ARTIKLER 88

ESTIMERING AV SYSSELSETTING I GEOGRAFISKE REGIONER: OM ESTIMATORENES SKJEVHET, VARIANS OG BRUTTOVARIANS

Av
Petter Laake og Hans Kristian Langva

ESTIMATION OF EMPLOYMENT WITHIN GEOGRAPHICAL REGIONS: ON THE BIAS, VARIANCE AND THE MEAN SQUARE ERROR OF THE ESTIMATES

OSLO 1976

STATISTISK SENTRALBYRÅ

ARTIKLER FRA STATISTISK SENTRALBYRÅ NR. 88

ESTIMERING AV SYSSELSETTING I GEOGRAFISKE REGIONER: OM ESTIMATORENES SKJEVHET, VARIANS OG BRUTTOVARIANS

Av
Petter Laake og Hans Kristian Langva

ESTIMATION OF EMPLOYMENT WITHIN GEOGRAPHICAL REGIONS: ON THE BIAS, VARIANCE AND THE MEAN SQUARE ERROR OF THE ESTIMATES

OSLO 1976

ISBN 82 - 537 - 0646 - 4

FORORD

Statistisk Sentralbyrå gjennomfører regelmessige utvalgsundersøkelser. Undersøkelsene er vanligvis planlagt slik at en publiserer statistikk for bestanden i hele landet. I mange situasjoner kan det imidlertid være av interesse å publisere tall for mindre geografiske områder. I denne artikkelen har en vurdert egenskapene til estimatorer for sysselsettingen innen små geografiske områder. De numeriske beregningene er foretatt for en bestand der alle kjennemerkene er kjente. Bestanden er opprettet på grunnlag av Folke- og boligtelling 1970.

Statistisk Sentralbyrå, Oslo, 19. oktober 1976

Odd Aukrust

PREFACE

At regular intervals the Central Bureau of Statistics of Norway carries out sample surveys. Usually a survey is planned to produce statistics for the whole country. On many occasions, however, there may be a need for statistics for smaller geographical areas. The contents of this article is a study of the properties of estimates of employment within smaller geographical areas. The numerical calculations are based on a population where the characteristics are known. The population is established from the Population and Housing Census 1970.

Central Bureau of Statistics, Oslo, 19 October 1976

Odd Aukrust

INNHOLD

Side

1. Innledning 7

2. Generelt om Byråets standard utvalgsplan og om estimatorer for totalverdier i bestanden i geografiske regioner 9

2.1. Generelt om Byråets standard utvalgsplan 9

2.2. Generelt om estimatorene 11

2.3. Sammenligning av to metoder for estimering i bestanden i fylker 14

2.4. Valg av delklasser til syntetisk estimering av sysselsettingen i fylker 16

2.5. Om metoden som er valgt til å vurdere estimatorenes egenskaper 16

3. Konklusjoner 18

3.1. Generelt om de numeriske beregningene 18

3.2. Konklusjoner til de numeriske beregningene 20

4. Formell beskrivelse av Byråets standard utvalgsplan og de foreslåtte estimatorene 22

4.1. Formell beskrivelse av Byråets standard utvalgsplan 22

4.2. Estimator for totalverdien i bestanden i fylker 26

4.3. Syntetiske estimatorer 27

Engelsk sammendrag 31

Referanser 32

Vedleggstabeller 33

T a b e l l e r

1. Sysselsetting etter næring i regioner. Beregningsbestand: 10 prosent av landets befolkning 33

2. Sysselsetting etter næring i fylkene, "nedbrutt" etter folketall. Beregningsbestand: 10 prosent av landets befolkning 39

3. Sysselsetting i fylkene, "nedbrutt" etter andel i hver næring. Beregningsbestand: 10 prosent av landets befolkning 47

4. Sysselsetting etter næring i fylkene, "nedbrutt" etter alders- og kjønnsklasser. Beregningsbestand: 10 prosent av landets befolkning 48

CONTENTS

Page

1. Introduction 7

2. On the sample survey of the Bureau and on the estimates of characteristics of the population within geographical regions 9

2.1. On the sample survey 9

2.2. On the estimates 11

2.3. A comparison between two methods for estimation within counties 14

2.4. On the choice of subclasses for synthetic estimation of the employment within counties 16

2.5. On the strategy for studying the properties of the estimates 16

3. Conclusions 18

3.1. On the numerical calculations 18

3.2. Concluding remarks 20

4. A formal description of the standard sample survey of the Bureau and of the estimates 22

4.1. A description of the standard sample survey of the Bureau 22

4.2. Estimates of population characteristics within a county 26

4.3. Synthetic estimates 27

English summary 31

References 32

Tables in the appendix 33

T a b l e s

1. Employment by industry within regions. Population: 10 per cent of the population of Norway 33

2. Employment by industry within counties, based on a priori information on the number of inhabitants. Population: 10 per cent of the population of Norway 39

3. Employment within counties, based on a priori information on the employment by industry. Population: 10 per cent of the population of Norway 47

4. Employment by industry within counties, based on a priori information on sex and age distribution. Population: 10 per cent of the population of Norway 48

1. INNLEDNING

Statistisk Sentralbyrå gjennomfører regelmessige utvalgsundersøkelser. Innsamlingen av data blir foretatt av intervjuere, og Byrået har derfor laget en standard utvalgsplan som brukes ved de fleste undersøkelsene. Siden det er brukt metoder fra matematisk statistikk ved konstruksjonen av utvalgsplanen, kan en lage estimatorer for totaler og gjennomsnitt i bestanden og finne variansen til disse estimatorene. Utvalgsundersøkelsene er vanligvis planlagt på en slik måte at en kan lage estimatorer for hele den bestanden en undersøker eller eventuelt en delklasse av den. Ved Byrået's arbeidskraftsundersøkelser (AKU) lager en eksempelvis estimater på landsbasis for sysselsettingen innen de ulike næringer for hele bestanden og for menn og kvinner separat.

