



*Anne Vedø og Leiv Solheim*

**En praktisk innføring i  
utvalgsplanlegging**

# Notater

## **Innhold**

<b>1. Innledning</b> .....	<b>3</b>
<b>2. Populasjonen</b> .....	<b>3</b>
2.1.    Registre .....	4
<b>3. De viktigste utvalgsplanene</b> .....	<b>4</b>
3.1.    Enkelt tilfeldig utvalg.....	4
3.2.    Stratifisert utvalg.....	6
3.3.    Systematisk utvalg .....	7
3.4.    Klyngeutvalg .....	8
3.5.    Pps-utvalg.....	8
3.6.    Totrinnsutvalg .....	8
3.7.    SSBs generelle utvalgsplan .....	9
3.8.    Selvveiende utvalg .....	10
<b>4. Utvalgsstørrelse</b> .....	<b>10</b>
4.1.    Stratifisert enkelt tilfeldig utvalg.....	11
4.2.    Andre utvalgsplaner .....	13
4.3.    I praksis .....	13
<b>5. Teknisk gjennomføring</b> .....	<b>13</b>
<b>6. Dokumentasjon</b> .....	<b>14</b>
<b>7. Noen enkle råd</b> .....	<b>14</b>
<b>8. Nasjonale prøver i norsk - utvalg av elever i 8. årstrinn. Et stratifisert tottrinns/ klyngeutvalg</b> .....	<b>15</b>
8.1.    Innledning .....	15
8.2.    Populasjon.....	15
8.3.    Utvalgsplan .....	18
8.4.    Oversikt over utvalget med 1. og 2. reserver .....	23
<b>9. Industriavfallsundersøkelsen 2003. Et stratifisert systematisk utvalg</b> .....	<b>25</b>
9.1.    Populasjon.....	25
9.2.    Utvalgsplan .....	25
9.3.    Utvalg - dekningsgrad .....	30
<b>10. Levekår i landbruket 2002. Et stratifisert systematisk utvalg, etter region og driftstype</b> .....	<b>33</b>
10.1.    Undersøkelsespopulasjon.....	33
10.2.    Trekkepopulasjon.....	33
10.3.    Utvalgsplan .....	33
10.4.    Utvalget.....	35
<b>11. Undersøkelsen om samvær og bidrag 2002</b> .....	<b>36</b>
11.1.    Undersøkelsespopulasjon .....	36
11.2.    Trekking .....	36
11.3.    Enheter og begreper .....	36
11.4.    Treksannsynligheter .....	37
11.5.    Utvalg .....	38
<b>12. Referanser</b> .....	<b>38</b>
<b>De sist utgitte publikasjonene i serien Notater</b> .....	



# 1. Innledning

Dette notatet er ment å være en enkel og praktisk innføring i utvalgsplanlegging, og er beregnet på personer uten tidligere erfaring i dette. Vi har forsøkt å bruke minst mulig matematikk og formler. For mer litteratur om emnet kan man se for eksempel [1] og [2].

Notatet er et resultat av en forelesningsrekke om utvalgsplanlegging holdt ved Seksjon 120 høsten 2005. I den første delen, kapitlene 1-7, beskrives prinsippene for utvalgsplanlegging. Deretter følger fire eksempler. Den første delen er skrevet av Anne Vedø, delvis basert på Leiv Solheims forelesning. Kapitlet om nasjonale prøver i norsk er skrevet av Leiv Solheim. Kapitlet om utvalgsplanen for industriavfallsundersøkelsen er opprinnelig skrevet av Anne Sofie Abrahamsen, men er senere blitt bearbeidet og oppdatert av Anne Vedø. Kapitlene om levekår i landbruket og samvær og bidrag er skrevet av Anne Vedø.

## 2. Populasjonen

Vi tar utgangspunkt i en populasjon vi ønsker å undersøke. Det kan være for eksempel personer bosatt i Norge mellom 18 og 74 år eller industribedrifter som var aktive et gitt år. Vi kaller populasjonen for  $U$  og nummererer enhetene fra 1 til  $N$ .

$$U = \{1, 2, \dots, N\}$$

Variabelen vi ønsker å måle kaller vi  $y$ . Hver enhet  $i$  i populasjonen har en verdi  $y_i$  på denne variabelen, så de ukjente verdiene er  $\{y_1, y_2, \dots, y_N\}$ .

Interessevariabelen  $y$  kan for eksempel være en kontinuerlig variabel som faktisk arbeidstid forrige uke, eller en kategorisk variabel som bare kan ta noen få bestemte verdier, som ja og nei.

Vi er interessert i *populasjonstotalen*

$$t = \sum_{i=1}^N y_i = y_1 + y_2 + \dots + y_N$$

eller *populasjonsgjennomsnittet*

$$\mu = t / N$$

En viktig egenskap ved populasjonen er *populasjonsvariansen*

$$\sigma^2 = \frac{1}{N-1} \sum_{i=1}^N (y_i - \mu)^2$$
<sup>1</sup>

Denne uttrykker hvor stor variasjon det er mellom enhetene i populasjonen. Vi er ikke alltid interessert i denne størrelsen i seg selv, men den har betydning for presisjonen på estimatene for populasjonstotalen og gjennomsnittet. *Standardavviket i populasjonen*,  $\sigma$ , er kvadratroten av populasjonsvariansen.

---

<sup>1</sup> Noen fremstillinger deler på  $N$  i stedet for  $N-1$  i definisjonen av  $\sigma^2$ . I så fall får man også litt andre formler for størrelser som avhenger av  $\sigma^2$ .

Når  $y$  er en binær variabel (en variabel som bare kan ta verdiene 0 og 1), kalles gjerne  $\mu$  for  $p =$  andelen som er 1. Hvis  $N_0$  er antall i populasjonen med  $y = 0$  og  $N_1$  er antall i populasjonen med  $y = 1$ , så er  $\mu = p = N_1 / N$ ,  $N_1 = Np$  og  $N_0 = N(1 - p)$ . Populasjonsvariansen kan da skrives:

$$\begin{aligned}\sigma^2 &= \frac{1}{N-1} \sum_{i=1}^N (y_i - p)^2 = \frac{1}{N-1} (N_1(1-p)^2 + N_0(0-p)^2) \\ &= \frac{1}{N-1} N(p(1-p)^2 + (1-p)p^2) \\ &= \frac{1}{N-1} Np(1-p)\end{aligned}$$

## 2.1. Registre

Når man skal trekke et tilfeldig utvalg fra en populasjon er det en stor fordel å ha alle enhetene i populasjonen samlet på en datafil. For noen populasjoner finnes det administrative registre som gir oss god oversikt over populasjonen.

I personstatistikken er Folkeregisteret det viktigste. Her finnes opplysninger om blant annet kjønn, alder, adresse, sivilstatus, innvandringsbakgrunn og familienummer. Ved folke- og bolig tellingen i 2001 ble det også opprettet et bolignummer. Det er fremdeles ikke alle boliger som har fått dette nummeret, men når dette er i orden, vil vi ha full oversikt over hvem som bor i samme bolig.

I Folkeregisteret, som i alle registre er det et visst problem med etterslep, dvs. det tar litt tid før endringer som for eksempel fødsler, dødsfall, og flytting blir registrert. For studenter er det også frivillig om de vil melde flytting. Folkeregisteret regnes likevel som en god tilnærming til den faktiske personpopulasjonen.

Et annet vesentlig personregister er registeret over befolkningens høyeste utdanning.

I bedriftsstatistikken er Bedrifts og Foretaksregisteret (BoF) det viktigste registeret. Hvert foretak er delt inn i en eller flere bedrifter. BoF inneholder opplysninger om blant annet næringskode og antall sysselsatte. I bedrifts populasjonen skjer endringer hyppigere enn blant personer, så BoF har litt større problemer med å holde seg oppdatert enn Folkeregisteret.

## 3. De viktigste utvalgsplanene

### 3.1. Enkelt tilfeldig utvalg

Et enkelt tilfeldig utvalg (ETU) kan lages ved å trekke en og en enhet fra populasjonen uten tilbakelegging. I hver trekning skal alle enheter ha lik sannsynlighet for å bli valgt ut. Ved enkelt tilfeldig utvalg blir alle mulige utvalg like sannsynlige, og alle enhetene i populasjonen får samme mulighet for å være med i utvalget. I praksis er det ikke så mange utvalg som trekkes rent enkelt tilfeldig, men man kan trekke enkelt tilfeldig innen forskjellige delpopulasjoner eller strata. ETU er enkelt å regne på, og er en grunnsten i utvalgsteorien.

Vi trekker et utvalg  $s$  på  $n$  enheter. Treksannsynligheten, dvs., sannsynligheten for at enhet  $i$  er med i utvalget  $s$ , betegnes med  $P(i \in s)$ . For ETU er  $P(i \in s) = n / N$  for alle enheter  $i$ .

Vi samler inn informasjon om interessevariabelen  $y$  for alle enheter i utvalget. Vi kjenner dermed verdiene  $\{y_i : i \in s\}$ .

En *estimator* er en metode eller en formel for hvordan man skal anslå en ukjent populasjonsparameter ut fra informasjon om et utvalg. En enkel estimator for det ukjente populasjonsgjennomsnittet  $\mu$ , er gjennomsnittet i utvalget:

$$\bar{y}_s = \frac{1}{n} \sum_{i \in s} y_i$$

Populasjonstotalen kan da estimeres med:

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

Et *estimat* for en populasjonsparameter er det tallet man får når man setter inn dataene fra et spesifikt utvalg i estimatoren. Estimaten for populasjonsgjennomsnittet og populasjonstotalen vil variere med utvalget. Denne variasjonen kalles *utvalgsvarians* og måles på samme måte som andre varianser i statistikken, dvs. man ser på hvor mye et estimat typisk avviker fra gjennomsnittet av alle estimater. For å finne dette tallet må man egentlig regne ut estimatet for alle mulige utvalg, finne gjennomsnittsverdien for estimater, finne avviket mellom estimat og gjennomsnitt for hvert estimat og kvadrere dette tallet, og til slutt ta gjennomsnittet av alle de kvadratiske avvikene. Heldigvis finnes det for ETU en enklere måte å komme fram til det samme tallet på:

$$Var(\bar{y}_s) = (1-f) \frac{\sigma^2}{n}, \quad f=n/N, \quad Var(\hat{t}) = N^2 Var(\bar{y}_s)$$

Variansen til estimatoren øker med økende populasjonsvarians. Dette betyr at det er vanskeligere å estimere gjennomsnittet i en populasjon der enhetene er veldig forskjellige. Man må da ha større utvalg for å oppnå samme presisjon.

Variansen minker med økende utvalgsstørrelse, som regel mest på grunn av faktoren  $\sigma^2/n$ , men også på grunn av den endelige korreksjonsfaktoren  $1-f$ . For de fleste personundersøkelser er  $n$  mye mindre enn  $N$ , slik at  $1-f$  blir tilnærmet lik 1 og utvalgsandelen  $f$  kan neglisjeres, men for bedriftsundersøkelser kan dette leddet bli merkbart mindre enn 1. Når utvalgsstørrelsen er lik populasjonsstørrelsen blir den endelige korreksjonsfaktoren 0, og all utvalgsusikkerhet forsvinner.

Når  $y$  er en binær variabel blir  $\bar{y}_s = \hat{p}$  = andel med verdi 1 i utvalget, og

$$Var(\hat{p}) = \frac{1-f}{n} \frac{N}{N-1} p(1-p) \approx \left(\frac{1}{n} - \frac{1}{N}\right) p(1-p)$$

Populasjonsvariansen  $\sigma^2$  er dessverre som regel ukjent. Den kan estimeres med den tilsvarende størrelsen for utvalget

$$\hat{\sigma}^2 = \frac{1}{n-1} \sum_{i \in S} (y_i - \bar{y}_s)^2$$

Dette estimatet settes så inn i uttrykket for  $Var(\bar{y}_s)$  i stedet for  $\sigma^2$  for å få et estimat på  $Var(\bar{y}_s)$ . På samme måte kan  $\hat{p}$  settes inn for  $p$  i uttrykket for  $Var(\hat{p})$ .

### 3.2. Stratifisert utvalg

For å få bedre kontroll over utvalget, er det vanlig å dele populasjonen inn i grupper, eller strata, og deretter trekke et utvalg fra hvert stratum. En populasjon av personer kan for eksempel stratifiseres etter kjønn, alder og fylke. For bedrifter er næring og størrelse vanlige stratifiseringsvariable. Fordeler ved stratifisering er:

- *Mer representative utvalg og mer presis estimering av totaltall*

Enkelt tilfeldig utvalg gir representative utvalg i gjennomsnitt, men vi kan være uheldige og trekke et skjevt utvalg. Stratifisering tillater oss å luke bort skjeve, uheldige utvalg allerede i trekkingen, og vil gi mer presise estimater av totaltall i populasjonen.

- *Behov for estimering av tall på mindre områder, detaljert statistikk*

Vanligvis er det ønskelig å publisere tall ned på mindre områder, for eksempel fylker eller næringsgruppe. Ved et stratifisert utvalg kan man på forhånd bestemme hvor store strata-utvalgene skal være og dermed sikre en ønsket presisjon for "områdeestimatene". Dette er særlig viktig når vi er avhengige av at flere mindre grupper, som for eksempel små næringer, er representert. Hvis tilfeldighetene får råde, er det ikke så usannsynlig at noen av gruppene som burde hatt for eksempel 25 enheter, ender opp med så få enheter at det blir vanskelig å publisere tall for dem.

