten i utvalsområdet (gitt i den gamle trekketettheitskatalogen). m er lik talet på dei personane som har merke i feltnr. 20.

Da blir

$$q_2 = \frac{M - m}{\frac{M}{q_1} - m}$$

trekkesannsynligheten vi skal bruke ved oppretting av nytt 10%-register. Ved kvar ny ajourføring (og trekking av nytt 10%-register) vil M og m ha endra seg, medan q_1 er fast. I appendiks gir vi ei formell grunngiving for at personane i 10%-registeret har fått rette trekkesannsyn.

Vi skal nå repetere metoden for merking og metoden for trekking av nye 10%-register:

Merking. Personane i utvalet (alltid generelt person- eller familieutval) skal merkast med tidspunkt (månad og år) for siste del av intervju-perioden (feltnr. 20). Vi skal bruke kodane 0, 1, 2, ..., 9, for kvart av åra i eit tiår, dvs. 0 for 1980, 1990, osv., 1 for 1981, 1991, osv.

Månadene skal kodast med bokstavar; a for januar, b for februar, c for mars, d for april, e for mai, f for juni, g for juli, h for august, i for september, j for oktober, k for november og l for desember. Når vi trekker familieutval, skal hovudpersonane også ha tildelt det same merket i feltnr. 18.

Ein merka person som seinare blir trekt ut i eit familieutval, skal få endra merket til tidspunktet for gjennomføringa av den siste undersøkinga.

Ajourføring og trekking av 10%-register

(a) Det gamle 10%-registeret skal rensast for merker som det er ønske om å fjerna. Dette vil stort sett dreia seg om dei undersøkingane som er minst 2 år gamle.

(b) Det rensa 10%-registeret skal koplant med det ajourførte utvalsregisteret. Dei merka personane (m) som framleis er med i utvalsregisteret skal vere med i 10%-registeret som merka personar.

(c) Ved å bruke den opphavlege trekkekatalogen og opplysningane om merka og umerka personar i det ajourførte utvalsregisteret, skal vi opprette ein ny trekkesannsynskatalog. Denne får vi ved å rekne ut q_2 -ane.

(d) Den nye trekkesannsynskatalogen skal vi bruke til å trekke eit utval av familiekjerner (via hovudpersonar) frå den totale massen i utvalsregisteret. Dei uttrekte umerka personane i dette utvalet skal saman med alle merka personar utgjere det nye 10%-registeret.

4. Vurdering av gammel og ny merkemetode på undersøkingane frå 1976 til 1981

For kvart av åra 1976, 1977, ..., 1981 skal vi gi ei oversikt over kva slags undersøkingar som kunne blitt merka ved den nye metoden.

Undersøking	Populasjon	Utvalgsstorleik	Merking av utvalet i 10%-registeret	Talet på personar i 10%-reg. som ville bli merka ved gammel prosedyre	Talet på personar i 10%-reg. som ville bli merka ved ny prosedyre
Forbruk	Alle hushald	1 500	Nei	4 100	4 100
Forbruk (tillegg)	Alle hushald som deltok i Forbruk 75	300	"	900	900
Husleige	Alle hushald	6 000	Ja	16 200	16 200
NRK I	15-79 år	1 400	Nei	1 400	1 900
NRK II	15-79 år	1 200	"	1 200	1 700
Fruktbarhet (prøve)	Kvinner (18-49 år)	200	"	220	1 200
AKU	Hushald (16-74 år)	6 000	Ja	16 200	22 500

1977

Undersøking	Populasjon	Utvalgsstorleik	Merking av utvalet i 10%-registeret	Talet på personar i 10%-reg. som ville bli merka ved gammel prosedyre	Talet på personar i 10%-reg. som ville bli merka ved ny prosedyre
Fruktbarhet	Kvinner (18-44 år)	4 900	Nei	4 900	26 500
Transport (Oslo)	17-75 år	2 400	"	2 400	25 600
Transport (landet)	17-75 år	1 600	"	1 600	2 300
Inntekt og formue	Alle hushald	4 700	"	12 700	12 700
Stortingsval	20-79 år	2 600	"	2 600	3 900
Varige forbruks- goder	16-80 år	5 000	Ja	5 000	18 300
AKU	Hushald (16-74 år)	6 000	"	16 200	22 500
Husleige	Alle hushald	5 500	"	14 900	14 900
NRK	15-79 år	2 800	Ja (landet) Nei (Oslo)	2 800	3 800
Forbruk	Alle hushald	1 200	Ja	3 200	3 200