I de undersøkelsene en gjennomfører, ønsker en vanligvis å vurdere usikkerheten til de estimatene en publiserer. Ved Underavdelingen for intervjuundersøkelser ble det derfor i 1974 laget et program for estimering av totaler og gjennomsnitt og estimering av variansen til disse estimatorene (Laake et al. (1974)). Estimeringsprogrammet bygger på den standard utvalgsplanen som ble brukt inntil 1975. Dagsvik (1974, side 4) har anvendt programmet til beregninger av standardavviket på estimatorer for sysselsettingen i hele landet innen noen hovednæringer. I Byrået's utvalgsundersøkelser publiserer en vanligvis estimater for bestanden i hele landet. Estimeringsprogrammet, som er omtalt foran, har derfor vært brukt bare til å beregne estimater og varianser for estimatorer for hele landet.

I den siste tiden har det imidlertid blitt økt interesse for å vurdere om det er mulig å publisere estimater for bestanden innen mindre geografiske regioner. Spesielt har det vært av interesse å vurdere om en i Byrået's arbeidskraftsundersøkelser kan publisere estimater for total sysselsetting og sysselsettingen innen hovednæringene for hvert fylke.

I Byrået's nye standard utvalgsplan (som er blitt brukt siden 1975) er stratifiseringen utført på en slik måte at en ikke direkte kan lage forventningsrette estimatorer for totalverdien i bestanden i fylker. Utvalgsplanen er konstruert slik at en kan lage forventningsrette estimatorer innen geografiske områder som i det vesentlige består av par av fylker. En har til nå ikke helt utnyttet denne egenskapen ved utvalgsplanen. Dette skyldes først og fremst at en ikke har utført noen nærmere studier av variansen til estimatorer innen så små geografiske områder som par av fylker. Det kan imidlertid være av interesse å

studere om en kan publisere estimater for sysselsettingen innen slike områder. I denne artikkelen har vi kalt disse områdene for regioner, og vi har studert egenskapene til estimatorer for sysselsettingen innen noen hovednæringer i regionene. Siden estimatorene er forventningsrette, har vi begrenset oss til å studere variansen til estimatorene.

Som nevnt foran har det blitt økt interesse for å vurdere om en kan publisere estimater for sysselsettingen innen hvert fylke. I tillegg til å studere usikkerheten til tall for regioner, har vi derfor vurdert muligheten av å publisere estimater for sysselsettingen innen fylker. Siden utvalgsplanen ikke direkte gir grunnlag for estimering av sysselsettingstall for fylker, har vi valgt å konstruere nye estimatorer som bruker a priori informasjon om bestanden i fylkene. Estimatorene er karakterisert ved at en under estimeringen av sysselsettingstall sammen med data fra AKU kan bruke data om alders- og kjønnsfordelingen innen hvert fylke. Slike estimatorer kalles syntetiske estimatorer, og de har tidligere blitt studert av Gonzalez & Waksberg (1973). Metoden har tidligere blitt utprøvd av National Center of Health Statistics (1968) og av Bureau of the Census i "1970 Census of Population and Housing" (se Gonzalez & Waksberg (1973, side 9)).

For bestanden innen fylker har vi foreslått tre sett syntetiske estimatorer. Det ene settet bygger på a priori informasjon om folketallet i fylkene, det andre bygger på informasjon om sysselsettingsstrukturen og det tredje bygger på informasjon om alders- og kjønnsfordelingen i bestanden i fylkene.

De syntetiske estimatorene for totalverdien i bestanden i fylker er ikke forventningsrette. Vi er ikke kjent med at en tidligere har studert skjevheter til slike estimatorer. I denne artikkelen vil vi imidlertid studere både forventningen og variansen til estimatorene. For å kunne gjøre dette, har vi konstruert en bestand der alle enhetene er kjente.

Vanligvis studeres estimatorenes varians ved å studere estimater for variansen. Byråets standard utvalgsplan er imidlertid konstruert slik at en ikke direkte kan lage forventningsrette estimatorer for variansen. De variansestimatorene som er foreslått, vil overestimere variansen til estimatorene. Når en bruker slike variansestimatorer, kan en altså ikke skille ut hvilken effekt som skyldes overestimering av variansen og hvilken som skyldes skjevhet i selve estimatoren for fylkestall. Vi ønsker derfor å vurdere egenskapene til estimatorene uten å

studere estimatorer for skjevhet, varians og bruttovarians, og vi oppnår det ved å bruke en bestand der alle enhetene er kjente.

Også i Sverige har det vist seg å være et stort behov for å lage regionale sysselsettingstall. I motsetning til den metoden vi har valgt å studere, har en der valgt å øke utvalget for å oppnå tilfredsstillende kvalitet på sysselsettingstall innen små geografiske områder. Dahmstrøm & Kristiansson (1976) har redegjort for denne utvidelsen av utvalget.

Framstillingen i denne artikkelen er delt i tre hovedavsnitt. Det første avsnittet gir en verbal oversikt over Byråets standard utvalgsplan og over estimatorene for totaler innen bestanden i hver region. I dette avsnittet gir vi også en oversikt over egenskapene til de foreslåtte estimatorene for totaler innen hvert fylke. Alle de numeriske beregninger av skjevhet, varians og bruttovarians er foretatt for en populasjon der alle enhetene er kjente. Tabeller for disse beregningene er tatt med i et vedlegg til artikkelen. Siden vi ønsker å gjøre resultatene av beregningene tilgjengelige for flest mulig, er konklusjonene av beregningene tatt med i slutten av det første hovedavsnittet. I det andre hovedavsnittet har vi gitt en formell beskrivelse av Byråets standard utvalgsplan og av de estimatorene som ble foreslått i det første hovedavsnittet.