- *Forskjellige innsamlingsmetoder*

Stratifisering gjør det mulig å benytte forskjellige måter å samle inn data på, for eksempel telefonintervjuer eller hjemmebesøk

- *Mulighet til å oversample visse grupper*

Noen ganger kan man ønske at noen grupper får mer en sin rettmessige andel i utvalget. For eksempel oversamples personer i Finnmark i arbeidskraftundersøkelsene. Dette er for å kunne gi gode estimater for ledighet og endring i ledighet i alle fylker. Hvis man har oversamlet noen strata, må man bruke vekter når man regner ut estimater på tvers av strata.

En annen grunn til å oversample noen grupper, er hvis man vet at disse gruppene har store  $y$ -verdier. For eksempel vet vi at store bedrifter ofte har store verdier på de fleste variable. Derfor oversampler man store bedrifter i bedriftsutvalg. Vi får da sikrere estimater i de strataene som bidrar mest til populasjonstotalen.

#### 3.2.1. Stratifiseringsvariable

Vanlige stratifiseringsvariable er

- For personer: Kjønn, alder, fylke, kommune, utdanning, innvandringsbakgrunn
- For bedrifter: Næring, antall ansatte
- For foretak: Eierform, kompleksitet

Når vi velger stratifiseringsvariable ser vi på

- Hvilke variable vi har tilgang til i det registeret det skal trekkes fra
- Hvilke grupper vi ønsker å publisere tall for
- Hvilke variable som kan tenkes å påvirke undersøkelsesvariablene så mye at det er viktig at utvalget er representativt med hensyn på disse

Ofte ønsker man å sikre seg at utvalget er representativt med hensyn på mange forskjellige variable. Et problem er imidlertid at det fort blir veldig mange strata, og dermed få enheter i hvert. Stratifisering på kjønn, 5 aldersgrupper og fylke gir for eksempel allerede  $2 \times 5 \times 19 = 190$  strata. Få enheter i endel strata trenger ikke være noe stort problem hvis man ikke skal estimere tall for hvert enkelt stratum, men det er et godt prinsipp å begrense seg til de viktigste variablene ved stratifisering.

### 3.2.2. Allokering

Når vi har bestemt oss for hvilke strata vi vil bruke og hvor stort utvalget skal være totalt, må vi finne ut hvordan utvalget skal fordeles, eller allokeres, på strataene.

#### 3.2.2.1. Proporsjonal allokering

Utvalgsstørrelsen i stratumet settes proporsjonal med populasjonsstørrelsen i stratumet.

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

der  $n_h$  og  $N_h$  er henholdsvis utvalgs- og populasjonsstørrelse i stratum  $h$ . Dette gir samme trekk sannsynlighet i alle strata.

#### 3.2.2.2. Optimal allokering

Hvis populasjonsvariansen i hvert stratum er kjent, eller kan estimeres på grunnlag av tidligere undersøkelser eller basert på et lite foreløpig utvalg (pilot), kan man sette utvalgsstørrelsen proporsjonal med produktet av stratumstørrelsen og standardavviket i populasjonen i hvert stratum.

$$n_h = n \frac{N_h \sigma_h}{\sum_{alle h} N_h \sigma_h}$$

Dette vil gi minimal usikkerhet på estimatene for gjennomsnitt og total for *hele* populasjonen. Hvis stratumvariansene er like, ser vi at proporsjonal allokering er optimal. Usikkerheten til den stratifiserte estimatoren er lite følsom overfor avvik fra optimal allokering. Det trengs derfor kun grove anslag på stratumvariansene. Et problem her er at populasjonsvariansene i strataene ofte varierer fra variabel til variabel. Man må da velge en primær interessevariabel for optimal allokering. Vi har ofte flere ”primære” interessevariable i hver undersøkelse. Derfor foretrekkes ofte proporsjonal allokering i praksis.

### 3.3. Systematisk utvalg

Ved systematisk trekking trekkes hver  $r$ -te enhet fra en liste. Steglengden  $r$  tilpasses slik at man får den utvalgsstørrelsen man ønsker. Man starter med å sortere lista etter passende variable, for eksempel kommunenummer og alder. Man velger så et tilfeldig startpunkt fra og med 1 til og med  $r$ , og trekker deretter hver  $r$ -te enhet. For eksempel kan man starte på enhet 4 og ta hver 100de enhet derfra, altså 4, 104, 204 osv. Utvalget vil da bli omtrent representativt med hensyn på sorteringsvariablene.



Systematisk utvalg gir noe av den samme effekten som stratifisering. Men merk at hvis man bruker svært mange sorteringsvariable, er det ikke sikkert utvalget blir representativt med hensyn på de siste. Det er fordi de bakerste sorteringsvariablene etter hvert skifter fra minimum til maksimum i løpet av færre enheter enn steglengden.

En ulempe ved systematisk trekking er at det ikke er mulig å estimere varianser direkte. Variansen til  $\bar{y}_s$  er ikke lenger bare avhengig av den ukjente populasjonsvariansen  $\sigma^2$  som ved ETU, men av den nøyaktige rekkefølgen enhetene i populasjonen ligger i før trekking. Det er derfor nødvendig med informasjon om populasjonens struktur for å estimere usikkerheten til estimatet. For mer detaljert informasjon om variansestimering viser vi til [2] (ch. 5.6)

### 3.4. Klyngeutvalg

I klyngeutvalg trekkes enhetene klyngevis. Et eksempel på klyngeutvalg er utvalget i SSBs arbeidskraftsundersøkelser, der det trekkes familier. Det er relativt lite ekstra arbeid å intervju hele familien når man først har fått tak i ett familiemedlem. Et annet eksempel er skoleklasser. Når man skal undersøke skoleelever, er det administrativt enklere å undersøke hele klasser enn bare noen i hver klasse.

Ulempen med klyngeutvalg er at vi må regne med at enhetene innen en klynge er likere enn enheter i forskjellige klynger. Dette gjelder i hvert fall for en del naturlige klynger. Et klyngeutvalg gir derfor i sin alminnelighet mindre informasjon og større usikkerhet på estimatene enn et utvalg der enhetene er trukket individuelt, gitt samme utvalgsstørrelse. De praktiske fordelene, bl.a., reduserte kostnader, kan likevel gjøre at vi velger klyngeutvalg.

### 3.5. Pps-utvalg

Pps står for probability proportional to size. I et pps-utvalg trekkes enhetene med sannsynlighet proporsjonal med størrelsen. Et eksempel på dette får vi når vi trekker husholdninger gjennom personer. Da trekker vi først et enkelt tilfeldig utvalg av personer, og lar så husholdningsutvalget bestå av husholdningene til de uttrukne personene. Når vi trekker på denne måten, vil store husholdninger (dvs. husholdninger med mange medlemmer) få større trekkingsannsynlighet, og derfor bli overrepresentert i utvalget. Dette er fordi store husholdninger har flere representanter, så det er større sjanse for at de er representert i det opprinnelige personutvalget. En en-personhusholdning er avhengig av at en bestemt person blir trukket ut til personutvalget, mens for en tre-personhusholdning er det tilstrekkelig at en av tre personer trekkes ut.

Et annet eksempel er å trekke bedrifter med sannsynlighet proporsjonal med antall sysselsatte. Dette er ofte fornuftig, fordi mange interessevariable, som regnskapstall og produksjon øker med økende sysselsetting. I praksis gir man som regel ikke hver enkelt bedrift trekkingsannsynlighet nøyaktig proporsjonal med størrelsen, men deler i stedet bedriftene inn i størrelsesgrupper, og trekker en større andel fra strata med store bedrifter enn fra strata med mindre bedrifter. Dette gir samme effekt og er teknisk litt enklere.

### 3.6. Totrinnsutvalg

Behovet for å trekke tottrinnsutvalg melder seg når vi ønsker å gjennomføre besøksintervjuer. Hvis vi trekker et enkelt tilfeldig utvalg i hver kommune, vil det medføre veldig mye reising per intervjuobjekt utenfor tettbygde strøk. For å unngå dette, trekker man først ut noen kommuner jevnt fordelt over landet, og deretter trekker man enheter innen disse.

En annen grunn til å trekke i to trinn kan være hvis det mangler et sentralt register over enhetene i populasjonen. Man kan da trekke ut et antall administrative områder i første trinn, og deretter trekke fra lokale registre.

Et viktig eksempel på en to-trinns utvalgsplan er SSBs generelle utvalgsplan beskrevet i neste avsnitt. Andre eksempler er:

- Trekke et utvalg av bedrifter og intervju noen ansatte i hver av disse
- Trekke noen skoler og undersøke et antall elever i hver av disse
- Trekke landsbyer og intervju beboerne i hvert tiende hus

### **3.6.1. Designeffekt**

Usikkerheten øker når vi trekker i to trinn. Variansforholdet mellom to-trinnsutvalg og ett-trinnsutvalg kalles designeffekten. Designeffekten er forskjellig for hver variabel, og er et uttrykk for hvor mye vi taper på å trekke i to trinn i forhold til om vi hadde brukt ett trinn. En designeffekt på 2 sier altså at vi får dobbelt så høy varians med to-trinnsplan. Jo mer forskjellige første-trinnsenheter (for eksempel kommunene) er med hensyn på en variabel, jo større blir designeffekten. Designeffekt øker også med økende utvalgsstørrelse. Man taper mer på å konsentrere et stort utvalg inn i et begrenset antall kommuner enn et lite.

## **3.7. SSBs generelle utvalgsplan**

Statistisk sentralbyrå har en generell to-trinns utvalgsplan som brukes når man skal gjennomføre besøksintervju av personer, familier og husholdninger.

Den primære utvalgsenheten, altså det som trekkes på første trinn i prosessen, er stort sett kommuner. I noen tilfeller er to eller flere mindre kommuner slått sammen til en primær utvalgsenhet (PUE).

De primære utvalgsenheterne er delt inn i 109 strata. Strataene er laget slik at PUEene innen et stratum er så like som mulig. Fra hvert stratum trekkes en PUE. Personene herfra skal representere alle personene i hele stratumet.

De største kommunene (byer) er alene i sitt stratum, som betyr at de trekkes med sikkerhet. Det er 34 strata som bare inneholder en PUE. I praksis har vi derfor egentlig en ett-trinnsplan i disse kommunene. Omtrent halvparten av befolkningen bor i slike kommuner.

I de andre 75 strataene trekkes en av flere PUEer. Trekk sannsynligheten for en PUE er proporsjonal med innbyggertallet.

Når første trinn er trukket, holdes de primære utvalgsenheterne fast i omtrent ti år, og intervjuere blir ansatt på de aktuelle stedene.

I trinn to trekkes personer i de uttrukne PUEene. Trekk sannsynligheten på annet trinn er justert slik at alle personer i landet får samme sannsynlighet for å være med i utvalget når man tar hensyn til begge trinn.

### **3.7.1. Designeffekt i SSBs generelle utvalgsplan**

I [3] er det beregnet designeffekt for sysselsetting etter næring og utvalgsstørrelse i alle fylker. Under er tabellen for Sogn og Fjordane.

**Tabell 3.7.1 Designeffekt for sysselsetting etter næring i Sogn og Fjordane**

Næring - Standard for næringsgruppering 83	Utvalgsstørrelse (hele landet)		
	4 000	12 000	40 000
1. Jordbruk, skogbruk, fiske og fangst	1,24	1,73	3,44
2. Olje og bergverk	1,04	1,13	1,43
3. Industri	1,97	3,98	11,01
4. Kraft og vannforsyning	1,11	1,34	2,14
5. Bygg og anlegg	1,44	2,35	5,54
6. Varehandel, hotell og restaurant	1,17	1,54	2,81
7. Transport, lager, post, tele	1,11	1,35	2,17
8. Finans, forretning, tjenesteyting m.m.	1,05	1,17	1,56
9. Offentlig og privat tjenesteyting	1,30	1,92	4,08
10. Uoppgitt næring for selvstendige	1,01	1,03	1,11
11. Uoppgitt næring for arbeidstakere	1,05	1,15	1,52

Vi ser at designeffekten øker med størrelsen på utvalget. Dette er fordi variansen på første trinn er konstant uansett utvalgsstørrelse, mens variansen på annet trinn minker med utvalgsstørrelsen. Hvis vi skulle redusert variansen på første trinn, måtte vi ha trukket flere PUEer.

Designeffekten blir spesielt stor for industrien. Dette er fordi andelen industriarbeidere varierer mye mellom PUEer, selv innen strata.

### 3.8. Selvveiende utvalg

I en selvveiende utvalgsplan har alle enheter i populasjonen lik sannsynlighet for å trekkes. Eksempler på selvveiende utvalgsplaner er enkelt tilfeldig utvalg, stratifisert enkelt tilfeldig utvalg med proporsjonal allokering, systematisk utvalg, enkel tilfeldig trekking av klynger og SSBs generelle utvalgsplan.