Undersøking	Populasjon	Utvalgsstorleik	Merking av utvalet i 10%-registeret	Talet på personar i 10%-reg, som ville bli merka ved gammel prosedyre	Talet på personar i 10%-reg. som ville bli merka ved ny prosedyre
Forbruk	Alle hushald	1 300	Ja	3 500	3 500
Husleige	Hushald (16-85 år)	6 000	"	16 200	21 000
NRK-prøve	3-8 år	50	Nei	50	600
Ferie	15-75 år	3 000	Nei (landet) Ja (Oslo)	3 000	4 100
NRK	3-14 år	1 500	Nei	1 500	8 200
19 Flyttekjede (Oslo)	Alle hushald i Oslo	500	"	1 400	12 800
Barnehage (Oslo)	Fam. m/barn 0-6 år (Oslo)	1 400	"	3 800	229 000
Tidsbruk (prøve)	20-74 år	300	"	300	500
AKU	Hushald (16-74 år)	6 000	Ja	16 200	22 500
Barneomsorg (prøve)	20-74 år	300	Nei	300	500

1979

Undersøking	Populasjon	Utvalgsstorleik	Merking av utvalet i 10%-registeret	Talet på personar i 10%-reg. som ville bli merka ved gammel prosedyre	Talet på personar i 10%-reg. som ville bli merka ved ny prosedyre
Forbruk	Alle hushald	2 500	Ja	6 800	6 800
Husleige	Hushald (16-85)	6 000	"	16 200	21 000
Om forholdet mellom den enkelte og myndighetene	19-75 år	1 500	Nei	1 500	2 300
Levekår (prøve)	17-66 år	600	"	600	1 000
Kvinneres forhold til arb.mark. (prøve)	20-67 år	300	"	300	1 100
Transport	16-79 år	3 000	"	3 000	4 100
Helse	16-79 år	2 000	"	2 000	2 800
Inform. tilgang i Østfold	16-74 år	1 800	"	1 800	45 500
Ferie	15-75 år	3 000	Ja	3 000	4 100
Tidsbruk (prøve)	20-74 år	600	Nei	600	1 000
AKU	15-74 år	6 000	Ja	16 200	22 500

Undersøking	Populasjon	Utvalgsstorleik	Merking av utvalet i 10%-registeret	Talet på personar i 10%-reg. som ville bli merka ved gammel prosedyre	Talet på personar i 10%-reg. som ville bli merka ved ny prosedyre
Levekår	16-79 år	5 000	Nei	5 000	6 800
Kvalitetskontroll	Alle hushald	4 000	Ja	10 800	10 800
Energi	Alle hushald	3 000	Nei	8 100	8 100
Likestilling	19-80 år	1 500	"	1 500	2 200
Ungdom	17-19 år	10 000	"	10 000	220 000
Kvinner forholds- til arbeidsmark.	Kvinner (20-67 år)	3 000	"	3 000	10 800
Kredittopplysning (prøve)	Alle hushald	200	"	600	600
NRK	10-80 år	4 600	"	4 600	5 500
Husleige	16-85 år	6 000	Ja	16 200	21 000
Røykevane	14-73 år	3 000	"	3 000	4 100
U-hjelp	14-73 år	3 000	"	3 000	4 100
Forbruk	Alle hushald	2 000	"	5 400	5 400
AKU	Hushald (14-73 år)	6 000	"	16 200	22 500

1981

Undersøking	Populasjon	Utvalgsstorleik	Merking av utvalet i 10%-registeret	Talet på personar i 10%-reg. som ville bli merka ved gammel prosedyre	Talet på personar i 10%-reg. som ville bli merka ved ny prosedyre
Tidsnytting	16-74 år	5 200	Ja	5 200	7 400
Aleineforeldres tidsbruk	Einslege over 18 år m/barn	700	Nei	700	40 000
Husleige	Hushald (16-85 år)	6 000	Ja	16 200	21 000
Forbruk	Hushald (under 85 år)	2 200	"	5 900	6 000
Demokrati	18-65 år	1 650	"	1 650	2 800
Boforhold	Alle hushald	3 000	"	8 100	8 100
22 Ungdom (prøve)	22-24 år	150	Nei	150	3 300
Ungdom	10 med i 1975 (22-24 år)	3 200	"	3 200	70 400
AKU	Hushald (14-73)	6 000	Ja	16 200	22 500
Val	18-79 år	2 200	Nei	2 200	3 200
Boforhold (Oslo)	Alle hushald (Oslo)	600	Ja	1 600	14 500
Boforhold (Akershus)	Alle hushald (Akershus)	600	Nei	1 600	17 700

For åra 1976 og 1977 ser vi at respektive 2 og 5 undersøkingar blir merka. I 1976 hadde ei av dei merka undersøkingane aldersavgrensing (AKU) medan 3 av dei 5 merka undersøkingane i 1977 hadde aldersavgrensing. Dette førde til at yngre og eldre personar fekk større sannsyn for å bli trekt ut i ei tilfeldig trekking frå 10%-registeret på eit seinare tidspunkt; registeret var blitt skeivt. Det same forholdet gjentek seg for åra 1978-1981.