2. GENERELT OM BYRÅETS UTVALGSPLAN OG OM ESTIMATORER FOR TOTALVERDIER I BESTANDEN I GEOGRAFISKE REGIONER

I dette hovedavsnittet vil vi gi noen generelle kommentarer til Byråets standard utvalgsplan og deretter kommentere de estimatorene som vi har foreslått for estimering av totalverdier i de geografiske regionene og fylkene. Denne framstillingen er hovedsakelig en verbal beskrivelse slik at den også er tilgjengelig for folk som ikke har en matematisk-statistisk bakgrunn.

2.1. Generelt om Byråets standard utvalgsplan

Byråets standard utvalgsplan bygger på to-trinns trekking av de enhetene som skal intervjues. På første trinn trekkes et sett faste utvalgsområder. Disse områdene kalles primære utvalgsområder. På annet trinn trekkes et utvalg av dem som skal intervjues, fra et register over alle som bor i de primære utvalgsområdene.

I utvalgsplanen har vi valgt å la kommunene være de primære utvalgsområdene. Kommuner med færre enn 3 000 innbyggere er slått sammen med andre kommuner slik at det er minst 3 000 innbyggere i hvert utvalgsområde. Utvalgsområdene er så stratifisert etter kommunetype og innbyggertall. I stratifiseringen er byer med over 30 000 innbyggere tatt ut som egne strata.

Landet er først inndelt i 5 landsdeler. Deretter er landsdelene inndelt i regioner, som hver består av ett eller to fylker. Disse regionene er:

Region	1:	Akershus,	
"	2:	Oslo,	
"	3:	Østfold og Vestfold,	
"	4:	Hedmark og Oppland,	
"	5:	Buskerud og Telemark,	
"	6:	Agder og Rogaland,	(2.1.1)
"	7:	Hordaland og Sogn og Fjordane,	
"	8:	Møre og Romsdal,	
"	9:	Sør- og Nord-Trøndelag,	
"	10:	Nordland	
og "	11:	Troms og Finnmark.	

Ved den videre stratifiseringen krysses ikke grensene for disse regionene. Det betyr altså at en lager nye strata innen hver region. En oversikt over stratifiseringen kan leseren finne i Thomsen & Rideng (1974).

I utvalgsplanen er nå byer med over 30 000 innbyggere tatt ut som egne strata, og utvalget i disse strataene er trukket rent lotterisk. Innen hvert av de øvrige strataene trekker en ut ett utvalgsområde med sannsynlighet proporsjonal med innbyggerantallet i utvalgsområdet, og en trekker så utvalget rent lotterisk innen hvert uttrukket utvalgsområde.

Av (2.1.1) ser vi at landet er inndelt i 11 regioner. Av disse regionene består 7 regioner av fylkespar og 4 av fylker. Siden stratifiseringen ikke krysser grensene for disse regionene, kan en altså lage forventningsrette estimatorer for totalverdien for bestanden innen hver slik region. Når en unntar Oslo, Akershus, Møre og Romsdal og Nordland, tar imidlertid utvalgsplanen ikke sikte på å lage forventningsrette estimatorer for totalverdier innen hvert fylke.

2.2. Generelt om estimatorene

Som nevnt innledningsvis er utvalgsområdene stratifisert, og vi trekker ett utvalgsområde fra hvert stratum. Vi ønsker å estimere en ukjent aggregatstørrelse i undersøkelsesbestanden. La da a_i være den ukjente totalverdien i undersøkelsesbestanden i stratum i. Vi lar videre a(p) være den ukjente totalen i bestanden i region p. Totalen a(p) framkommer altså som en uveid sum av stratumtotalene i region p. La da S(p) være mengden som består av strataene i region p. Vi ønsker altså å estimere totalen

$$a(p) = \sum_{i \varepsilon S(p)} a_i .$$

Laake (1974, side 3) har foreslått en forventningsrett estimator, $\hat{a}_i$, for a_i. Da er altså

$$\hat{a}(p) = \sum_{i \varepsilon S(p)} \hat{a}_i \qquad (2.2.1)$$

en forventningsrett estimator for a(p). Laake (1974, side 4) har videre funnet et uttrykk for var $\hat{a}_i$, og ut fra dette finner vi da enkelt et uttrykk for var $\hat{a}(p)$. Vi har dermed funnet en forventningsrett estimator for totalen innen hver region, og vi har funnet estimatorens varians. Både forventningen og variansen til estimatoren avhenger av ukjente parametre i undersøkelsesbestanden.

Vi har tidligere påpekt at det ikke er mulig å konstruere estimatorer som bygger på en uveid sum av estimatorer for stratumtotaler og som er forventningsrette for totalene i fylkene. En kan selvfølgelig konstruere forventningsrette estimatorer for totalen i fylker til tross for at en i utvalgsplanen ikke helt ut har tatt hensyn til at en skulle gjøre det. Disse estimatorene vil imidlertid ha stor varians, og vi har derfor valgt ikke å studere denne klassen av estimatorer.

Vi går derfor over til å studere den ene av de to klassene av estimatorer vi i denne artikkelen har valgt å studere. Det første settet av estimatorer bygger på informasjon om folketallet i fylkene. La da a(q,p) være totalen i undersøkelsesbestanden i fylke q i region p. Anta at det er henholdsvis N(q,p) og N(p) innbyggere i fylke q i region p og i hele regionen p. Vi vil nå studere egenskapene til estimatoren

$$\hat{a}^*(q,p) = \frac{N(q,p)}{N(p)} \hat{a}(p). \qquad (2.2.2)$$

Skjevheten til estimatoren (2.2.2) er

$$E\hat{a}^{*}(q,p) - a(q,p) = \frac{N(q,p)}{N(p)} a(p) - a(q,p). \tag{2.2.3}$$

Vi ser at estimatoren er forventningsrett dersom det relative forholdet mellom totalverdiene i undersøkelsesbestanden i fylke q i region p og hele regionen p er lik det relative forholdet mellom innbyggertallene.