De fleste personutvalg og familieutvalg i SSB trekkes selvveiende. Dette er ikke tilfelle for bedrifter og foretak. Her favoriseres store enheter. Det er til og med vanlig å ta med *alle* enheter over en viss størrelse.

## 4. Utvalgsstørrelse

Et vanlig spørsmål er hvor stort utvalget må være. Utvalgets størrelse er en viktig faktor når det gjelder kostnad i forbindelse med en undersøkelse. De viktigste statistiske faktorer, foruten selve utvalgsplanen som er bestemt ut fra kriterier nevnt tidligere, er

- Krav til estimeringsnøyaktighet for mange variable. Fokuser på en eller to variable av størst interesse. Presise estimater krever større utvalg.
- Homogenitet i populasjonen. Trenger mindre utvalg hvis det er liten variasjon i populasjonen
- Estimering for mindre områder, grupper i populasjonen

Det er ofte den siste faktoren som setter de største kravene til undersøkelsen.

## 4.1. Stratifisert enkelt tilfeldig utvalg

### 4.1.1. Binær undersøkelsesvariabel

Anta at vi vil måle andelen som har en bestemt egenskap, for eksempel andelen arbeidsledige, i et gitt stratum av størrelse  $N$ , og vi ønsker å bestemme stratumutvalgets størrelse  $n$ .

Undersøkelsesvariabelen  $y$  er 1 for arbeidsledige, 0 ellers. Den sanne andelen er  $p = \frac{1}{N} \sum_{i=1}^N y_i$ . Den

estimeres med andelen i utvalget  $\hat{p} = \frac{1}{n} \sum_{i \in S} y_i$ . Vi ønsker å ha utvalget så stort at vi oppnår en på forhånd bestemt presisjon. Som mål på presisjon er det vanligste å se på nøyaktigheten til et 95% konfidensintervall for  $p$ , tilnærmet lik

$$\hat{p} \pm 2\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n} - \frac{1}{N}\right)} \approx \hat{p} \pm 2\sqrt{\hat{p}(1-\hat{p})/n}, \text{ for store } N$$

Størrelsen som skal trekkes fra og legges til  $\hat{p}$  kalles *feilmarginen*. Et mulig krav er at den forventede lengden til konfidensintervallet er proporsjonal med  $p$  (når  $p \leq 0,5$ , ellers se på konfidensintervall for  $1-p$ ), dvs. at den forventede feilmarginen er lik  $d \cdot p$ :

$$2\sqrt{\frac{p(1-p)}{n}} = d \cdot p$$

En annen måte å formulere det samme kravet på er:

$$CV(\hat{p}) = d/2 = cv$$

Her er CV estimatorens relative standardfeil eller variasjonskoeffisienten, som det også kalles. Standardfeilen til  $\hat{p}$  er kvadratroten av variansen til  $\hat{p}$ , og relativ standardfeil er standardfeil delt på  $p$ . Den uttrykker hvor stor standardfeilen er i forhold til det som skal estimeres, og den kan uttrykkes i prosent. Man kan for eksempel kreve en relativ standardfeil på høyst fem prosent. Da vet man at standardfeilen ikke blir mer enn fem prosent av den verdien man estimerer. Det betyr også at den forventede feilmarginen er ti prosent av  $p$ .

$$CV(p) = \frac{\sqrt{p(1-p)/n}}{p} = cv \Leftrightarrow n = \frac{(1-p)/p}{cv^2}$$

Siden den sanne andelen  $p$  er ukjent, brukes en planverdi  $p_0$ . Vi får nå utvalgsstørrelsen uttrykt som en funksjon av presisjonskrav  $cv$  og planverdi:

$$n = \frac{(1-p_0)/p_0}{cv^2}$$

For små strata og/eller små andeler, så blir planlagt utvalgsstørrelse i stratomet lik

$$n = \frac{(1 - p_0) / p_0}{cv^2 + \frac{1}{N}(1 - p_0) / p_0}$$

**Tabell 4.1.1 Nødvendig utvalgsstørrelse ved estimering av andel som funksjon av presisjonskrav og planverdi når  $N=100\ 000$**

Presisjon(cv)	$p_0=0,01$	$p_0=0,02$	$p_0=0,05$	$p_0=0,25$	$p_0=0,5$
<b>0,025</b>	61 301	43 946	23 313	4 580	1 575
<b>0,05</b>	28 367	16 388	7 063	1 186	398
<b>0,1</b>	9 008	4 671	1 865	299	100
<b>0,15</b>	4 215	2 131	837	133	44
<b>0,2</b>	2 415	1 210	473	75	25
<b>0,25</b>	1 559	778	303	48	16

For å estimere små andeler presist, kreves det veldig store utvalg. Store andeler krever mye mindre utvalg for samme relative presisjon. Presisjonskravet for CV betyr at konfidensintervallet er lik  $\hat{p} \pm 2cv \cdot p_0$ . For eksempel,  $cv = 0,1$  tilsvarer at

Når  $p_0 = 0,5$ : 95% konfidensintervall =  $\hat{p} \pm 0,10$  og  $n=100$

Når  $p_0 = 0,05$ : 95% konfidensintervall =  $\hat{p} \pm 0,01$  og  $n=1865$

Som eksempel, i estimering av arbeidsledighet for hele populasjonen kunne man bruke  $p_0 = 0,05$ . Siden det er viktig å oppdage endringer i arbeidsledigheten så kunne et rimelig krav være at feilmarginen er lik 0,002, dvs. at  $CV = 0,002/0,1 = 0,02 = 2\%$  og  $n = 47500$ .

#### 4.1.2. Ikke binær undersøkelsesvariabel

Vi skal estimere gjennomsnittlig  $y$ -verdi ved observert gjennomsnitt. La oss uttrykke presisjonskravet ved at variasjonskoeffisienten skal være lik  $cv$ , dvs.:

$$CV(\bar{y}_s) = \frac{\sqrt{V(\bar{y}_s)}}{\mu} = \sqrt{\frac{N-n}{N} \frac{(\sigma / \mu)}{\sqrt{n}}} = cv$$

La  $\sigma_0 / \mu_0$  være en planverdi for det relative standardavviket i stratomet. Når vi løser ut  $n$  får vi

$$n = \frac{(\sigma_0 / \mu_0)^2}{cv^2 + N^{-1}(\sigma_0 / \mu_0)^2}$$

Hvis man har god kunnskap om den relative variasjonen i populasjonen, så kan  $n$  bestemmes fra kravet til nøyaktighet. Vi ser også at CV-kravet betyr at kravet til stratumgjennomsnittets 95% konfidensintervall er at den forventede feilmarginen er lik  $2 \cdot cv \cdot \mu_0$ :

$$\bar{y}_s \pm 2 \cdot cv \cdot \mu_0$$

**Tabell 4.1.2 Nødvendig utvalgsstørrelse ved estimering av gjennomsnitt som funksjon av presisjonskrav og planverdi når N=100 000**

Presisjon(cv)	$\sigma_0 / \mu_0 = 0,1$	$\sigma_0 / \mu_0 = 0,2$	$\sigma_0 / \mu_0 = 0,3$	$\sigma_0 / \mu_0 = 0,5$	$\sigma_0 / \mu_0 = 1$	$\sigma_0 / \mu_0 = 2$	$\sigma_0 / \mu_0 = 5$
<b>0,005</b>	398	1 575	3 475	9 091	28 571	61 538	90 909
<b>0,01</b>	100	398	892	2 439	9 091	28 571	71 429
<b>0,02</b>	25	100	224	621	2 439	9 091	38 462
<b>0,03</b>	11	44	100	277	1 099	4 255	21 739
<b>0,05</b>	4	16	36	100	398	1 575	9 091
<b>0,1</b>	1	4	9	25	100	398	2 439
<b>0,25</b>	0	1	1	4	16	64	398
<b>0,5</b>	0	0	0	1	4	16	100

Hvis man ønsker andre presisjonskrav og relative standardavvik, er formelen enkel å programmere i Excel.

## 4.2. Andre utvalgsplaner

Andre utvalgsplaner fører til andre estimatorer og standardfeil . Det finnes formler for dette i bøker (se for eksempel [2]), og disse kan brukes til å lage tabeller over nødvendig utvalgsstørrelse, på samme måte som vi gjorde for stratifisert enkelt tilfeldig utvalg. I endel tilfeller kan det imidlertid vise seg å bli uforholdsmessig komplisert. En mulighet er da å bruke tabellene her allikevel som en grov tilnærming.

## 4.3. I praksis

I praksis har man gjerne mange undersøkelsesvariable, der noen krever større utvalgsstørrelse enn andre. Det kan også være vanskelig å finne fornuftige planverdier  $p_0$  og  $\sigma_0 / \mu_0$  for de sanne populasjonsparametrene. Vi må derfor konsentrere oss om de viktigste variablene, og forsøke å anslå planverdier på grunnlag av tidligere undersøkelser eller registervariable eller det man måtte ha av kunnskap på området.

Det er heller ikke alltid så lett å spesifisere på forhånd akkurat hva man vil ha ut av dataene, og hva som er tilstrekkelig nøyaktighet. Man kan lett tenke seg mange delpopulasjoner og sammenhenger som kan være interessante å undersøke.

## 5. Teknisk gjennomføring

Det finnes mange programmer som er i stand til å trekke tilfeldige utvalg. I SSB er SAS et av de mest brukte verktøyene. I SAS finnes det en prosedyre PROC SURVEYSELECT som kan håndtere alle utvalgsplanene beskrevet over. Her er et eksempel på hvordan man trekker et stratifisert, enkelt tilfeldig utvalg:

```
PROC SURVEYSELECT data=populasjon out=utvalg
method=srs n=(100,100);
strata kjonn;
run;
```

En detaljert beskrivelse av prosedyren finnes i SAS-manualen.

Når man har trukket utvalget, er det lurt å gjøre noen enkle kontroller. Se på utvalget, og lag frekvenstabeller over de viktigste kjennetegnene og sammenlign med fordelingen på populasjonsfila.

## 6. Dokumentasjon

Trekkeprosessen bør dokumenteres så nøyaktig som mulig. Når man senere skal beregne estimater og eventuelt anslå utvalgsvarians for disse estimatene, er det vesentlig å vite nøyaktig hvordan utvalget er trukket. Punkter som bør være med i dokumentasjonen er:

- En beskrivelse av populasjonen. Nevn enhet og hvordan populasjonen er avgrenset. Oppgi register eller annet datagrunnlag som er brukt til å konstruere en trekkepopulasjon, og på hvilket tidspunkt uttrekkene er tatt. Lag gjerne frekvenstabeller over viktige egenskaper i populasjonen hvis mulig.
- En beskrivelse av trekkingen:
  - Hvilke variable er brukt til å lage strata, og hvordan er utvalget allokert
  - Trekkemetode, dvs. hvordan det trekkes i hvert stratum. Er utvalget enkelt tilfeldig, systematisk eller pps, og trekkes det enkeltenheter eller klynger.
  - Hvis utvalget trekkes i flere trinn, beskriv alle trinn.
- Noen frekvenstabeller over det ferdige utvalget.

## 7. Noen enkle råd

Hvordan går man frem når man skal lage en utvalgsplan og trekke et utvalg? Under er noen enkle råd.

- Få oversikt over hva som skal måles. Les spørreskjemaet hvis det finnes.
- Finn ut hvilke resultater som er viktigst. Hva er hovedvariablene, og hvilke delpopulasjoner vil man brette tallene ned på.
- Få klarhet i hva slags enhet(er) og hvilken/hvilke populasjon(er) som undersøkes.
- Del inn i strata
- Ved beregning av utvalgsstørrelse tenk enkelt tilfeldig utvalg i strataene
- Prøv å lage utvalgsplanen så enkel som mulig. Ofte skal utvalget brukes av andre til mange ting, og det er en stor fordel at planen er enkel å gjennomføre og forstå.
- Trekk utvalget
- Kontroller utvalget
- Dokumenter hva du gjorde

Resten av dette notatet består av fire eksempler på utvalgsplaner.

## 8. Nasjonale prøver i norsk - utvalg av elever i 8. årstrinn. Et stratifisert totrinns/klyngeutvalg

### 8.1. Innledning

Statistisk sentralbyrå ved Seksjon for statistiske metoder og standarder har utarbeidet utvalgsplan og trukket et utvalg av skoler med drøyt 500 elever der nasjonale prøver i norsk skal prøves ut. I tabell 8.1.1. er de 22 skolene som er trukket ut fordelt etter målform, antall elever og full-/fådelte.

**Tabell 8.1.1. Oversikt over skolene til utvalget fordelt målform, antall elever og full-/fådelte.**

Målform	Antall elever	Totalt	Fådelte/fulldelte		
			Full	Få	Ukjent
<b>Totalt</b>	<b>Totalt</b>	<b>22</b>	<b>15</b>	<b>4</b>	<b>3</b>
	1-30	6	2	4	0
	31-60	3	3	-	0
	61-	13	10	-	3
<b>Bokmål</b>	<b>Totalt</b>	<b>16</b>	<b>11</b>	<b>2</b>	<b>3</b>
	1-30	3	1	2	0
	31-60	2	2	-	0
	61-	11	8	-	3
<b>Nynorsk</b>	<b>Totalt</b>	<b>6</b>	<b>4</b>	<b>2</b>	<b>-</b>
	1-30	3	1	2	-
	31-60	1	1	-	-
	61-	2	2	-	-

I avsnitt 8.2 gir vi en beskrivelse av skolene og elevene i 8. til 10. årstrinn i skoleåret 2004/2005 som er trekkegrunnlaget. I avsnitt 8.3 er utvalgsplanen og resultatet av trekkingen dokumentert. Til slutt i avsnitt 8.4 er skolene som ble trukket sammen med 1. og 2. reserver listet ut.