Dersom vi hadde merka alle undersøkingane for åra 1976 og 1977 ved den nye prosedyren, finn vi at til saman 163 800 personar ville blitt merka. Dette resultatet får vi ved å summere tala i kolonne 6 i tabellane for åra 1976 og 1977. Dei tilsvarande tala for 1978 og 1979 er respektive 302 700 og 112 200. For begge desse åra var det ei undersøking som burde unnatakast merking, Barnehageundersøkinga (Oslo) i 1978 og Undersøking om informasjonstilgang (Østfold) i 1979. Desse to undersøkingane ville ha krevd merking av til saman 274 500 personar.

Om vi ser bort frå dei to sistnemnde undersøkingane ville den nye prosedyren medført at 140 400 personar hadde blitt merka for åra 1978 og 1979. I 1980 og 1981 kunne vi ha merka alle utval unnateke Ungdomsundersøkingane og Aleneforeldres tidsbruk. I så fall ville til saman 205 100 personar blitt merka.

Konklusjon: 28 av i alt 63 undersøkingar for åra 1976-1981 er blitt merka. 18 av dei merka undersøkingane har ikkje tilfredsstilt kravet om å vere generelle utval; dei har derfor gitt skeive register. Ved den nye merke-metoden kunne vi ha merka 58 av dei 63 undersøkingane utan at registeret blei skeivt.

Den nye metoden for bevaring og fjerning av merker vil i motsetnad til den gamle metoden unngå å skape skeive 10%-register.

I enkelte undersøkingar bruker vi berre eitt eller to av dei tre setta frå kvart utvalsområde. Vi må da trekke eit generelt utval (frå alle setta) og gjennomføre merking for alle dei uttrekte personane. Dei (det) ubrukne setta (settet) kan vi ta vare på for seinare bruk. Det ideelle er å planleggje (på førehand) ei undersøking basert på dei (det) ubrukne setta (settet).

I praksis må vi basere oss på at om lag halvparten av personane i 10%-registeret vil vere merka til ei kvar tid. For å unngå alvorlege overskridingar, må Metodegruppa alltid vurdere om undersøkingar som krev meir enn 25 000 merker skal merkast eller ikkje.

Appendiks. Formell grunngiving av trekkesannsyna for umerka personar til 10%-registeret

Vi føreset at vi har tilgjengeleg eit ajourført utvalsregister. Dette er sett saman av både merka og umerka personar. Alle dei merka personane skal vere med i det nye 10%-registeret. Frå den totale massen i det ajourførte utvalsregisteret skal vi trekke eit utval av familiekjerner. Dei uttrekte umerka personane i dette utvalet skal saman med alle merka personar utgjere det nye 10%-registeret.

Gitt at ein person frå eit bestemt utvalsområde er med i det ajourførte utvalsregisteret, skal vi etablere eit uttrykk for sannsynet (p_1) for at denne personen er med i det nye 10%-registeret. Frå dette uttrykket skal vi finne sannsynet for umerka personar til 10%-registeret.

La A = personen er med i 10%-registeret, B = personen er merka, $p_1 = \Pr(A)$ og $p_2 = \Pr(A|\bar{B})$. Da har vi at

$$p_1 = \Pr(A) = \Pr(A \cap B) + \Pr(A \cap \bar{B}) = \Pr(A|B) \Pr(B) + \Pr(A|\bar{B}) \Pr(\bar{B}).$$

Vi føreset at alle personar i utvalsområdet har hatt like stort sannsyn for å bli merka. Dermed får vi at

$$\Pr(B) = \frac{m}{M}.$$

Vidare finn vi at -

$$\Pr(A|B) = 1 \quad \text{og} \quad \Pr(\bar{B}) = 1 - \frac{m}{M}.$$

Dette gir

$$p_1 = \frac{m}{M} + p_2 \left(1 - \frac{m}{M}\right)$$

eller

$$p_2 = \frac{M p_1 - m}{M - m}.$$

p_1 er trekkesannsynet på trinn 2 for den generelle utvalsplanen i Byrået. p_1 er kjent for kvart utvalsområde og finst som $q_1 = 1/p_1$ i trekketettheitskatalogen. Vi kan derfor rekne ut verdiane (q_2) til den nye trekkekatalogen ved

$$q_2 = \frac{M - m}{\frac{M}{q_1} - m}.$$

HANDTERING AV REGISTERFEIL, FØRING AV FRAFALL OG AVGANG

Av

Tor Haldorsen

INNHold

	Side
-	
Innledning	26
Begreper	26
Tidspunkter	27
Momenter ved planleggingen	28
Vanlige typer utvalg i Byrådet	30
Gjennomføring av standardundersøkelser	31
Ajourføring av utvalg	32
Regnskap over frafall og avgang	33

Innledning

De fleste utvalg vi bruker er trukket fra registre. Disse kan ha feil. Enheter som skulle vært med, er utelatt og/eller enheter som ikke skulle vært med, er tatt med. Opplysningene om enkelte enheter kan være gale.