Når vi konstruerer estimatoren (2.2.2), antar vi at N(q,p) og N(p) er kjente tall. Da er

$$\text{var } \hat{a}^{*}(q,p) = \{\frac{N(q,p)}{N(p)}\}^2 \text{ var } \hat{a}(p). \tag{2.2.4}$$

Siden estimatoren (2.2.2) er forventningsskjev, er det naturlig å studere estimatorens bruttovarians i tillegg til dens varians. Med estimatorens bruttovarians mener vi da

$$B \text{ var } \hat{a}^{*}(q,p) = E\{\hat{a}^{*}(q,p) - a(q,p)\}^2. \tag{2.2.5}$$

Siden nå

$$E\{\hat{a}^{*}(q,p) - a(q,p)\}^2 = \text{var } \hat{a}(q,p) + \\ + \{E\hat{a}^{*}(q,p) - a(q,p)\}^2, \tag{2.2.6}$$

er både skjevheten, variansen og bruttovariansen avhengig av de ukjente parametrene a(p) og a(q,p) i undersøkelsesbestanden og de kjente vektene N(q,p) og N(p).

Den andre estimeringsmetoden vi studerer, er en modifisert versjon av metoden med syntetisk estimering. Som nevnt innledningsvis, er syntetisk estimering en metode der en kombinerer a priori informasjon om de geografiske områdene med estimatorer fra en utvalgsundersøkelse. Bakgrunnen for at vi studerer syntetiske estimatorer er at vi ved konstruksjon av estimatorer for sysselsettingen i fylkene ønsker å bruke mer av den informasjonen vi har om fylkene enn bare data om innbyggerantallet. Syntetisk estimering har tidligere vært anvendt av National Center of Health Statistics (1968) og seinere av Gonzales & Waksberg (1973). National Center for Health Statistics (NCHS) utviklet metoden med syntetisk estimering for å estimere antall uføre i alle statene i USA i 1967. Utvalget ble da inndelt i 78 delklasser etter alder, kjønn, størrelse på husholdningen, inntekt osv., og ved en utvalgsundersøkelse oppnådde en estimater for proporsjoner av uføre i bestanden i hver av de 78 delklassene. For å oppnå estimater for totalbestanden i hver stat ble hvert av disse estimatene så veid sammen med relativt antall som tilhørte delklassen i staten. Disse vektene var kjent fra Census of Population i 1960. Egenskapene til de syntetiske estimatorene har seinere blitt studert av Gonzalez & Waksberg (1973).

Da vi dannet oss estimatoren (2.2.2), brukte vi bare informasjonen om forholdet mellom innbyggerantallet i fylkene og de tilsvarende regionene. I likhet med metoden som ble foreslått av NCHS i 1968, vil vi dele utvalget inn i delklasser etter demografiske variable som alder, kjønn o.l. Formålet med en slik delklasseinndeling er først og fremst å prøve å konstruere estimatorer som har mindre skjevhet enn estimatorene (2.2.2).

For å estimere totalverdien innen fylker veies nå estimatorene for regioner sammen med de kjente vektene for forholdet mellom antall i delklassen innen fylket og innen regionen. La altså $\hat{a}^{(\ell)}(p)$ være en estimator for totalen innen delklasse ℓ i region p. Anta videre at $r^{(\ell)}(q,p)$ er det relative forholdet mellom antall i undersøkelsesbestanden i delklassen i henholdsvis fylke q og hele regionen p. Som en estimator for a(q,p) foreslår vi da

$$\hat{a}^{**}(q,p) = \sum_{\ell} r^{(\ell)}(q,p)\ \hat{a}^{(\ell)}(p), \tag{2.2.7}$$

der $\sum_{q} r^{(\ell)}(q,p) = 1$ for $\ell = 1, 2, \ldots L$.

Dersom vi antar at det bare er en delklasse, og vi lar $r^{(1)}(q,p) = N(q,p)/N(p)$, vil (2.2.7) bli identisk med (2.2.2).

Skjevheten til (2.2.7) er

$$E\ \hat{a}^{**}(q,p) - a(q,p) = \sum_{\ell}\{r^{(\ell)}(q,p) - 1\}\ a^{(\ell)}(q,p), \tag{2.2.8}$$

som i likhet med (2.2.3) avhenger av ukjente parametre i bestanden. Dersom vi nå eksempelvis ønsker å estimere total sysselsetting i et fylke, deler vi utvalget i den aktuelle regionen inn i delklasser og "blåser" opp estimatene for hver delklasse. Oppblåsingsfaktorene er kjente, ikke-stokastiske vekter som er bestemt av totalt antall personer i delklassen i fylket i forhold til totalt antall i regionen. Ved å summere over alle delklassene finner vi en estimator $\hat{a}^{**}(q,p)$ for a-verdien i fylke q.

Siden vektene er ikke-stokastiske er

$$\begin{aligned} \operatorname{var} \hat{a}^{**}(q,p) = {} & \sum_{\ell}\{r^{(\ell)}(q,p)\}^2 \operatorname{var} \hat{a}^{(\ell)}(p) + \\ & + \sum_{\ell \neq \ell'} r^{(\ell)}(q,p) r^{(\ell')}(q,p) \operatorname{cov}\{\hat{a}^{(\ell)}(p),\ \hat{a}^{(\ell')}(p)\}. \end{aligned} \tag{2.2.9}$$

Fra definisjonen av bruttovariansen til en estimator i (2.2.5), finner vi også bruttovariansen til estimatoren (2.2.7). I likhet med skjevheten (2.2.8) avhenger både variansen og bruttovariansen til estimatoren av ukjente parametre i bestanden.