### 8.2. Populasjon

I tabell 8.2.1 og 8.2.2. er populasjonen av aktuelle skoler fordelt etter målform, antall elever i gjennomsnitt i årstrinnene 8 til 10, morsmålsundervisning (innvandrere) og om skolen er fådelte eller fulldelte. Skoler med bokmål dominerer som ventet. En legger også merke til at andelen skoler som er fådelte er langt høyere når det ikke gis morsmålsundervisning, altså ingen innvandrere på skolen.



**Tabell 8.2.1. Antall skoler etter målform, morsmålsundervisning, fulldelt/fådelt og antall elever**

Mål- form	Antall elever	Morsmålsundervisning				Ikke morsmålsundervisning			
		Totalt	Full	Få	Ukjent	Totalt	Full	Få	Ukjent
<b>Totalt</b>	<b>Totalt</b>	<b>615</b>	<b>464</b>	<b>58</b>	<b>93</b>	<b>592</b>	<b>323</b>	<b>188</b>	<b>81</b>
	A: 1-30	161	83	58	20	370	137	188	45
	B: 31-60	138	113	0	25	94	82	0	12
	C: 61-90	132	107	0	25	65	53	0	12
	D: 91-120	102	91	0	11	40	34	0	6
	E: 121-	82	70	0	12	23	17	0	6
<b>Bokmål</b>	<b>Totalt</b>	<b>500</b>	<b>362</b>	<b>53</b>	<b>85</b>	<b>444</b>	<b>236</b>	<b>138</b>	<b>70</b>
	A: 1-30	120	48	53	19	272	93	138	41
	B: 31-60	96	76	0	20	60	52	0	8
	C: 61-90	114	91	0	23	57	46	0	11
	D: 91-120	94	83	0	11	33	28	0	5
	E: 121-	76	64	0	12	22	17	0	5
<b>Nynorsk</b>	<b>Totalt</b>	<b>113</b>	<b>101</b>	<b>4</b>	<b>8</b>	<b>146</b>	<b>87</b>	<b>48</b>	<b>11</b>
	A: 1-30	40	35	4	1	96	44	48	4
	B: 31-60	41	36	0	5	34	30	0	4
	C: 61-90	18	16	0	2	8	7	0	1
	D: 91-120	8	8	0	0	7	6	0	1
	E: 121-	6	6	0	0	1	0	0	1
<b>Samisk</b>	<b>Totalt</b>	<b>2</b>	<b>1</b>	<b>1</b>	<b>0</b>	<b>2</b>	<b>0</b>	<b>2</b>	<b>0</b>
	A: 1-30	1	0	1	0	2	0	2	0
	B: 31-60	1	1	0	0	0	0	0	0
	C: 61-90	0	0	0	0	0	0	0	0
	D: 91-120	0	0	0	0	0	0	0	0
	E: 121-	0	0	0	0	0	0	0	0

**Tabell 8.2.2. Antall skoler etter fulldelt/fådelt og antall elever**

Målform	Antall elever	Fådelt/fulldelt			
		Totalt	Full	Få	Ukjent
<b>Totalt</b>	<b>Totalt</b>	<b>1 207</b>	<b>787</b>	<b>246</b>	<b>174</b>
	A: 1-30	531	220	246	65
	B: 31-60	232	195	0	37
	C: 61-90	197	160	0	37
	D: 91-120	142	125	0	17
	E: 121-	105	87	0	18
<b>Bokmål</b>	<b>Totalt</b>	<b>944</b>	<b>598</b>	<b>191</b>	<b>155</b>
	A: 1-30	392	141	191	60
	B: 31-60	156	128	0	28
	C: 61-90	171	137	0	34
	D: 91-120	127	111	0	16
	E: 121-	98	81	0	17
<b>Nynorsk</b>	<b>Totalt</b>	<b>259</b>	<b>188</b>	<b>52</b>	<b>19</b>
	A: 1-30	136	79	52	5
	B: 31-60	75	66	0	9
	C: 61-90	26	23	0	3
	D: 91-120	15	14	0	1
	E: 121-	7	6	0	1
<b>Samisk</b>	<b>Totalt</b>	<b>4</b>	<b>1</b>	<b>3</b>	<b>0</b>
	A: 1-30	3	0	3	0
	B: 31-60	1	1	0	0
	C: 61-90	0	0	0	0
	D: 91-120	0	0	0	0
	E: 121-	0	0	0	0

I tabellene 8.2.3 og 8.2.4 ser vi på en tilsvarende fordeling av elevene. Trekkene fra tabellene ovenfor over skolene blir enda tydeligere. Drøyt femten prosent av elevene bruker nynorsk, men nynorsk er målformen på over 20 prosent av skolene. Blant de skolene som ikke har morsmålsundervisning er målformen til hver femte elev nynorsk. På fådelte skoler uten morsmålsundervisning er målformen nynorsk til hver tredje elev, men det er kun drøyt 1 100 elever på slike skoler i alt. På fådelte skoler med morsmålsundervisning bruker under hver tiende elev nynorsk.

**Tabell 8.2.3. Antall elever etter målform, fremmedspråklige, fulldelt/fådelt og antall elever**

Mål- form	Antall elever	Morsmålsundervisning				Ikke morsmålsundervisning			
		Totalt	Full	Få	Ukjent	Totalt	Full	Få	Ukjent
<b>Totalt</b>	<b>Totalt</b>	<b>41 593</b>	<b>34 864</b>	<b>466</b>	<b>6 263</b>	<b>20 363</b>	<b>16 065</b>	<b>1 111</b>	<b>3 187</b>
	A: 1-30	2 317	1 641	466	210	3 829	2 419	1 111	299
	B: 31-60	6 165	5 042	0	1 123	4 067	3 530	0	537
	C: 61-90	10 011	8 132	0	1 879	4 800	3 879	0	921
	D: 91-120	10 830	9 647	0	1 183	4 269	3 641	0	628
	E: 121-	12 270	10 402	0	1 868	3 399	2 596	0	803
<b>Bokmål</b>	<b>Totalt</b>	<b>35 965</b>	<b>29 664</b>	<b>425</b>	<b>5 876</b>	<b>16 317</b>	<b>12 906</b>	<b>727</b>	<b>2 684</b>
	A: 1-30	1 587	965	425	197	2 622	1 618	727	277
	B: 31-60	4 386	3 483	0	903	2 661	2 287	0	374
	C: 61-90	8 706	6 981	0	1 725	4 239	3 386	0	853
	D: 91-120	9 969	8 786	0	1 183	3 556	3 019	0	537
	E: 121-	11 317	9 449	0	1 868	3 240	2 596	0	644
<b>Nynorsk</b>	<b>Totalt</b>	<b>5 588</b>	<b>5 162</b>	<b>39</b>	<b>387</b>	<b>4 035</b>	<b>3 159</b>	<b>373</b>	<b>503</b>
	A: 1-30	728	676	39	13	1 196	801	373	22
	B: 31-60	1 741	1 521	0	220	1 406	1 243	0	163
	C: 61-90	1 305	1 151	0	154	561	493	0	68
	D: 91-120	861	861	0	0	713	622	0	91
	E: 121-	953	953	0	0	159	0	0	159
<b>Samisk</b>	<b>Totalt</b>	<b>40</b>	<b>38</b>	<b>2</b>	<b>0</b>	<b>11</b>	<b>0</b>	<b>11</b>	<b>0</b>
	A: 1-30	2	0	2	0	11	0	11	0
	B: 31-60	38	38	0	0	0	0	0	0
	C: 61-90	0	0	0	0	0	0	0	0
	D: 91-120	0	0	0	0	0	0	0	0
	E: 121-	0	0	0	0	0	0	0	0

Totalt sett er målformen til mer enn hver fjerde elev på fådelte skoler nynorsk. Totalt går kun knapt 1 600 av elevene på fådelte skoler mot knapt 51 000 på fulldelte skoler, mens når de bruker nynorsk er det 412 på fådelte og drøyt 8 300 på fulldelte skoler.

Mens knapt 10 prosent av elevene går på skoler med under 30 elever per årstrinn utgjør disse skolene over 40 prosent av alle skolene. For de som har valgt nynorsk går hver femte elev på skoler med under tretti elever i gjennomsnitt per årstrinn. Disse skolene utgjør over halvparten av alle skolene der elevene bruker nynorsk som målform.

Vi må regne med at de fleste skolene der vi ikke kjenner til om skolen er fådelt eller fulldelt er en del av skolene med under 30 elever i gjennomsnitt på årstrinnet fådelt, mens resten er fulldelt.

**Tabell 8.2.4. Antall elever etter fulldelt/fådelt og antall elever**

Målform	Antall elever	Totalt	Fådelt/fulldelt		
			Full	Få	Ukjent
<b>Totalt</b>	<b>Totalt</b>	<b>61 956</b>	<b>50 929</b>	<b>1 577</b>	<b>9 450</b>
	A: 1-30	6 146	4 060	1 577	509
	B: 31-60	10 232	8 572	0	1 660
	C: 61-90	14 811	12 011	0	2 800
	D: 91-120	15 099	13 288	0	1 811
	E: 121-	15 669	12 998	0	2 671
<b>Bokmål</b>	<b>Totalt</b>	<b>52 282</b>	<b>42 570</b>	<b>1 152</b>	<b>8 560</b>
	A: 1-30	4 209	2 583	1 152	474
	B: 31-60	7 047	5 770	0	1 277
	C: 61-90	12 945	10 367	0	2 578
	D: 91-120	13 525	11 805	0	1 720
	E: 121-	14 557	12 045	0	2 512
<b>Nynorsk</b>	<b>Totalt</b>	<b>9 623</b>	<b>8 321</b>	<b>412</b>	<b>890</b>
	A: 1-30	1 924	1 477	412	35
	B: 31-60	3 147	2 764	-	383
	C: 61-90	1 866	1 644	-	222
	D: 91-120	1 574	1 483	-	91
	E: 121-	1 112	953	-	159
<b>Samisk</b>	<b>Totalt</b>	<b>51</b>	<b>38</b>	<b>13</b>	<b>-</b>
	A: 1-30	13	-	13	-
	B: 31-60	38	38	-	-
	C: 61-90	-	-	-	-
	D: 91-120	-	-	-	-
	E: 121-	-	-	-	-

### 8.3. Utvalgsplan

Vi skal først regne ut en utvalgsplan der elevene er fordelt på strata proporsjonalt med antallet elever i hvert stratum. I tabell 8.3.1 er fordelingen regnet ut med en total størrelse på utvalget lik 500. Vi ser at en slik utvalgsplan vil gi

- få nynorskelever i de fleste aktuelle strata
- nok bokmåselever for fulldelte skoler i alle strata
- få elever ellers i de andre strata

Særlig er det få elever som ville blitt trukket fra fådelte skoler.

Utvalget skal trekkes som et klynge/totrinnsutvalg etter følgende opplegg:

- På skoler med inntil 30 elever i åttende klassetrinn trekkes alle ut
- På skoler med mellom 31 og 60 elever på 8. klassetrinn trekkes halvparten
- På skoler med over 60 elever trekkes 30 elever ut

I tabell 8.3.3 er antallet skoler som trengs i hvert stratum regnet ut. Vi ser at i mange strata skal det trekkes under en skole.