De fleste registre ajourføres til bestemte tider. Hvis vi vil ha situasjonen andre tider, vil det fortone seg som det er feil i registeret.

I notatet vil vi beskrive hva vi vanligvis betrakter som registerfeil og hvordan vi behandler disse. Vi skal også komme inn på frafall. Først vil vi innføre en del nyttige begreper.

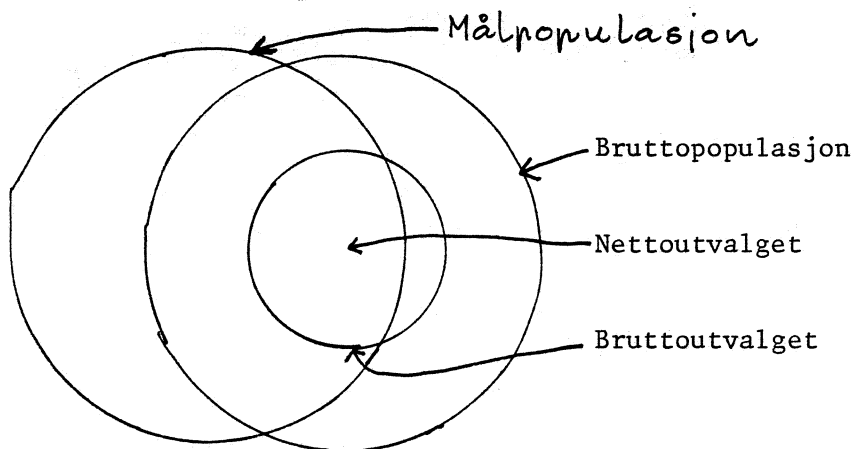
Begreper

For hver undersøkelse vil det foreligge en målpopulasjon. Det er den mengde enheter vi vil uttale oss om. Enhetene kan være personer, familier, husholdninger, ferieturer, syketilfelle osv.

Av og til vil enhetene i registeret som vi trekker fra, være av en annen type enn de i målpopulasjonen. Vi må f.eks. gå om personer for å få tak i syketilfelle. For å understreke dette snakker vi om trekkeenheter og analysenheter.

Som regel vil dessverre ikke registeret (eller de enheter som indirekte gis av registeret når vi har spesielle trekkeenheter), være identisk med målpopulasjonen. Enheter som på grunnlag av opplysninger i registeret, ikke tilhører målpopulasjonen kan vi ekskludere forbindelse med trekkingen, så de byr ikke på problemer. F.eks. kan vi rense registeret for uaktuelle aldersgrupper hvis alder er kjennetegn i registeret. Enheter som direkte eller indirekte gis av registeret etter eventuell opprenskning utgjør bruttopopulasjonen. Generelt må vi regne med at enkelte i målpopulasjonen ikke er med i bruttopopulasjonen. Disse utgjør en underdekning. På den annen side vil også bruttopopulasjonen kunne inneholde enkelte som ikke er med i målpopulasjonen. Disse representerer en overdekning. Fra bruttopopulasjonen trekkes et utvalg. Det kalles bruttoutvalget.

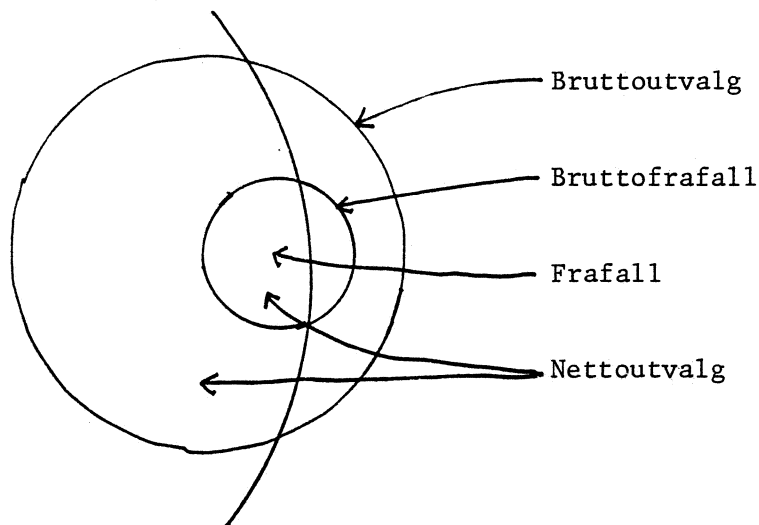
Prinsipielt vil bruttoutvalget kunne deles i to, de som tilhører målpopulasjonen og de som ikke gjør det. Den første delen kaller vi netto-utvalget. Begrepene kan vi illustrere:



(Piler til linjer navngir alt innenfor den sirkellinjen, piler til sektorer navngir bare sektoren).