Estimatoren som ble foreslått av NCHS, er av en litt annen type enn estimatoren (2.2.7). Vi skal derfor studere estimatoren til NCHS litt nærmere. La som tidligere $\hat{a}^{(\ell)}(p)$ være en estimator for totalverdien i delklasse ℓ i region p, men la nå $s^{(\ell)}(q,p)$ betegne proporsjonen i bestanden i fylke q som tilhører delklasse ℓ. Når vi lar $N^{(\ell)}(q,p)$ betegne antallet i delklasse ℓ i fylke q, er altså

$$s^{(\ell)}(q,p) = \frac{N^{(\ell)}(q,p)}{N(q,p)} ,$$

og

$$\sum_{\ell} s^{(\ell)}(q,p) = 1.$$

NCHS har da foreslått å estimere totalverdien i fylke q med

$$\hat{a}^{***}(q,p) = \sum_{\ell} s^{(\ell)}(q,p)\ \hat{a}^{(\ell)}(p). \qquad (2.2.10)$$

La oss nå anta at vi deler inn utvalget i delklasser etter kjønn og alder. I Byråets standard utvalgsplan har en stratifisert de primære utvalgsområdene etter kommunetype. Ved stratifiseringen har en derfor prøvd å oppnå at sysselsettingsstrukturen for kommunene innen regionene er så homogen som mulig. Dersom en studerer estimatorer for total sysselsetting, er det derfor rimelig å anta at

$$a^{(\ell)}(q,p) \approx \frac{N^{(\ell)}(q,p)}{N^{(\ell)}(p)}\ a^{(\ell)}(p) \quad \text{for alle } \ell. \qquad (2.2.11)$$

Under antagelsen (2.2.11) vil estimatoren (2.2.7) være tilnærmet forventningsrett (se avsnitt 4.3.). Forventningen til estimatoren (2.2.10) vil under antagelsen (2.2.11) ha formen

$$E\hat{a}^{***}(q,p) = \sum_{\ell} \frac{N^{(\ell)}(q,p)}{N(q,p)}\ a^{(\ell)}(p).$$

Vi ser her at dersom antallet delklasser er stort, kan estimatoren (2.2.10) gi en betydelig underestimering av a(q,p). Vi vil derfor konsentrere oss om å studere egenskapene til estimatoren (2.2.7).

2.3. Sammenligning av to metoder for estimering i bestanden i fylker

Som nevnt i avsnitt 2.2. er formålet med å innføre syntetiske estimatorer at en med en inndeling av utvalget etter demografiske variable som alder, kjønn o.l., først og fremst prøver å redusere skjevheten til den enkle estimatoren (2.2.2). I dette avsnittet vil vi studere skjevheten til de to settene av estimatorer som ble foreslått i avsnitt 2.2. Skjevheten til estimatorene $\hat{a}^{*}(q,p)$ og $\hat{a}^{**}(q,p)$ er gitt ved (2.2.3) og (2.2.8). Da er altså differansen i skjevhet de to estimeringsmetodene

gitt ved

$$S\{\hat{a}^{*}(q,p),\ \hat{a}^{**}(q,p)\} = E\hat{a}^{**}(q,p) - E\hat{a}^{*}(q,p)$$

$$= \sum_{\ell} \{\frac{N^{(\ell)}(q,p)}{N^{(\ell)}(p)} - \frac{N(q,p)}{N(p)}\}\ a^{(\ell)}(p).$$

La oss nå anta at

$$\frac{1}{L} \sum_{\ell} \frac{N^{(\ell)}(q,p)}{N^{(\ell)}(p)} \approx \frac{N(q,p)}{N(p)}. \tag{2.4.1}$$

Da kan $S\{\hat{a}^{*}(q,p),\ \hat{a}^{**})q,p)\}$ uttrykkes som

$$S\{\hat{a}^{*}(q,p),\ \hat{a}^{**}(q,p)\} = \sum_{\ell} \{\frac{N^{(\ell)}(q,p)}{N^{(\ell)}(p)} - \frac{N(q,p)}{N(p)}\}\ \{a^{(\ell)}(p) - \frac{1}{L}\ a(p)\}. \tag{2.4.2}$$

Under antagelsen (2.4.1) ser vi at (2.4.2) kan uttrykkes som "kovariansen" mellom $N^{(\ell)}(q,p)/N^{(\ell)}(p)$ og $a^{(\ell)}(p)$. Vi ser derfor umiddelbart at differansen i skjevheten mellom de foreslåtte estimatorene er positiv når $N^{(\ell)}(q,p)/N^{(\ell)}(p)$ og $a^{(\ell)}(p)$ er "store" samtidig eller er "små" samtidig. Tilsvarende er differansen i skjevheten negativ når ett av uttrykkene $N^{(\ell)}(q,p)/N^{(\ell)}(p)$ og $a^{(\ell)}(p)$ er "lite" og ett er "stort".

Siden $\sum_{q} \frac{N(q,p)}{N(p)} = 1$, vil estimatorene for totaltallet i fylkene basert på "nedbryting" av estimatorene for regionene ved bruk av folketallet i fylkene overestimere sysselsettingen i det ene fylket og underestimere i det andre fylket. Anta da at skjevheten til estimatoren $\hat{a}^{*}(q,p)$ er negativ. Dersom vi nå velger en inndeling av bestanden i delklasser slik at $N^{(\ell)}(q,p)/N^{(\ell)}(p)$ og $a^{(\ell)}(p)$ er "store" eller "små" samtidig, vil vi i de fleste situasjoner få redusert skjevheten ved å bruke estimatoren $\hat{a}^{**}(q,p)$. Hvis vi antar at skjevheten til estimatoren $\hat{a}^{*}(q,p)$ er positiv, oppnår vi å få en reduksjon i skjevheten ved å velge en delklasseinndeling der ett av uttrykkene $N^{(\ell)}(q,p)/N^{(\ell)}(p)$ eller $a^{(\ell)}(p)$ er "stort" og ett "lite".