**Tabell 8.3.1. Antall elever som skal trekkes etter målform, fremmedspråklige, fulldelt/fådelte og antall elever**

Målform	Antall elever	Fremmedspråklige				Ikke fremmedspråklige			
		Totalt	Full	Få	Ukjent	Totalt	Full	Få	Ukjent
<b>Totalt</b>	<b>Totalt</b>	<b>335,7</b>	<b>281,4</b>	<b>3,8</b>	<b>50,5</b>	<b>164,3</b>	<b>129,6</b>	<b>9,0</b>	<b>25,7</b>
	A: 1-30	18,7	13,2	3,8	1,7	30,9	19,5	9,0	2,4
	B: 31-60	49,8	40,7	-	9,1	32,8	28,5	-	4,3
	C: 61-90	80,8	65,6	-	15,2	38,7	31,3	-	7,4
	D: 91-120	87,4	77,9	-	9,5	34,5	29,4	-	5,1
	E: 121-	99,0	83,9	-	15,1	27,4	21,0	-	6,5
<b>Bokmål</b>	<b>Totalt</b>	<b>290,2</b>	<b>239,4</b>	<b>3,4</b>	<b>47,4</b>	<b>131,7</b>	<b>104,2</b>	<b>5,9</b>	<b>21,7</b>
	A: 1-30	12,8	7,8	3,4	1,6	21,2	13,1	5,9	2,2
	B: 31-60	35,4	28,1	-	7,3	21,5	18,5	-	3,0
	C: 61-90	70,3	56,3	-	13,9	34,2	27,3	-	6,9
	D: 91-120	80,5	70,9	-	9,5	28,7	24,4	-	4,3
	E: 121-	91,3	76,3	-	15,1	26,1	21,0	-	5,2
<b>Nynorsk</b>	<b>Totalt</b>	<b>45,1</b>	<b>41,7</b>	<b>0,3</b>	<b>3,1</b>	<b>32,6</b>	<b>25,5</b>	<b>3,0</b>	<b>4,1</b>
	A: 1-30	5,9	5,5	0,3	0,1	9,7	6,5	3,0	0,2
	B: 31-60	14,1	12,3	-	1,8	11,3	10,0	-	1,3
	C: 61-90	10,5	9,3	-	1,2	4,5	4,0	-	0,5
	D: 91-120	6,9	6,9	-	-	5,8	5,0	-	0,7
	E: 121-	7,7	7,7	-	-	1,3	-	-	1,3
<b>Samisk</b>	<b>Totalt</b>	<b>0,3</b>	<b>0,3</b>	<b>0,0</b>	<b>-</b>	<b>0,1</b>	<b>-</b>	<b>0,1</b>	<b>-</b>
	A: 1-30	0,0	-	0,0	-	0,1	-	0,1	-
	B: 31-60	0,3	0,3	-	-	-	-	-	-
	C: 61-90	-	-	-	-	-	-	-	-
	D: 91-120	-	-	-	-	-	-	-	-
	E: 121-	-	-	-	-	-	-	-	-

**Tabell 8.3.2. Antall elever som skal trekkes etter fulldelt/fådelte og antall elever.**

Målform	Antall elever	Totalt	Fådelte/fulldelt		
			Full	Få	Ukjent
<b>Totalt</b>	<b>Totalt</b>	<b>500,0</b>	<b>411,0</b>	<b>12,7</b>	<b>76,3</b>
	A: 1-30	49,6	32,8	12,7	4,1
	B: 31-60	82,6	69,2	-	13,4
	C: 61-90	119,5	96,9	-	22,6
	D: 91-120	121,9	107,2	-	14,6
	E: 121-	126,4	104,9	-	21,6
<b>Bokmål</b>	<b>Totalt</b>	<b>421,9</b>	<b>343,6</b>	<b>9,3</b>	<b>69,1</b>
	A: 1-30	34,0	20,8	9,3	3,8
	B: 31-60	56,9	46,6	-	10,3
	C: 61-90	104,5	83,7	-	20,8
	D: 91-120	109,2	95,3	-	13,9
	E: 121-	117,4	97,2	-	20,3
<b>Nynorsk</b>	<b>Totalt</b>	<b>77,7</b>	<b>67,2</b>	<b>3,3</b>	<b>7,2</b>
	A: 1-30	15,6	11,9	3,3	0,3
	B: 31-60	25,4	22,3	-	3,1
	C: 61-90	15,0	13,3	-	1,8
	D: 91-120	12,7	12,0	-	0,7
	E: 121-	9,0	7,7	-	1,3
<b>Samisk</b>	<b>Totalt</b>	<b>0,4</b>	<b>0,3</b>	<b>0,1</b>	<b>-</b>
	A: 1-30	0,1	-	0,1	-
	B: 31-60	0,3	0,3	-	-
	C: 61-90	-	-	-	-
	D: 91-120	-	-	-	-
	E: 121-	-	-	-	-

**Tabell 8.3.3. Antall skoler som skal trekkes etter målform, fremmedspråklige, fulldelt/fådelt og antall elever**

Mål- form	Antall elever	Fremmedspråklige				Ikke fremmedspråklige			
		Totalt	Full	Få	Ukjent	Totalt	Full	Få	Ukjent
<b>Totalt</b>	<b>Totalt</b>	<b>12,7</b>	<b>10,3</b>	<b>0,5</b>	<b>1,9</b>	<b>7,9</b>	<b>5,2</b>	<b>1,5</b>	<b>1,2</b>
	A: 1-30	1,3	0,7	0,5	0,2	3,0	1,1	1,5	0,4
	B: 31-60	2,2	1,8	-	0,4	1,5	1,3	0,0	0,2
	C: 61-90	2,7	2,2	-	0,5	1,3	1,0	-	0,2
	D: 91-120	2,9	2,6	-	0,3	1,1	1,0	-	0,2
	E: 121-	3,3	2,8	-	0,5	0,9	0,7	-	0,2
<b>Bokmål</b>	<b>Totalt</b>	<b>10,6</b>	<b>8,4</b>	<b>0,4</b>	<b>1,8</b>	<b>6,1</b>	<b>4,0</b>	<b>1,1</b>	<b>1,0</b>
	A: 1-30	1,0	0,4	0,4	0,2	2,2	0,8	1,1	0,3
	B: 31-60	1,5	1,2	-	0,3	1,0	0,8	-	0,1
	C: 61-90	2,3	1,9	-	0,5	1,1	0,9	-	0,2
	D: 91-120	2,7	2,4	-	0,3	1,0	0,8	-	0,1
	E: 121-	3,0	2,5	-	0,5	0,9	0,7	-	0,2
<b>Nynorsk</b>	<b>Totalt</b>	<b>1,8</b>	<b>1,7</b>	<b>0,0</b>	<b>0,1</b>	<b>1,7</b>	<b>1,1</b>	<b>0,4</b>	<b>0,2</b>
	A: 1-30	0,3	0,3	0,0	0,0	0,8	0,4	0,4	0,0
	B: 31-60	0,7	0,6	-	0,1	0,5	0,5	-	0,1
	C: 61-90	0,4	0,3	-	0,0	0,2	0,1	-	0,0
	D: 91-120	0,2	0,2	-	-	0,2	0,2	-	0,0
	E: 121-	0,3	0,3	-	-	0,0	-	-	0,0
<b>Samisk</b>	<b>Totalt</b>	<b>0,0</b>	<b>0,0</b>	<b>-</b>	<b>-</b>	<b>0,0</b>	<b>-</b>	<b>0,0</b>	<b>-</b>
	A: 1-30	-	-	-	-	0,0	-	0,0	-
	B: 31-60	0,0	0,0	-	-	-	-	-	-
	C: 61-90	-	-	-	-	-	-	-	-
	D: 91-120	-	-	-	-	-	-	-	-
	E: 121-	-	-	-	-	-	-	-	-

**Tabell 8.3.4. Antall skoler som skal trekkes etter fulldelt/fådelt og antall elever.**

Målform	Antall elever	Totalt	Fådelt/fulldelt		
			Full	Få	Ukjent
<b>Totalt</b>	<b>Totalt</b>	<b>20,6</b>	<b>15,5</b>	<b>2,0</b>	<b>3,1</b>
	A: 1-30	4,3	1,8	2,0	0,5
	B: 31-60	3,7	3,1	0,0	0,6
	C: 61-90	4,0	3,2	-	0,8
	D: 91-120	4,1	3,6	-	0,5
	E: 121-	4,2	3,5	-	0,7
<b>Bokmål</b>	<b>Totalt</b>	<b>16,7</b>	<b>12,4</b>	<b>1,5</b>	<b>2,8</b>
	A: 1-30	3,2	1,1	1,5	0,5
	B: 31-60	2,5	2,1	-	0,5
	C: 61-90	3,5	2,8	-	0,7
	D: 91-120	3,6	3,2	-	0,5
	E: 121-	3,9	3,2	-	0,7
<b>Nynorsk</b>	<b>Totalt</b>	<b>3,5</b>	<b>2,8</b>	<b>0,4</b>	<b>0,3</b>
	A: 1-30	1,1	0,6	0,4	0,0
	B: 31-60	1,2	1,1	-	0,1
	C: 61-90	0,5	0,4	-	0,1
	D: 91-120	0,4	0,4	-	0,0
	E: 121-	0,3	0,3	-	0,0
<b>Samisk</b>	<b>Totalt</b>	<b>0,0</b>	<b>0,0</b>	<b>0,0</b>	<b>-</b>
	A: 1-30	0,0	-	0,0	-
	B: 31-60	0,0	0,0	-	-
	C: 61-90	-	-	-	-
	D: 91-120	-	-	-	-
	E: 121-	-	-	-	-

I tabell 8.3.5 er utvalgsplanen vi skal bruke for å trekke skoler stilt opp. Fra tabell 8.3.4 er det gjort fire grep:

- Antallet nynorsk skoler er økt til 6 fra 3,5
- Antallet bokmålsskoler er redusert fra 16,7 til 16
- Samiske skoler er fjernet
- Fem strata er fjernet fra trekkgrunnet, de er markert med '0' i tabellen

**Tabell 8.3.5. Justert utvalgsplan over antall skoler som skal trekkes etter fulldelt/fådelt og antall elever.**

Målform	Antall elever	Totalt	Fådelt/fulldelt		
			Full	Få	Ukjent
<b>Totalt</b>	<b>Totalt</b>	<b>22</b>	<b>15</b>	<b>3</b>	<b>4</b>
	A: 1-30	6	2	3	0
	B: 31-60	3	4	-	0
	C: 61-90	5	4	-	1
	D: 91-120	5	4	-	1
	E: 121-	3	1	-	1
<b>Bokmål</b>	<b>Totalt</b>	<b>16</b>	<b>10</b>	<b>2</b>	<b>4</b>
	A: 1-30	3	1	2	0
	B: 31-60	3	3	-	0
	C: 61-90	5	3	-	2
	D: 91-120	3	2	-	1
	E: 121-	2	1	-	1
<b>Nynorsk</b>	<b>Totalt</b>	<b>6</b>	<b>5</b>	<b>1</b>	<b>0</b>
	A: 1-30	2	1	1	0
	B: 31-60	1	1	-	0
	C: 61-90	1	1	-	0
	D: 91-120	2	2	-	0
	E: 121-	0	0	-	0

Trekkingen av skolene ble gjennomført ved at skolene ble

- inndelt i strata med hensyn på målform, antall elever i gjennomsnitt (fem grupper) og få-/fulldelt - i alt 14 strata
- sortert etter morsmålsundervisning (ja/nei), fylke og tilfeldig innen hvert fylke

Det ble trukket ut 22 skoler med hver sin 1. og 2. reserve i samme strata og med samme status på morsmålsundervisning og i samme fylke i tillegg. Disse er listet opp i vedlegget.

I tabell 8.3.6 er antallet elever vi forventer blir med i undersøkelsen fra de 22 utrukne skolene fordelt etter samme kriterier som i tabell 8.3.5.

**Tabell 8.3.6. Oversikt over elevene som skal være med i undersøkelsen fordelt etter målform, fulldelt/fådelte og antall elever ut fra de 22 skolene som er trukket ut.**

Målform	Antall elever	Totalt	Fådelte/fulldelt		
			Full	Få	Ukjent
<b>Totalt</b>	<b>Totalt</b>	<b>521</b>	<b>373</b>	<b>27</b>	<b>120</b>
	A: 1-30	<b>50</b>	23	27	0
	B: 31-60	<b>81</b>	81	-	0
	C: 61-90	<b>180</b>	120	-	60
	D: 91-120	<b>150</b>	120	-	30
	E: 121-	<b>60</b>	30	-	30
<b>Bokmål</b>	<b>Totalt</b>	<b>386</b>	<b>250</b>	<b>16</b>	<b>120</b>
	A: 1-30	<b>21</b>	5	16	0
	B: 31-60	<b>65</b>	65	-	0
	C: 61-90	<b>150</b>	90	-	60
	D: 91-120	<b>90</b>	60	-	30
	E: 121-	<b>60</b>	30	-	30
<b>Nynorsk</b>	<b>Totalt</b>	<b>135</b>	<b>123</b>	<b>12</b>	<b>0</b>
	A: 1-30	<b>29</b>	1	2	0
	B: 31-60	<b>16</b>	1	-	0
	C: 61-90	<b>30</b>	1	-	0
	D: 91-120	<b>60</b>	1	-	0
	E: 121-	<b>0</b>	0	-	0