Ved en undersøkelse prøver vi å innrette oss slik at vi kan avgjøre hvilke enheter i bruttoutvalget som ikke er med i nettoutvalget. Disse vil vi ha som avgang. Her skaper frafallet spesielle problemer.

Vi definerer bruttofrafall som den del av bruttoutvalget som vi ikke får intervjuet/opplysninger om. Den del av bruttofrafallet som tilhører nettoutvalget, kan kalles nettofrafall eller bare frafall. For å illustrere de siste begrepene forstørrer vi en del av forrige figur.



Tidspunkter

I forbindelse med utvalgsundersøkelser basert på registre er det noen ulike tidspunkter en bør ha klart for seg. For det første er det registeret, i prinsippet skal det være ajour ved tid T_1 , men av ulike årsaker inneholder det noen feil. Videre vil som regel målpopulasjonen referere til et tidspunkt, f.eks. personer over 16 år (registrert), bosatt i Norge 1. november 1980. Dette tidspunkt vil vi kalle T_2 . Som det tredje tidspunkt, T_3 , har vi den dag intervjueren oppsøker IO. Av og til kan noen av tidspunktene falle sammen. Oftest vil de tre tidspunktene være i rekkefølge T_1 før T_2 før T_3 .

Momenter ved planleggingen

Ved en undersøkelse vil de ulike gruppene på Underavdelingen ha bestemte krav og interesser til gjennomføringen. Noen ganger vil interessene være delvis motstridende. Vi skal nevne noen av problemene i tilknytning til de mengder vi har definert.

Viktig er det selvsagt med en presis og praktisk definisjon av målpopulasjonen. For å få til dette, kan det være nødvendig at oppdragsgiver, saksbehandler (Gruppe for analyse og planlegging), Gruppe for metoder og Intervjukontoret (Feltgruppa/Bearbeidingsgruppa) samarbeider om definisjonen. Om nødvendig, må det være saksbehandlerens ansvar å hente og koordinere synspunkter fra de ulike hold. Det er ikke uvanlig at en i definisjonen av målpopulasjonen skjeler til hva som er praktisk mulig. F.eks. vil en gjerne i en helseundersøkelse uttale seg om helse-tilstanden i hele befolkningen, men på grunn av problemer med å intervju personer som er innlagt i helseinstitusjoner nøyer en seg med å definere målpopulasjonen som personer utenfor helseinstitusjoner på et bestemt tidspunkt.

Det er ønskelig at bruttopopulasjonen fullstendig dekker målpopulasjonen. Hvis noen i målpopulasjonen faller utenfor bruttopopulasjonen, representerer de en underdekning i undersøkelsen. Spesielt ille er det når underdekningen er systematisk og en ikke er klar over dette.

Videre vil en at alle i målpopulasjonen skal ha kjent sannsynlighet for å trekkes til utvalget. Dette er en grunnleggende forutsetning for de slutninger en kan trekke på grunnlag av undersøkelsen. Det er også viktig for bruken av materialet at de kjente sannsynlighetene er av en enkel form, f.eks. selvveiende utvalg.

En metode for å oppnå liten underdekning kan være å lage en stor bruttopopulasjon. Men det kan også medføre problemer. Vi får da ofte en stor overdekning. Denne kan bety vesentlig merarbeid for Intervjukontoret (sentralt og/eller i felten) og kan innføre komplikasjoner i trekkesannsynlighetene for målpopulasjonen. Men av og til er målpopulasjonen slik definert at dette ikke er til å unngå. Ved undersøkelsen "Alene foreldres tidsbruk" var målpopulasjonen de som bodde helt alene med barn, det skulle ikke bo andre voksne i husholdningen. En slik masse kunne ikke defineres ved hjelp av registeropplysninger. Det ble derfor trukket et bruttoutvalg av ugifte/før gifte med barn. Disse ble tilskrevet av Feltgruppa og spurt om de oppfylte betingelsene. De som svarte benektende antok man ikke tilhørte målpopulasjonen, og det ble ikke sendt intervjuere til disse.

Ellers er det oftest intervjuerne som må skaffe opplysninger så en kan avgjøre om enhetene er med i nettoutvalgene. Av og til innarbeides det spørsmål om dette i begynnelsen av intervjueskjemaet.