Eksempelvis kan vi anta at vi ønsker å estimere sysselsettingen i et fylke og at skjevheten til $\hat{a}^{*}(q,p)$ for et bestemt valg av q og p er negativ. Dette er ekvivalent med at sysselsettingen i fylket er "stor i forhold til folketallet der". For å oppnå en reduksjon i skjevheten vil vi derfor prøve å lage en delklasseinndeling der samvariasjonen mellom $N^{(\ell)}(q,p)/N^{(\ell)}(p)$ og $a^{(\ell)}(p)$ er stor. Vi prøver altså å finne en delklasseinndeling som best mulig "forklarer" den høye sysselsettingen i fylket. Intuitivt vil det derfor være best å "forklare" sysselsettingen ved hjelp av andre data om sysselsettingen. Dette oppnår vi selvfølgelig

ved å inndele utvalget i delklasser etter noen hovednæringer og bruke kjente vekter som angir sysselsettingsandeler.

2.4. Valg av delklasser til syntetisk estimering av sysselsettingen i fylker

Estimatoren (2.2.2) krever kjennskap bare til det relative forholdet mellom antall innbyggere i hvert fylke i en region og hele regionen. For å konstruere estimatoren (2.2.7) trenger vi å kjenne forholdstallet innen mindre delklasser. Vi har valgt å inndele bestanden i delklasser etter alder og kjønn. Inndelingen etter alder har blitt foretatt i følgende delklasser: 16-24 år, 25-34 år, 35-44 år, 45-54 år, 55-64 år, 65-74 år. Kombinert med kjønn gir dette 12 delklasser.

I denne artikkelen har vi valgt å studere egenskapene til estimatorer for total sysselsetting og sysselsetting i de ulike næringene, og det er derfor naturlig å inndele bestanden i delklasser etter hvilken næring personene i bestanden er sysselsatt i. Vektene som brukes i konstruksjonen av estimatoren (2.2.7) er da sysselsettingsandelen i en bestemt næring i ett fylke i forhold til den aktuelle regionen. Disse vektene antar vi er kjente fra annen statistikk. Den estimatoren som vi på denne måten kan konstruere, vil være en forventningsrett estimator for total sysselsetting i hvert fylke.

De tre estimatorene som altså er foreslått, bygger på en "nedbryting" basert på henholdsvis folketallsandel (2.2.2), andel i kjønns- og aldersgrupper (2.2.7) og sysselsettingsandel i de ulike næringene (2.2.7).

2.5. Om metoden som er valgt til å vurdere estimatorenes egenskaper

Formålet med dette prosjektet har vært å vurdere de statistiske egenskapene til estimatorene for totalverdiene i fylkene og i regionene. Spesielt har det vært av interesse å studere estimatorenes skjevhet, varians og bruttovarians. Som nevnt i avsnitt 2.2. er imidlertid både skjevheten, variansen og bruttovariansen avhengig av ukjente parametre i bestanden. De statistiske egenskapene kan derfor studeres bare dersom en lager estimater for en bestand der alle enhetene er kjente. Vi har laget en slik bestand, idet vi har trukket et 10 prosent utvalg blant alle personer bosatt i Norge 1/11 1970 (Folke- og boligtelling 1970). Dette

10 prosent utvalget består av ca. 400 000 enheter. Vi lar imidlertid vår bestand bestå bare av personer som er med i 10 prosent utvalget og er mellom 16 og 74 år. Dermed blir bestanden ytterligere redusert i størrelse. Bestanden inndeles nå i primære utvalgsområder og stratifiseres slik det er gjort i Byråets standard utvalgsplan. Stratifiseringen av kommunene bygger på kommuneinndelingen av 1/1 1974. Siden tallmaterialet er fra 1/11 1970, har vi måttet endre noe på den opprinnelige kommuneinndelingen i folketellingsmaterialet.

Fra denne "kjente" bestanden trekkes nå utvalg på 12 000 personer i to trinn. På denne måten oppnår vi å få en utvalgsplan som er identisk med den i AKU, og vi kan dermed studere de statistiske egenskapene til estimatorene i AKU for små geografiske regioner.

Vi ønsker å studere egenskapene til estimatorene for total sysselsetting og sysselsetting i de ulike næringene. Det er imidlertid ikke mulig å finne en variabel i Folke- og boligtelling 1970 som er identisk med variabelen sysselsetting i AKU. Vi har valgt å definere en person som sysselsatt dersom han har inntekt av eget arbeid som viktigste kilde til livsopphold. Denne variabelen er deretter kombinert med variabelen som angir hvilken næring personen er sysselsatt i.

De sysselsatte i bestanden er inndelt i følgende næringsgrupper:

Næringsgruppe 1: Jordbruk,
" 2: Skogbruk,
" 3: Fiske og fangst,
" 4: Bergverksdrift,
" 5: Industri,
" 6: Kraft- og vannforsyning,
" 7: Bygge- og anleggsvirksomhet,
" 8: Varehandel, hotell- og restaurantvirksomhet,
" 9: Transport, lagring, post og telekommunikasjon,
" 10: Bank- og finansieringsvirksomhet, forsikringsvirksomhet, eiendomsdrift og forretningsmessig tjenesteyting,
og " 11: Offentlig, sosial og privat tjenesteyting.

Ved inndelingen i næringsgrupper er det tatt hensyn til at en i Folke- og boligtelling 1970 brukte den gamle standarden for næringsgruppering, mens AKU bruker den nye grupperingen.