#### 8.4. Oversikt over utvalget med 1. og 2. reserver

Navn	Målform	Antall elever -snitt	Full/fådelte	Morsmål	Elever intervju	Løpe- nummer
Sira Skole	B	A:1-30	full	nei	5	1
Mokollen skole	B	A:1-30	få	nei	9	2
Oksfjord skole	B	A:1-30	få	nei	7	3
Hauge skole	B	B:31-60	full	ja	24	4
Våler Ungdomsskole	B	B:31-60	full	nei	22	5
Kautokeino ungdomsskole	B	B:31-60	full	nei	19	6
Groruddalen skole	B	C:61-90	full	ja	30	7
St.Svithun skole	B	C:61-90	full	ja	30	8
Reinsvoll Ungdomsskole	B	C:61-90	full	nei	30	9
Vigvoll skole	B	C:61-90	ukje	ja	30	10
Langnes skole	B	C:61-90	ukje	nei	30	11
Kråkerøy Ungdomsskole	B	D:91-120	full	ja	30	12
Håvåsen skole	B	D:91-120	full	ja	30	13
Råde Ungdomsskole	B	D:91-120	ukje	nei	30	14
Torstad ungdomsskole	B	E:121-150	full	ja	30	15
Rosenborg skole	B	E:121-150	ukje	ja	30	16
Hauso barne- og ungdomsskule	N	A:1-30	full	nei	17	17
Brekke oppvekstsenter	N	A:1-30	få	nei	12	18
Samnanger ungdomsskule	N	B:31-60	full	nei	16	19
Øygarden ungdomsskule	N	C:61-90	full	ja	30	20
Sykkylven ungdomsskule	N	D:91-120	full	ja	30	21
Nærbø ungdomsskule	N	D:91-120	full	nei	30	22
<b>1. reserver</b>						
Samfundets Skole, Sentrum	B	A:1-30	full	nei	27	101
Sentrum skole	B	A:1-30	få	nei	2	102
Tromvik Skole	B	A:1-30	få	nei	2	103
Harestad skole	B	B:31-60	full	ja	26	104
Brøttum Skole	B	B:31-60	full	nei	25	105
Storsteinnes Skole	B	B:31-60	full	nei	15	106
Majorstuen skole	B	C:61-90	full	ja	30	107
Lindesnes Ungdomsskole	B	C:61-90	full	ja	30	108
Søndre Land Ungdomsskole	B	C:61-90	full	nei	30	109
Karuss skole	B	C:61-90	ukje	ja	30	110
Saltvern skole	B	C:61-90	ukje	nei	30	111
Tananger ungdomsskole	B	D:91-120	full	ja	30	112
Skaun ungdomsskole	B	D:91-120	ukje	ja	30	113
Selsbakk skole	B	E:121-150	full	ja	30	114
Ringstabekk Skole	B	E:121-150	full	ja	30	115
Gimle skole	B	E:121-150	ukje	ja	30	116
Jondal skule	N	A:1-30	full	nei	12	117
Våge skule	N	A:1-30	få	nei	9	118
Øystese ungdomsskule	N	B:31-60	full	nei	22	119
Ulveset skule	N	C:61-90	full	ja	30	120
Førde ungdomsskule	N	D:91-120	full	ja	30	121
Jørpeland ungdomsskole	N	D:91-120	full	nei	30	122



Navn	Målform	Antall elever -snitt	Full/fådelte	Morsmål	Elever intervju	Løpe- nummer
<b>2. reserver</b>						
Steinerskolen i Kristiansand	B	A:1-30	full	nei	8	201
Elihu Kristne Grunnskole	B	A:1-30	få	nei	10	202
Stonglandet skole	B	A:1-30	få	nei	12	203
Ny Krohnborg skole	B	B:31-60	full	ja	25	204
Åmot Ungdomsskole	B	B:31-60	full	nei	23	205
Håvåsen skole	B	C:61-90	full	ja	30	206
Høyenhall Skole	B	C:61-90	full	ja	30	207
Kjøkkelvik skole	B	C:61-90	full	ja	30	208
Søndre Modum Ungdomsskole	B	C:61-90	full	nei	30	209
Tastarustå skole	B	C:61-90	ukje	ja	30	210
Sandfallet ungdomsskole	B	C:61-90	ukje	nei	30	211
Tindlund ungdomsskole	B	D:91-120	full	ja	30	212
Gosen skole	B	D:91-120	full	ja	30	213
Risum Ungdomsskole	B	D:91-120	ukje	nei	30	214
Stav skole	B	E:121-150	full	ja	30	215
Blussuvoll skole	B	E:121-150	ukje	ja	30	216
Lægreid skule	N	A:1-30	full	nei	17	217
Eivindvik skule	N	A:1-30	få	nei	9	218
Hafslo barne- og ung.skule	N	B:31-60	full	nei	15	219
Halbrend skule	N	C:61-90	full	ja	30	220
Ulstein ungdomsskule	N	D:91-120	full	ja	30	221
Klepp ungdomsskule	N	D:91-120	full	nei	30	222

## 9. Industriavfallsundersøkelsen 2003. Et stratifisert systematisk utvalg

### 9.1. Populasjon

Populasjonen består av aktive bedrifter med næringskode 10-37 i Standard for næringsgruppering. Populasjonsfila er et uttrekk fra Bedrifts- og Foretaksregisteret, 1. situasjonsuttak i desember 2002. Det er følgende begrensninger:

- Næring: 10-37
- Status: B, dvs. aktiv
- Bedriftstype:
  - 01 bedrift i enbedriftsforetak
  - 02 bedrift i flerbedriftsforetak
  - 04 hjelpebedrift
- Oppdatering: Har omsetning og/eller sysselsatte datert til 1998 eller seinere eller har ansatte > 0

Populasjonsfila inneholder 23 383 bedrifter. Før trekking ble følgende bedrifter fjernet:

- Bedrifter som verken har sysselsatte eller ansatte (2209 bedrifter)
- Bedrifter med konkursermeldinger (status\_foretak lik F eller S, 575 bedrifter). Vi har ikke fjernet disse fra totaltellingene. Det gjelder 4 bedrifter.

Av de 20 662 trekkbare, er det 19 077 som har registrert omsetning større enn 0.

### 9.2. Utvalgsplan

Det skal trekkes 1 800 bedrifter.

#### 9.2.1. Trekkemetode

Vi deler inn i strata etter

- 2-siffer næring

Næring 16 (har 10 bedrifter) slås sammen med næring 15. Næring 10 og 13 har også veldig få bedrifter (henholdsvis 13 og 9), men vi har latt dem være for seg selv, siden disse næringene er med for første gang. Vi har dermed 26 næringsgrupper.

- 5 størrelsesgrupper (ansatte brukes der sysselsetting mangler)

0	-	0 sysselsatte (trekkes ikke)
1	-	1 - 9 sysselsatte
2	-	10 - 19 sysselsatte
3	-	20 - 49 sysselsatte
4	-	50 - 99 sysselsatte
5	-	100 eller flere sysselsatte

Innen hvert stratum (utenom totaltellingene) sorteres bedriftene etter fylke og trekkes systematisk med konstant trekkesannsynlighet.

#### 9.2.2. Allokering

Utvalgsstørrelse pr. stratum skal beregnes slik at strata som har størst betydning for total avfallsmengde inneholder størst antall bedrifter, men samtidig må alle strata bestå av et rimelig antall bedrifter. Omsetning brukes her som hjelpevariabel for avfallsmengde.

- 1) Bedrifter uten sysselsatte eller ansatte trekkes ikke
- 2) Alle bedrifter med minst 100 sysselsatte skal være med i utvalget. I følgende næringer senkes denne grensen til 50: 20, 21, 24, 26, 27, 32, 35.
- 3) Bedrifter med omsetning over 10% av næringens totale omsetning skal være i utvalget (dette gjelder 5 bedrifter som ikke allerede er totaltelling pga. sysselsettingen)
- 4) Utvalgsantallet (1800) med fradrag av antall store bedrifter deles på 2.
  - a) Den ene halvdel fordeles likt på 97 strata (= 26 næringsgrupper x 4 størrelsesgrupper minus 7 fulltellingsstrata med størrelsesgruppe 4)
  - b) den andre halvdel fordeles på strata proporsjonalt med omsetning
- 5) Antall bedrifter fra forrige undersøkelse fordeles innen hvert stratum slik at:
  - alle kommer med hvis forrige utvalg har færre bedrifter enn  $\frac{1}{4}$  av nytt utvalg
  - halvparten kommer med hvis forrige utvalg har færre bedrifter enn  $\frac{1}{2}$  av nytt utvalg
  - fjerdeparten kommer med hvis forrige utvalg har færre bedrifter enn nytt utvalg
  - antall fra forrige utvalg er lik fjerdeparten av ny utvalgsstørrelse hvis forrige utvalg har flere bedrifter enn nytt utvalg

Det resterende antallet bedrifter trekkes blant bedrifter som ikke var med i 99 - utvalget. Hvis populasjonen er så liten i forhold til forrige utvalg at det ikke kan trekkes tilstrekkelig antall nye enheter, økes antall bedrifter fra forrige utvalg slik at antall bedrifter i utvalget blir som bestemt under pkt. 4 for hvert stratum.

Resultat av allokering gjengis i tabell 9.2.1. I tabellene 9.2.2 og 9.2.3 har vi aggregert tabell 9.2.1 opp til henholdsvis sysselsettingsgrupper og næringer. Populasjonstallene refererer seg til trekkepopulasjonen, altså bedrifter med sysselsatte eller ansatte og uten konkursermeldinger. De 5 totaltellingene utenfor totaltellingsstrata kommer ikke med her. De 5 fordeler seg på næring og sysselsettingsgruppe på følgende måte:

Næring\syssgr	3	4
13	0	2
19	1	1
30	1	0

**Tabell 9.2.1 - Resultat av allokering på strata**

Næring	Syssgr	Antall i trekkepop	Omsetning pop	Antall i utvalg99	Antall i utvalg 03	Antall fra 99 med i 03	Utvalgsandel (prosent)
Alle	Alle	20 657	523 637 161	1 282	1 795	740	8,7
10	1	11	29 142	0	6	0	54,5
10	5	1	218 204	0	1	0	100,0
11	1	228	428 630	0	7	0	3,1
11	2	25	231 895	0	6	0	24,0
11	3	33	3 830 505	0	17	0	51,5
11	4	23	1 175 228	0	9	0	39,1
11	5	44	12 231 958	0	44	0	100,0
13	1	3	4 257	0	3	0	100,0
13	2	1	497	0	1	0	100,0
13	3	1	0	0	1	0	100,0
13	5	1	348 084	0	1	0	100,0
14	1	529	1 614 990	0	10	0	1,9
14	2	54	1 348 095	0	9	0	16,7
14	3	26	1 219 380	0	9	0	34,6
14	4	4	379 172	0	4	0	100,0
14	5	3	926 923	0	3	0	100,0
15	1	1 256	7 114 811	18	26	5	2,1
15	2	369	8 441 945	18	30	5	8,1
15	3	369	23 283 255	62	72	16	19,5
15	4	148	23 440 011	68	72	17	48,6
15	5	101	56 462 022	80	101	80	100,0
17	1	573	815 411	5	8	1	1,4
17	2	49	703 355	6	8	2	16,3
17	3	40	1 191 603	10	9	2	22,5
17	4	11	689 603	7	8	2	72,7
17	5	8	981 468	7	8	7	100,0
18	1	427	336 325	5	7	1	1,6
18	2	15	102 081	4	6	1	40,0
18	3	13	456 713	6	7	2	53,8
18	4	5	292 736	5	5	5	100,0
19	1	67	78 704	3	6	2	9,0
19	2	4	32 291	.	4	0	100,0
19	3	1	41 974	3	1	3	100,0
19	5	1	204 901	1	1	1	100,0
20	1	1 664	2 974 162	12	14	3	0,8
20	2	183	2 690 230	14	13	3	7,1
20	3	111	4 555 871	14	19	4	17,1
20	4	39	4 038 412	18	39	18	100,0
20	5	21	4 800 784	20	21	20	100,0
21	1	61	131 758	4	6	1	9,8
21	2	13	263 726	5	6	1	46,2
21	3	21	1 459 720	7	10	2	47,6
21	4	7	843 222	5	7	5	100,0
21	5	28	18 037 662	25	28	25	100,0

Næring	Syssgr	Antall i trekkepop	Omsetning pop	Antall i utvalg99	Antall i utvalg 03	Antall fra 99 med i 03	Utvalgsandel (prosent)
22	1	2 759	5 735 185	16	22	4	0,8
22	2	265	4 067 013	11	17	3	6,4
22	3	178	5 346 730	19	21	5	11,8
22	4	49	3 978 176	19	17	4	34,7
22	5	51	16 018 799	37	51	37	100,0
23	1	14	14 054	1	6	1	42,9
23	2	1	0	.	1	0	100,0
23	3	1	1 222 019	1	1	1	100,0
23	5	2	37 647 101	2	2	2	100,0
24	1	157	464 722	6	7	2	4,5
24	2	29	755 729	8	8	2	27,6
24	3	33	2 208 128	9	12	2	36,4
24	4	20	2 682 353	8	20	8	100,0
24	5	47	29 523 722	35	47	35	100,0
25	1	304	776 864	8	8	2	2,6
25	2	65	1 091 523	5	9	1	13,8
25	3	56	2 231 444	12	12	3	21,4
25	4	20	1 752 822	11	11	3	55,0
25	5	12	1 604 961	10	12	10	100,0
26	1	734	2 785 077	12	14	3	1,9
26	2	94	1 738 451	9	11	2	11,7
26	3	62	2 486 213	9	13	2	21,0
26	4	30	3 328 470	15	30	15	100,0
26	5	14	3 718 340	13	14	13	100,0
27	1	110	416 032	5	7	1	6,4
27	2	24	419 156	5	7	1	29,2
27	3	19	678 357	7	8	2	42,1
27	4	13	2 039 275	11	13	11	100,0
27	5	38	43 954 177	34	38	34	100,0
28	1	1 641	2 839 469	14	14	4	0,9
28	2	253	2 576 818	12	13	3	5,1
28	3	201	5 673 123	26	22	6	10,9
28	4	74	5 031 114	22	20	5	27,0
28	5	23	4 805 819	17	23	17	100,0
29	1	1 977	3 862 825	12	17	3	0,9
29	2	193	3 028 204	9	14	2	7,3
29	3	140	6 936 339	14	25	4	17,9
29	4	58	6 100 508	20	23	5	39,7
29	5	39	12 872 982	29	39	29	100,0
30	1	40	98 436	4	6	1	15,0
30	2	4	73 068	1	4	1	100,0
30	5	1	865 231	1	1	1	100,0
31	1	382	931 524	8	8	2	2,1
31	2	49	732 444	6	8	2	16,3
31	3	42	1 297 499	4	9	2	21,4
31	4	14	1 389 667	7	10	2	71,4
31	5	23	7 488 963	23	23	23	100,0