Mål- og bruttopopulasjonen, brutto- og nettoutvalg er "teoretiske" mengder, i noen tilfelle kan vi angi størrelsen, i andre må vi estimere eller gjette.

Hvis målpopulasjonen er alle personer bosatt i Norge, har vi brukbare tall for det, men er målpopulasjonen alle kjevhendte personer bosatt i Norge, må vi skaffe et anslag fra andre kilder eller estimere etter at undersøkelsen er gjennomført.

I en personundersøkelse med basis i ett personregister, vil vi ha tall for bruttopopulasjon og bruttoutvalg hvis intervjuerne ikke får intervju andre enn dem på IO-listene. Men en annen sak er at det i noen få tilfelle kan føre til en relativt stor underdekning.

Som nevnt skaper frafallet spesielle problemer når vi skal avgjøre hvem i bruttoutvalget som er med i nettoutvalget. Vi vil ha få opplysninger om enhetene i frafallet. Vi følger der en "konservativ" praksis, hvis vi ikke positivt vet at bruttofrafallsenheten er utenfor målpopulasjonen, så regnes den med i nettofrafallet.

Generelt kan problemene med de ulike tidspunktene (T_1 , T_2 og T_3) minimeres ved å innrette seg slik at de er så nær hverandre som mulig. Forandringer (flytting, dødsfall, sammensetning i husholdning o.a.) som skjer mellom T_1 og T_2 vil ha samme virkning som rene feil i registeret og kan betraktes som samme type problem. Forandringer som skjer mellom T_2 og T_3 kan skape finurligheter og praktiske problemer, men det kan begrenses ved å gjøre tid mellom T_2 og T_3 så kort som mulig. Ofte vil det være et slingringsmann i valg av T_2 som kan utnyttes og dessuten er det ofte mulig å begrense intervjuperioden. For noen av de IO som volder problemer i vår sammenheng (flytting, skifte av status) vil det være vanskelig for intervjuere å få informasjon om status ved T_2 og i noen tilfelle er det unødvendig å lage regler for intervjuer som skjelner mellom hva han skal gjøre ettersom ting har skjedd for eller etter T_2 .

Et moment med de definisjoner som vi har innført er at bruttopopulasjon er knyttet til trekkingen. For å beholde oversikten må Feltgruppa følge alle i bruttoutvalget. Saksbehandleren er mer opptatt av nettoutvalg og (netto)frafall. For å hjelpe seg bruker Feltgruppa ofte et (brutto-) frafallsskjema som dekker både frafall og avgang.

I disse tilfelle vil det for alle i bruttoutvalget som er sendt intervjuerne, til slutt foreligge et intervjuuskjema eller et frafalls/avgangsskjema. Det siste må inneholde opplysninger om enheten er i eller utenfor målpopulasjonen. Intervjuuskjema for enheter utenfor målpopulasjonen må siles ut ved revisjon/kontroll.

Vanlige typer utvalg i Byrået

Som regel bruker vi et spesielt trekkeregister som er lagd på grunnlag av folkeregisterenes magnetbåndregister. Trekkeregisteret har blitt ajourført én gang i året, men dette skal nå gjøres hvert halvår. Utvalgene trekkes i to trinn. I første trinn trekkes 102 områder, mens i annet trinn trekkes personer/husholdninger fra disse områdene. Områdene som trekkes på første trinn holdes fast og trekkeregisteret omfatter derfor bare disse områdene.

Folkeregisterene har detaljerte regler for hvor folk skal regnes bosatt. Siden vi bruker deres registre, følger vi også deres definisjon av bosted. Begreper som flytting og midlertidig fravær får sin betydning ut fra bostedsreglene som er beskrevet i et hefte utgitt av Sentralkontoret for folkeregistrering, Statistisk Sentralbyrå.

Vi kan skjelne mellom tre ulike typer utvalg; direkte personutvalg, husholdningsutvalg og indirekte personutvalg trukket som "alle" i en husholdning. Definisjonen av husholdning kan variere litt fra undersøkelse til undersøkelse. Men vi forutsetter at det til husholdningen knyttes et bosted.

Direkte personutvalg er det enkleste. Da holder vi fast ved de personer som blir trukket. De utgjør bruttoutvalget og ingen andre personer skal med i undersøkelsen. Ved intervjutidspunktet (T_3) evt. T_2 kan enkelte personer i utvalget bo utenfor utvalgsområdene. Det skyldes registerfeil eller flytting etter T_1 .