3. KONKLUSJONER

Vi vil nå se nærmere på de numeriske beregningene av forventing, varians og bruttovarians til de foreslåtte estimatorene. Først vil vi imidlertid gi noen generelle kommentarer om estimatorenes egenskaper og finne kriterier for enkelte estimater vi kan publisere.

3.1. Generelt om de numeriske beregningene

Av praktiske og økonomiske hensyn har vi i denne artikkelen foretatt de numeriske beregningene for en undersøkelsesbestand som består av et 10 prosent utvalg av alle personer bosatt i Norge pr. 1. november 1970. La oss imidlertid anta at totalverdiene for alle de variable vi studerer i hele bestanden er 10 ganger så store som den i vår undersøkelsesbestand, og at variasjonen pr. enhet innen kommunene er den samme i de to bestandene. De sammenligninger som vi har foretatt mellom undersøkelsesbestanden og resultatene fra Folke- og boligtelling 1970, viser godt samsvar mellom bestandene. Det vil derfor ikke bety noe tap av generalitet ved å studere estimater fra vår undersøkelsesbestand istedenfor for hele Folke- og boligtelling 1970. Det kan nå lett vises at det relative standardavviket, dvs. forholdet mellom standardavviket og estimanden, er tilnærmet det samme i undersøkelsesbestanden og den egentlige bestanden. Dersom vi ønsker å vurdere kvaliteten av estimatorene ved hjelp av det relative standardavviket, kan vi altså begrense oss til å vurdere estimatene fra undersøkelsesbestanden.

Vi ønsker nå å estimere en totalverdi $a(p)$, og vi angir estimatoren $\hat{a}(p)$. La oss anta at

$$E\hat{a}(p) = a(p)$$

og

$$\text{var}\ \hat{a}(p) = \sigma_{\hat{a}(p)}.$$

Dersom vi nå antar av $\hat{a}(p)$ er tilnærmet normalfordelt, vil $\{\hat{a}(p) - a(p)\}/\sigma_{\hat{a}(p)}$ være tilnærmet normalfordelt med forventning lik 0 og varians lik 1. Videre er

$$\left[\hat{a}(p) - z_{1-\varepsilon/2}\,\sigma_{\hat{a}(p)},\ \hat{a}(p) + z_{1-\varepsilon/2}\,\sigma_{\hat{a}(p)}\right], \tag{3.1.1}$$

der $z_{1-\varepsilon/2}$ er $(1-\varepsilon/2)$-fraktilen i den standardiserte normalfordelingen,

et tilnærmet $(1-\varepsilon)$-konfidensintervall for $a(p)$. Vi ser at konfidensintervallet er symmetrisk omkring $\hat{a}(p)$, og vi lar da $z_{1-\varepsilon/2}\,\sigma_{\hat{a}(p)}$ betegne konfidensavviket.

En ønsker ikke å publisere estimater som er sterkt beheftet med samplingfeil. Vi velger derfor å publisere forventningsrette estimatorer bare for de estimatene der

$$\{z_{1-\varepsilon/2}\,\sigma_{\hat{a}(p)}\}/a(p) \leq 0.4, \tag{3.1.2}$$

dvs. at konfidensavviket skal være mindre enn 40 prosent av estimanden. Dette kriteriet har tidligere vært brukt av Brenna & Hoem (1973) i en litt annen sammenheng. Ved å sette $\varepsilon = 0.05$ gir kriteriet (3.1.2) at vi for forventningsrette estimatorer bare publiserer estimater der

$$\sigma_{\hat{a}(p)}/a(p) \leq 0.2. \tag{3.1.3}$$

Vi har i avsnittet 2.2. vist at estimatorene for totalverdiene innen fylkene ikke er forventningsrette. Disse estimatorene vil derfor kreve en spesiell behandling. Ifølge (2.2.3) er

$$E\hat{a}^{*}(q,p) = a(q,p) + B,$$

der $B = \{E\hat{a}^{*}(q,p) - a(q,p)\}$. Under forutsetning av at $\hat{a}^{*}(q,p)$ er normalfordelt med forventning $a(q,p) + B$ og varians $\sigma_{\hat{a}^{*}(q,p)}$ er

$$\left[\hat{a}^{*}(q,p) - z_{1-\varepsilon/2}\,\sigma_{\hat{a}^{*}(q,p)},\ \hat{a}^{*}(q,p) + z_{1-\varepsilon/2}\,\sigma_{\hat{a}^{*}(q,p)}\right] \tag{3.1.4}$$

et tilnærmet $(1-\varepsilon)$-konfidensintervall for $a(q,p) + B$. Vanligvis vil B være ukjent, og vi vil derfor være interessert i sannsynligheten for at konfidensintervallet (3.1.4) dekker den "sanne" parameter $a(q,p)$ i bestanden. Denne sannsynligheten er

$$\begin{aligned}
P_1 &= \Pr\{\hat{a}^{*}(q,p) - z_{1-\varepsilon/2}\sigma_{\hat{a}^{*}(q,p)} \leq a(q,p) \leq \\
&\qquad \leq \hat{a}^{*}(q,p) + z_{1-\varepsilon/2}\sigma_{\hat{a}^{*}(q,p)}\} \\
&= \Pr\{-z_{1-\varepsilon/2} - \frac{B}{\sigma_{\hat{a}^{*}(q,p)}} \leq \frac{\hat{a}^{*}(q,p) - E\hat{a}^{*}(q,p)}{\sigma_{\hat{a}^{*}(q,p)}} \leq \\
&\qquad \leq z_{1-\varepsilon/2} - \frac{B}{\sigma_{\hat{a}^{*}(q,p)}}\} \\
&= \Phi\{z_{1-\varepsilon/2} - \frac{B}{\sigma_{\hat{a}^{*}(q,p)}}\} - \Phi\{-z_{1-\varepsilon/2} - \frac{B}{\sigma_{\hat{a}^{*}(q,p)}}\},
\end{aligned}$$

der Φ betegner den kumulative standardiserte normalfordelingen. Vi beregner nå sannsynligheten P_1 for ulike verdier av $B / \sigma_{\hat{a}^*(q,p)}$. Resultatene er gitt i oppstillingen nedenfor, som er delvis omarbeidet fra en tabell i Cochran (1963, side 14).