Næring	Syssgr	Antall i trekkepop	Omsetning pop	Antall i utvalg99	Antall i utvalg 03	Antall fra 99 med i 03	Utvalgsandel (prosent)
32	1	95	137 741	4	6	1	6,3
32	2	6	98 900	2	6	2	100,0
32	3	19	776 864	6	8	2	42,1
32	4	10	1 267 847	5	10	5	100,0
32	5	16	9 207 332	13	16	13	100,0
33	1	394	755 930	7	8	2	2,0
33	2	49	650 115	5	7	1	14,3
33	3	34	1 123 676	9	9	2	26,5
33	4	10	1 544 383	4	10	4	100,0
33	5	12	5 056 194	9	12	9	100,0
34	1	75	313 461	4	7	1	9,3
34	2	29	406 708	5	7	1	24,1
34	3	19	746 205	9	8	2	42,1
34	4	8	404 858	5	7	1	87,5
34	5	16	4 804 384	14	16	14	100,0
35	1	708	1 613 233	8	10	2	1,4
35	2	105	1 791 481	6	11	2	10,5
35	3	126	3 653 049	17	16	4	12,7
35	4	52	3 705 184	20	52	20	100,0
35	5	88	37 406 896	65	88	65	100,0
36	1	1412	1 689 926	7	10	2	0,7
36	2	107	1 217 938	7	9	2	8,4
36	3	94	2 800 765	11	14	3	14,9
36	4	31	2 152 852	14	12	3	38,7
36	5	19	4 110 305	16	19	16	100,0
37	1	131	702 820	2	8	2	6,1
37	2	23	391 712	3	7	2	30,4
37	3	20	1 148 984	5	9	1	45,0
37	4	2	216 776	1	2	1	100,0

**Tabell 9.2.2 - Fordeling på sysselsettingsgrupper**

Syssgr	Antall i trekkepop	Omsetning pop	Antall i utvalg99	Antall i utvalg 03	Antall fra 99 med i 03	Utvalgsandel (prosent)
1	15 752	36 665 489	165	251	46	1,6
2	2 009	32 853 375	141	222	39	11,1
3	1 659	74 368 416	260	332	70	20,0
4	628	66 452 669	265	381	134	60,7
5	609	313 297 212	451	609	451	100,0

**Tabell 9.2.3 - Fordeling på næring**

Næring	Antall i trekkepop	Omsetning pop	Antall i utvalg99	Antall i utvalg 03	Antall fra 99 med i 03	Utvalgsandel (prosent)
10	12	247 346	0	7	0	58,3
11	353	17 898 216	0	83	0	23,5
13	6	352 838	0	6	0	100,0
14	616	5 488 560	0	35	0	5,7
15	2 243	118 742 044	246	301	123	13,4
17	681	4 381 440	35	41	14	6,0
18	460	1 187 855	20	25	9	5,4
19	73	357 870	7	12	6	16,4
20	2 018	19 059 459	78	106	48	5,3
21	130	20 736 088	46	57	34	43,8
22	3 302	35 145 903	102	128	53	3,9
23	18	38 883 174	4	10	4	55,6
24	286	35 634 654	66	94	49	32,9
25	457	7 457 614	46	52	19	11,4
26	934	14 056 551	58	82	35	8,8
27	204	47 506 997	62	73	49	35,8
28	2 192	20 926 343	91	92	35	4,2
29	2 407	32 800 858	84	118	43	4,9
30	45	1 036 735	6	11	3	24,4
31	510	11 840 097	48	58	31	11,4
32	146	11 488 684	30	46	23	31,5
33	499	9 130 298	34	46	18	9,2
34	147	6 675 616	37	45	19	30,6
35	1 079	48 169 843	116	177	93	16,4
36	1 663	11 971 786	55	64	26	3,8
37	176	2 460 292	11	26	6	14,8

### 9.3. Utvalg - dekningsgrad

Utvalget trukket etter planen ovenfor dekker ca. 71 % av total omsetning og ca. 62 % av total sysselsetting i populasjonen. Dekningsgrad fordelt på sysselsettingsgrupper, næring og fylke er gitt i tabellene 9.3.1-2-3 (her er de med konkursanmerkninger tatt med i populasjonen, men ikke de uten sysselsatte/ansatte). Det er 4 bedrifter i utvalget (totaltelling) med konkursanmerkninger.

**Tabell 9.3.1 - Dekningsgrader etter næring**

Næring		Andel bedrifter	Andel sysselsetting	Andel omsetning
<b>Alle</b>		0,09	0,62	0,71
<b>10</b>	Bryting av steinkull og brunkull. Utvinning av torv	0,58	0,94	0,94
<b>11</b>	Utvinning av råolje og naturgass. Tjenester tilknyttet olje- og gassutvinning	0,23	0,88	0,90
<b>13</b>	Bryting av metallholdig malm	1,00	1,00	1,00
<b>14</b>	Bergverksdrift ellers	0,06	0,32	0,37
<b>15/16</b>	Produksjon av næringsmidler og drikkevarer Produksjon av tobakksvarer	0,13	0,55	0,61
<b>17</b>	Produksjon av tekstiler	0,06	0,40	0,43
<b>18</b>	Produksjon av klær. Beredning og farging av pelskinn	0,05	0,39	0,51
<b>19</b>	Beredning av lær. Produksjon av reiseeffekter, salmakerartikler og skotøy	0,18	0,67	0,80
<b>20</b>	Produksjon av trelast og varer av tre, kork, strå og flettematerialer, unntatt møbler	0,05	0,44	0,51
<b>21</b>	Produksjon av papirmasse, papir og papirvarer	0,43	0,94	0,96
<b>22</b>	Forlagsvirksomhet, grafisk produksjon og reproduksjon av innspilte opptak	0,04	0,56	0,53
<b>23</b>	Produksjon av kull og petroleumsprodukter og kjernebrensel	0,53	0,99	1,00
<b>24</b>	Produksjon av kjemikalier og kjemiske produkter	0,32	0,90	0,93
<b>25</b>	Produksjon av gummi og plastprodukter	0,11	0,43	0,40
<b>26</b>	Produksjon av andre ikke-metallholdige mineralprodukter	0,09	0,53	0,55
<b>27</b>	Produksjon av metaller	0,35	0,93	0,98
<b>28</b>	Produksjon av metallvarer, unntatt maskiner og utstyr	0,04	0,27	0,32
<b>29</b>	Produksjon av maskiner og utstyr	0,05	0,47	0,49
<b>30</b>	Produksjon av kontor- og datamaskiner	0,25	0,83	0,89
<b>31</b>	Produksjon av andre elektriske maskiner og apparater	0,11	0,65	0,71
<b>32</b>	Produksjon av radio-, fjernsyns- og annet kommunikasjonsutstyr	0,31	0,90	0,95
<b>33</b>	Produksjon av medisinske instrumenter, presisjonsinstrumenter, optiske instrumenter, klokker og ur	0,09	0,60	0,77
<b>34</b>	Produksjon av motorkjøretøyer, tilhengere og deler	0,30	0,84	0,84
<b>35</b>	Produksjon av andre transportmidler	0,16	0,82	0,86
<b>36</b>	Produksjon av møbler. Annen industriproduksjon	0,04	0,38	0,44
<b>37</b>	Gjenvinning	0,14	0,35	0,33



**Tabell 9.3.2 - Dekningsgrader etter sysselsettingsgruppe**

Sysselsettings- gruppe	Andel bedrifter	Andel sysselsetting	Andel omsetning
1	0,02	0,02	0,02
2	0,11	0,11	0,13
3	0,20	0,20	0,25
4	0,61	0,60	0,58
5	1,00	1,00	1,00

**Tabell 9.3.3 - Dekningsgrader etter fylke**

Fylke	Andel bedrifter	Andel sysselsetting	Andel omsetning
01	0,09	0,63	0,76
02	0,07	0,60	0,67
03	0,06	0,66	0,74
04	0,08	0,55	0,62
05	0,07	0,55	0,52
06	0,09	0,64	0,74
07	0,09	0,57	0,80
08	0,08	0,62	0,79
09	0,08	0,57	0,65
10	0,09	0,64	0,81
11	0,11	0,69	0,74
12	0,09	0,64	0,78
14	0,11	0,61	0,71
15	0,10	0,53	0,60
16	0,08	0,54	0,68
17	0,09	0,50	0,61
18	0,09	0,52	0,60
19	0,06	0,35	0,41
20	0,07	0,43	0,38
21	0,20	0,97	0,99
23	0,34	0,98	1,00

## 10. Levekår i landbruket 2002. Et stratifisert systematisk utvalg, etter region og driftstype

### 10.1. Undersøkelsespopulasjon

- Bruk med minst 20 dekar jordbruksareal i drift.
- Bruk med dyr og minst 5 dekar jordbruksareal i drift.

### 10.2. Trekkepopulasjon

Utvalget ble trukket fra en foreløpig versjon av registeret over de som søkte om produksjonstilskudd 31.07.2001. Bruk utenfor undersøkelsespopulasjonen, og bruk utenfor SSBs intervjuerområder ble fjernet før trekking.

Fjerning av bruk utenfor SSBs intervjuerområder innebærer en betydelig forskjell mellom undersøkelsespopulasjon og trekkepopulasjon. Bare 37 prosent av brukene er trekkbare. Dette er ikke ideelt, men intervjuing av bruk utenfor intervjuerområdene ville økt kostnadene betydelig. Vi mener at vi ved hjelp av stratifisering likevel får et forsvarlig utvalg.

### 10.3. Utvalgsplan

#### 10.3.1. Stratifisering

Populasjonen av bruk ble delt inn i 40 strata etter region og driftstype. Utvalgsstørrelsen er 2000 bruk.

**Region:** NILFs soneinndeling som lå på fila er brukt.

- 1 Østlandet, flatbygder
- 2 Østlandet, andrebygder
- 3 Jæren
- 4 Agder/Rogaland, andre bygder
- 5 Vestlandet
- 6 Trøndelag, flatbygder
- 7 Trøndelag, andrebygder
- 8 Nord-Norge

**Driftstype:** Variabelen "produksjon" som lå på fila er brukt

- 1 melk
- 2 kjøtt (storfe/svin/fjørfe)
- 3 sau
- 4 korn
- 5 annet

Det er viktig med gode tall i alle regioner, så det ble trukket like mange, dvs. ca. 250 bruk, i hver region. Innenfor hver region ble det stratifisert på driftstype med proporsjonal allokering, slik at fordelingen på driftstype ble riktig i hver region, dvs. lik fordelingen i undersøkelsespopulasjonen. Merk at fasitfordeling er den i undersøkelsespopulasjonen, ikke trekkepopulasjonen, siden det er undersøkelsespopulasjonen vi vil at utvalget skal være representativt for.

Innenfor hvert stratum ble brukene sortert etter størrelse (jordbruksareal i drift), og det ble deretter trukket systematisk. Dette gir en jevn fordeling på størrelse innen hvert stratum.

**Tabell 10.3.1 Antall bruk i hvert stratum for hele populasjonen, trekkepopulasjonen og utvalget**  
Stratum xy betyr region x og driftstype y.

Stratum	Hele populasjonen	Trekkepopulasjonen	Utvalget
11	1 430	930	28
12	844	516	17
13	888	603	18
14	6 847	3 765	134
15	2 773	1 769	55
21	3 234	748	68
22	753	203	16
23	2 701	640	57
24	2 739	654	58
25	2 551	695	54
31	1 430	739	136
32	215	117	21
33	361	158	35
34	85	55	9
35	548	247	52
41	1 903	447	85
42	598	229	27
43	2 137	737	96
44	109	41	5
45	866	271	39
51	5 099	1 441	93
52	1 354	420	25
53	5 181	1 451	95
54	67	5	2
55	2 006	506	37
61	1 535	912	90
62	400	254	24
63	324	197	19
64	1 145	731	68
65	865	589	51
71	2 235	609	125
72	341	87	19
73	956	251	54
74	298	82	17
75	662	167	37
81	2 058	612	91
82	365	112	17
83	2 194	711	97
84	4	0	0
85	1 038	340	46
<b>Totalt</b>	<b>61 139</b>	<b>23 041</b>	<b>2 017</b>

## 10.4. Utvalget

I tabellen under sammenlignes størrelsesfordelingen i utvalget og i undersøkelsespopulasjonen innenfor hver region. Frekvenstabellen for utvalget er delt på frekvenstabellen for populasjonen celle for celle, slik at når de to fordelingene er like, er alle forholdstallene under lik 1.