For husholdningsutvalg og (indirekte) personutvalg trukket som "alle" i en husholdning er trekkeenheten husholdning og da tenker vi på en annen måte. Vi holder fast ved adressene/bopelene til de husholdningene som blir trukket. Det er husholdninger/personer som bebor de trukne adressene/bopelene på T_3 evt. T_2 som skal med i undersøkelsen. Det betyr at det i bruttoutvalget kan være husholdninger/personer som ikke finnes i trekkeregisteret. Dette kan skyldes feil i registeret eller at husholdningen/personene har flyttet til de trukne adressene etter T_1 . Alle husholdninger/personer som skal med i undersøkelsen vil være bosatt i utvalgsområdene, ved T_3 evt. T_2 , men enkelte kan selvsagt være midlertidig fraværende, p.g.a. arbeid, utdanning m.m.

Gjennomføring av standardundersøkelser

Når vi gjennomfører undersøkelser vil det opptre en del problemtilfeller. Vi skal her gjengi noen av disse og nevne hvordan de skal behandles i følge tenkesettet fra forrige avsnitt.

Personutvalg

Problem	Behandling
1) IO er død	Avgang
2) IO er flyttet	
a) Til samme utvalgsområde	Forfølges
b) Til annet utvalgsområde	"
c) Utenfor utvalgsområdene	Forfølges/fracfall
d) Utlandet	Avgang
e) Uvisst hvorhen	Frafall/forfølges
3) IO endret status så utenfor målpopulasjonen	Avgang
4) Huset revet, nedbrent, "fracflyttet"	Forfølges
5) Finner ikke huset/adressen/leiligheten/personen	Frafall/forfølges

Kommentarer:

- 1) I de fleste tilfelle vil det bety lite å skille mellom før og etter T_2 , derfor har vi "avgang" som generell regel.
- 2) Disse regnes med i utvalget så lenge en ikke vet det motsatte. Som regel vil målpopulasjonen dekke personer bosatt i Norge, derfor er 2d satt som avgang. Feltgruppa vil bestemme hvilke IO intervjueren selv skal forfølge og hvilke som skal oppsøkes av andre intervjuere evt. fra sentralt hold. For de IO som ikke intervjueren forfølger selv, sender han frafalls/avgangsskjema for med de aktuelle opplysninger, eventuelt et spesielt flytteskjema. Aksjonen på 2e) avhenger av hvilke muligheter Feltgruppa har til å etterspore nye adresser.
- 3) Her antar vi også som regel at det er liten vits i å skille mellom før og etter T_2 .
- 4) Intervjuer søker ytterligere opplysninger og behandler dette som punktene 1)-3) (oftest 2)).
- 5) Intervjuer sender frafalls/avgangsskjema med aktuelle opplysninger og Feltgruppa avgjør videre behandling.

Husholdningsutvalg og (indirekte) personutvalg trukket som "alle" personer i husholdningen

For disse typene utvalg holder vi altså fast ved adressene til dem som er trukket og tar med i undersøkelsen de som bor på adressene ved tid T_3 evt. T_2 . På IO-listene vil adressene være definert med adressebetegnelse og navn på de som ifølge registeret bor på adressen.

Problem	Behandling
1) Helt ny/bare deler av uttrukket familie bor på adressen (dvs. hele eller deler av uttrukket familie flyttet, død).	Oppsøke de som bor der nå. Andre blir avgang
2) Huset revet, nedbrent, "fracflyttet"	Avgang
3) Finner ikke huset/adressen/leilighetene	Frafall
4) Husholdning endret status så utenfor målpopulasjonen	Avgang

Kommentarer:

Ved disse undersøkelser holder vi regning med flere enn de som skal være med i undersøkelsen. Hvis en husholdning har flyttet ut og ingen ny har flyttet inn på uttrukket adresse, sendes frafalls/avgangsskjema for husholdningen med passende avgangsårsak. Ved indirekte personutvalg skal det meldes avgang for enkeltpersoner som har flyttet fra en husholdning hvis disse personene står på IO-lista. Årsaken til at vi bryr oss om disse tilfellene er kontroll av feltarbeidet og vurdering av metoden.

Ajourføring av utvalg

Selv om trekkeregisteret nå skal ajourføres hvert halvår, er det fortsatt aktuelt å henge seg på de månedlige kjøringene på Personregisteret og få ajourført opplysninger om enhetene i trukne utvalg.

For personutvalg kan vi behandle eventuelle nye opplysninger på samme måte som reglene for intervjuerne tilsier. Vi slipper bl.a. å sende intervjuerne til noen fraflyttede adresser og får tatt ut noen døde personer. For de andre to typene utvalg kan vi ikke bruke flytteopplysninger på samme måte. Vi holder fast ved de uttrukne adressene selv om personene i trekkeregisteret er flyttet. I tillegg har vi også her nytte av ajourførte personopplysninger om døde og skifte av status.

Regnskap over frafall og avgang

Frafall og avgang skal rapporteres etter årsak. Hvis ikke spesielle forhold tilsier det, skal en bruke de vanlige kategorier og koder. Da oppnår vi sammenlignbarhet i den løpende vurdering av feltarbeid og metoder. Opplysninger om frafallet har selvsagt også betydning for analyser av den enkelte undersøkelse.

Siden vi opererer med to ulike tenkesett ved standardundersøkelsene, har vi to ulike normer for føring av frafall og avgang. En del kategorier er ikke aktuelle for alle typer undersøkelser.

En del av kategoriene er det bare Byrået som skal bruke.

Nedenfor har vi satt frafalls-/avgangsårsaker som skal brukes. For hver kategori er angitt for hvilke typer undersøkelse de er aktuelle. P= Personundersøkelse, H=Husholdningsundersøkelse, X=Personer trukket som "alle" i en husholdning Umerkede årsaker er aktuelle for alle typer. Ring rundt koden på de kategorier som vanligvis er reservert i Byrået.

FRAFALL

- 00 IO nekter (ved besøk eller telefonkontakt)
- 01 IO nekter ved melding til Byrået.
- 02 Andre nekter for IO
- 10 IO er kortvarig syk
- 11 IO er langvarig syk
- 20 Sykdom eller dødsfall i IO's familie
- 30 IO er bortreist, på ferie o.l.
- 31 IO er borte på arbeid, forretningsreise o.l.
- 32 IO er borte på skole, studieopphold, militærtjeneste o.l.
- 33 IO er ikke å treffe, ikke nærmere opplysninger
- AKU 34 IO er på fiske
- AKU 35 IO er til sjøs (innenriks- eller utenriksfart)
- 40 Ikke funnet IO's bolig/adresse
- P 41 IO er flyttet, adresse ukjent
- P (50) IO er ikke oppsøkt, bor utenfor utvalgsområdene
- (51) IO er ikke oppsøkt, andre kostnadmessige årsaker
- (52) Mangler intervjuer som kan oppsøke IO
- (60) Skjema er ufullstendig besvart/mangler opplysninger fra intervjuer
- (61) Mangler opplysning pr. post.
- 70 Annet

AVGANG

	80	IO er død
	81	Skiftet status så utenfor målpopulasjon
PX	90	IO er flyttet til utlandet
HX	91	Tomt hus, revet, nedbrent
X	92	IO oppført på IO-liste flyttet
AKU	93	IO/hush., nylig vært med i AKU.
	99	Annet,.....

Når en fyller ut skjemaet skal en bruke så spesifiserte kategorier som mulig. "Annet" og andre restkategorier brukes bare i de tilfelle en ikke vet mer etter å ha prøvd å skaffe opplysninger.

Feltgruppa går igjennom frafallet og vil følge opp noen av tilfellene. F.eks. kan intervjuer finne at tildelt IO er flyttet, adresse ukjent (41) og sender inn skjema om dette. Feltgruppa finner ny adresse, sender annen intervjuer og IO nekter. Da skal dette IO føres som frafall (00). Generelt skal "siste" årsak gjelde for IO som er utsatt for flertrinnsbehandling.

Over tid har flere frafalls/avgangsårsaker vært i bruk. Vi har ikke tatt med alle. Noen er noe uklare og/eller lite utnyttet. Men ved undersøkelser med spesiell definisjon av målpopulasjon/husholdning kan spesielt andre avganger være aktuelle (bosatt institusjon, bosatt fellehusholdning). Ellers vil personer/husholdninger uten fast bopel regnes som avgang i alle vanlige undersøkelser. Til en "vanlig" undersøkelse er neppe alle kategoriene aktuelle. En slags minimumsløsning for frafallsårsakene kan være

- 0 Nekting
- 10 IO er kortvarig syk
- 11 IO er langvarig syk
- 2 Sykdom eller dødsfall i IOs familie
- 30 IO er bortreist, på ferie o.l.
- 31 IO er borte på arbeid, forretningsreise o.l.
- 32 IO er borte på skole, studieopphold, militærtjeneste o.l.
- 33 IO er ikke å treffe, ikke nærmere opplysninger.
- 4 Ikke funnet IOs bolig/adresse
- 5 IO ikke oppsøkt
- 6 Mangler opplysning
- 7 Annet

Og det er neppe aktuelt med alle disse i en standardpublisering.

Da kan det holde med:

0 Nekting

1-2 Sykdom IO eller familie

3 IO ikke å treffe (p.g.a. midlertidig fravær).

4-5 Ikke kontakt med IO, andre årsaker

Rest.Annet

Avgangsårsaken skal ikke publiseres, de brukes til intern kontroll og vurdering.