**Tabell 10.4.1 Størrelsesfordeling utvalg/populasjon**

Region	Størrelsesgruppe (dekar)				
	0-49	50-99	100-149	150-199	200 og over
Østlandet, flatbygder	1,13	1,07	0,99	1,03	0,94
Østlandet, andre	0,95	1,00	1,03	0,98	1,02
Jæren	1,03	1,10	1,01	0,96	0,97
Agder/Rogaland	1,09	1,05	0,83	0,93	1,05
Vestlandet (rest)	0,96	1,04	1,09	0,98	0,83
Trøndelag, flatbygder	0,96	0,94	1,00	0,97	1,04
Trøndelag, andre	0,76	0,80	0,82	0,92	1,73
Nord-Norge	0,97	1,01	1,01	0,99	1,00

Vi ser at størrelsesfordelingen i utvalget er omtrent lik størrelsesfordelingen i undersøkelsespopulasjonen, bortsett fra i regionen "Trøndelag, andre", der vi har fått med litt for mange store bruk.

## 11. Undersøkelsen om samvær og bidrag 2002

### 11.1. Undersøkelsespopulasjon

I denne undersøkelsen er det flere interessante populasjoner. De viktigste er populasjonene av bidragsmottakere og bidragspliktige, ev. delpopulasjonene kvinnelige bidragsmottakere (også kalt enslige mødre) og mannlige bidragspliktige (samværsfedre). Eventuelt kan man se på mottakere og pliktige under ett, dvs. alle voksne involvert i bidragsforhold. Populasjonen av alle "skilsmissebarn" kan også være aktuelt. Man kan også endre enhet fra person til par, og undersøke populasjonen av foreldrepar som ikke bor sammen.

### 11.2. Trekking

Først ble det etablert et trekkegrunnlag ut i fra folkeregisteret. Dette besto av alle barn som var 0-17 år pr. 31.12.2002 der mor og far er bosatte, men ikke på samme (numeriske) adresse. Barnet skulle også bo sammen med en av foreldrene. Trekkegrunnlaget ble sortert etter kommunenummer, postnummer og familienummer, og deretter ble det trukket ut 2 100 barn. Det skal bare være ett uttrukket barn per foreldrepar i utvalget, så det ble fjernet 12 barn for å unngå samme foreldre. Til slutt ble utvalget redusert til 2 000 barn.

Ut i fra utvalget av barn konstrueres det så to utvalg av voksne, nemlig mottakerne til de uttrukne barna (som skal besvare skjema for mottakere), og bidragspliktige til de uttrukne barna (som skal besvare skjema for bidragspliktige). Man kan slå de to utvalgene sammen til ett utvalg av voksne.

Begge foreldre får spørsmål om de samme tingene. For eksempel er spørsmål 1 på skjemaet for mottakere:

*Har du barn under 18 år som **du bor sammen med** til daglig, og som ikke bor sammen med den andre forelderen? Regn også med barn som bor halvparten av tiden hos hver av dere.*

For pliktige er det tilsvarende spørsmålet:

*Har du barn under 18 år som **du ikke bor sammen med**, men som bor sammen med den andre forelderen? Regn også med barn som bor halvparten av tiden hos hver av dere.*

Så blir man bedt om å liste dem opp fra yngst til eldst. Begge foreldre skal da liste opp de samme barna. Noen av de etterfølgende spørsmålene, som for eksempel hvordan samværsordningen fungerer, skal besvares for hvert enkelt barn.

### 11.3. Enheter og begreper

Det kan være nyttig å skille mellom *trekkeenhet*, *recordenhet* og *analyseenhet*. I denne undersøkelsen er trekkeenheten barn, mens recordenheten på basisfilene er voksne personer, fordi det lages en record per skjema (som er utfyllt enten av en mottaker eller pliktig). Som analyseenhet er det flere aktuelle muligheter: personer, foreldrepar, barn. I tillegg er det for personer interessant å se på delpopulasjonene mottakere og pliktige hver for seg, og også alle samlet.

Noen begreper:

*Bidragmottaker:* Foreldren det uttrukne barnet er folkeregistrert hos

*Bidragspliktig:* Den foreldren det uttrukne barnet ikke er folkeregistrert hos

*Foreldrepar:* To personer med felles barn

*Bidragsrelasjon:* En bidragsrelasjon består av et foreldrepar og de felles barn som er registrert hos en bestemt av dem. Dette betyr at foreldre som har delt et kull, er i to bidragsrelasjoner med hverandre. Begge foreldre er da mottaker i en relasjon og pliktig i den andre.

## 11.4. Trekksannsynligheter

I selve trekkingen har alle barna på trekkefila samme trekksannsynlighet, som er lik antall uttrukne barn delt på antall barn totalt på trekkefila. Kall denne trekksannsynligheten for  $\pi$ .

For de forskjellige utvalgene av voksne og par, får vi et sett av trekksannsynligheter (eller sannsynligheter for å komme med) for hvert utvalg.

### 11.4.1. Bidragmottakere/pliktige

En mottaker kommer med i utvalget av mottakere nøyaktig når et barn han/hun mottar bidrag for blir trukket ut fra trekkefila. Sannsynligheten for å være med i utvalget av mottakere er derfor lik sannsynligheten for at et av de barna man mottar bidrag for blir trukket ut. Denne sannsynligheten er (tilnærmet) lik  $\pi$  ganger antall barn i trekkegrunnlaget man er mottaker for. Dette betyr at mottakere med mange barn er overrepresentert i utvalget.

Sannsynligheten for å være med i utvalget av bidragspliktige blir på samme måte lik  $\pi$  ganger antall barn i trekkegrunnlaget man er bidragspliktig for.

### 11.4.2. Alle voksne

Sannsynligheten for å være med i dette utvalget er  $\pi$  ganger antall barn man er forelder til i trekkegrunnlaget.

Noen personer er både mottaker og pliktig. Dette kan skje hvis man har flere ekspartnere eller har splittet et kull. For disse personene vil sannsynligheten for å være med i utvalget av alle voksne være forskjellig fra sannsynligheten for å være med som mottaker/pliktig.

### 11.4.3. Par

Et foreldrepar kommer med i bruttoutvalget hvis og bare hvis et felles barn blir trukket fra trekkefila, så trekksannsynligheten for paret er  $\pi$  ganger antall felles barn på trekkefila.

### 11.4.4. Alle barn

Man kan danne et utvalg av alle barn det er svart for på skjemaene. Dette blir flere enn de opprinnelig uttrukne barna.

Mottaker (altså den forelderen som er registrert bosatt på samme sted som det uttrukne barnet) skal svare for de barn under 18 år som hun/han bor sammen med til daglig, og som ikke bor sammen med den andre forelderen, pluss barn som bor halvparten av tiden hos hver. Bidragspliktig skal svare for de samme barna, så det kommer ikke noen flere barn inn i utvalget gjennom pliktig.

Vi må her skille mellom barn som bor bare på ett sted og barn som bor halvparten hos hver forelder. Barn som bor på ett kommer med hvis de selv blir trukket ut eller hvis helsøsken registrert på samme sted blir trukket ut, men de kan ikke komme med gjennom andre helsøsken. Barn som bor halvparten hos hver kan komme med gjennom alle helsøsken. Trekk sannsynlighetene blir derfor  $\pi$  ganger (antall helsøsken registrert på samme sted + 1) for barn som bor på ett sted hele tiden, og  $\pi$  ganger (antall helsøsken + 1) for barn som bor halvparten av tiden hos hver av foreldrene.

## 11.5. Utvalg

**Tabell 11.5.1 Bruttoutvalg av bidragsmottakere og bidragspliktige etter kjønn, alder, landsdel, innvandringsstatus og antall bidragsrelasjoner. Prosent**

	Mottakere	Pliktige
<b>IOs kjønn</b>		
Mann	11	89
Kvinne	89	11
<b>Aldersgruppe</b>		
18-29 år	14	9
30-34 år	23	17
35-39 år	28	26
40-44 år	22	24
45 år el. eldre	13	24
<b>Innvandrerkategori</b>		
Uten innvandrerbakgrunn	88	86
Med innvandrerbakgrunn	12	14
<b>Landsdel</b>		
Østlandet ellers	28	28
Akershus og Oslo	21	22
Agder og Rogaland	14	13
Vestlandet	16	16
Trøndelag	9	9
Nord-Norge	12	12
<b>Antall relasjoner</b>		
1 relasjon	89	87
2 relasjoner	11	12
3 eller flere relasjoner	1	2
<b>Antall personer</b>	1 956	1 950

Tallene er tatt fra Tabell F, side 14, i Aina Helen Sætre: Undersøkelsen om samvær og bidrag 2002, Dokumentasjons- og tabellrapport, Notater 2004/26.

## 12. Referanser

- [1] Jan Bjørnstad: En innføring i utvalgsundersøkelser, Notater 2000/50
- [2] Sharon L. Lohr: Sampling: Design and Analysis, Duxbury Press 1999
- [3] Margareta Stålnacke, Jenny-Anne Sigstad Lie og Leiv Solheim: En analyse av SSBs generelle utvalgsplan fra 1995 basert på næringsvise sysselsettingstall, Notater 99/36

## De sist utgitte publikasjonene i serien Notater

- 2006/11 E.C. Rauan og R. Nyggård Johnsen: Forventningsindikator - Konsumprisene. November 2005 - mai 2006. 18s.
- 2006/12 S. Lien og Ø. Sivertstøl: Veier ut av langtidsmottak av sosialhjelp. 47s.
- 2006/13 M. Hansen-Møllerud, A. Kalvøy, G. M. Pilskog og A-H. Sølverud: Informasjonssamfunnet 2005. 49s.
- 2006/14 K.A. Kjesbu: Dokumentasjon av tidsseriebase for FoU-data. 47s.
- 2006/15 B.O. Lagerstrøm og M. Høstmark: Kultur- og mediebruksundersøkelsen 2004. Dokumentasjonsrapport. 55s.
- 2006/16 H. Skullerud: Metanutslipp fra norske avfallsfyllinger. Reviderte beregninger av deponert avfall 1945 - 2004\*. 15s.
- 2006/17 S.K. Boateng og S. Ferstad: Dokumentasjonsnotat for FylkesKOSTRA videregående opplæring. Publisering av 2004-tallene. 312.s
- 2006/18 K.I. Bøe, S. Lien og Ø. Sivertstøl: Fd-Trygd. Dokumentasjonsrapport. Demografi revidert. 1992-2003. 130s.
- 2006/19 A. Holmøy og B.O. Lagerstrøm: Interkommunalt legevaktsamarbeid - en forundersøkelse. 19s.
- 2006/20 H. Tønseth: Årsrapport for kontaktutvalget for helse- og sosialstatistikk. 19s.
- 2006/21 D. Gronna og S. Todsén: Nasjonalregnskap: Beregning av olje- og gassnæringene. 31s.
- 2006/22 D. Gronna, S. Todsén og K. Erlandsen Kolshus: Beregning av olje- og gassnæringene i KNR. 24s.
- 2006/23 T. Tveiekrem Sæter og I. T. Holmen: Prisindeks for bilutleie. 31s.
- 2006/24 Ø. Linnestad og G.M. Molseth: Forprosjekt "Godstransport på kysten". 66s.
- 2006/25 K. Loe Hansen: Indikatorer på kjemikalieområdet- Risiko for skade på helse og miljø grunnet bruk av kjemiske stoffer. 46s.
- 2006/26 A. Akselsen, S. Lien og Ø. Sivertstøl: FD-Trygd. Variabelliste. 58s.
- 2006/27 J. Heldal og A. Rusti: Om samordning av utvalg ved bruk av PRN-tall. 29s.
- 2006/28 C. Nordseth og Ø. Sivertstøl: FD - Trygd. Dokumentasjonsrapport. Fødsels- og sykepengene, 1992-2003. 134s.
- 2006/29 A. Linderud: Verdipapirstatistikk. Dokumentasjonsnotat. 54s.
- 2006/30 V.V. Holst Bloch, H. Høye, M. Steinnes og J.K. Undelstveit: Kartbasert rapportering i KOSTRA - en mulighetsstudie. 50s.
- 2006/31 E. Høydal: Monitor for sekundærflytting. En deskriptiv analyse av sekundærflyttinger blant flyktninger bosatt i Norge i 1995-2004. 67s.
- 2006/32 E.Cometa Rauan: Undersøking om foreldrebetaling i barnehagar, januar 2006. 46s.
- 2006/33 T. Skarøhamar: Kriminalitet gjennom ungdomstiden blant nordmenn og ikke-vestilige innvandrere. En analyse av fødselskullet 1977. 36s.
- 2006/34 N. Hagesæther og L-C. Zhang: Om arbeidsledighet i AKU og Arena. 19s.
- 2006/35 T. Hægeland, Lars J. Kirkebøen og Oddbjørn Raaum: Skoleresultater 2005. En kartlegging av karakterer fra grunnskoler og videregående skoler i Norge. 83s.
- 2006/36 S. Skaare: Undersøkelse om «Utbrenthet i enkelte yrker» 2005. Dokumentasjonsrapport. 68s.
- 2006/37 O.F. Vaage: Barn og unges idrettsdeltakelse og foreldres inntekt. Analyse med data fra Levekårsundersøkelsen 2004. 31s.