$B/\sigma_{\hat{a}^*(q,p)}$	0	0,1	0,2	0,4	0,6	0,8
P_1	0,9500	0,9489	0,9454	0,9315	0,9079	0,8741

$B/\sigma_{\hat{a}^*(q,p)}$	1,0	1,2	1,4	1,6	1,8	2,0
P_1	0,8300	0,7756	0,7119	0,6404	0,5634	0,4841

Vi ser altså at selv for meget store skjevheter (helt opp til $B/\sigma_{\hat{a}^*(q,p)} \approx 2$) vil konfidensintervallet (3.1.4) dekke den "sanne" parameter med en sannsynlighet på minst 0,5. Vi velger derfor å si at et estimat $\hat{a}^*(q,p)$ kan publiseres dersom det maksimale konfidensavviket fra $a(q,p)$ i forhold til estimanden selv er mindre enn 40 prosent. Altså publiserer vi bare estimater der vi har oppfylt at

$$\frac{z_{1-\varepsilon/2}\,\sigma_{\hat{a}^*(a,p)} + |B|}{a(q,p)} < 0{,}40,$$

eller omregnet at

$$\frac{\sigma_{a^*(q,p)}}{a(q,p)} + \frac{|B|}{2a(q,p)} < 0{,}20. \qquad (3.1.5)$$

Tilsvarende finner vi et kriterium for å publisere estimater for $\hat{a}^{**}(q,p)$ ved å erstatte $\hat{a}^*(q,p)$ i (3.1.5) med $\hat{a}^{**}(q,p)$.

Vi har brukt kriteriene (3.1.3) og (3.1.4) for alle estimatene i tabellene i vedlegget. Dersom vi ifølge kriteriene har funnet at et estimat <u>ikke</u> kan publiseres, har vi satt parentes rundt tallet i tabellene.

3.2. Konklusjoner til de numeriske beregningene

I tabellvedlegget har vi publisert tabeller med estimater for total sysselsetting og sysselsetting etter næring i de geografiske regionene og alle fylkene. Alle tabellene viser forventningen til de foreslåtte estimatorene for sysselsetting, korrekt sysselsetting og standardavviket, skjevheten og bruttovariansen til estimatorene.

Tabell 1 viser estimator for sysselsettingen i de ulike regionene. Siden estimatoren (2.2.1) er en forventningsrett estimator for total sysselsetting i en region, vil forventet sysselsetting og korrekt sysselsetting være like og dermed skjevheten lik null. Siden skjevheten er lik null, vil dermed standardavviket og bruttostandardavviket være like.

Tabellene viser at en med den publiseringsregelen vi har valgt i avsnittet 3.1., kan publisere estimatene i de store næringene i hver region. For å få tilfredsstilt kravet (3.1.3) må vi ha at sysselsettingen i en av hovednæringene i en region bør være over ca. 7 000 for at en kan publisere estimatet. Til sammenligning fant Brenna & Hoem (1973) at en ikke kunne publisere estimater som var mindre enn ca. 10 000.

Tabell 2 viser estimater for sysselsettingen i hvert fylke beregnet etter (2.2,2). De fylkene som er identiske med regionene, som er definert i avsnittet 2.1., er ikke tatt med i disse tabellene. Siden vektene $\{N(q,p)/N(p)\}$ i disse tilfellene er lik 1, blir da estimatene identiske med dem i tabell 1.

Vi ønsker å publisere estimater som tilfredsstiller (3.1.5). Med den publiseringsregelen vi har valgt, kan vi altså publisere estimater for sysselsettingen innen de store næringene også i fylkene. Vi kan lage en generell publiseringsregel ved å si at vi publiserer estimater i fylkene der de tilsvarende estimatene i regionene er større enn 10 000. Unntatt fra denne publiseringsregelen er imidlertid estimatene for sysselsatte i jordbruk i Hedmark, Aust- og Vest-Agder, Rogaland, Hordaland, Sogn og Fjordane og Sør-Trøndelag. I Troms må vi unnta estimatet for sysselsatte i industri, i Østfold transport, lagring, post- og telekommunikasjon. I Finnmark kan vi verken publisere estimater for industri eller offentlig, sosial og privat tjenesteyting og i Nord-Trøndelag verken jordbruk eller varehandel, hotell- og restaurantvirksomhet. I unntakstilfellene må vi i enkelte tilfeller kreve over 25 000 sysselsatte i hovednæringen i regionen for å kunne publisere de tilsvarende estimatene i fylkene.

Tabell 3 viser estimert total sysselsetting i fylkene. Estimatorene bygger på en "nedbryting" av sysselsettingen etter næring. Standardavviket på disse estimatene er av samme størrelsesorden som standardavviket for estimatorene som er nedbrutt etter befolkningsandel. Estimatorene i tabell 3 er imidlertid forventningsrette, og bruttostandardavviket til estimatene blir derfor mindre enn for dem i tabell 2. En skal imidlertid være klar over at for å konstruere slike estimatorer trenger

en a priori informasjon om sysselsettingsandelen i fylkene i hver næring.

Estimatene i tabell 4 bygger på en "nedbryting" etter alder og kjønn. Bakgrunnen for en slik "nedbryting" er forventningen om at hvis en tar med informasjonen om forskjeller i alders- og kjønnsfordelingen i regionene i konstruksjonen av estimatorene, kan en redusere bruttostandardavviket til estimatorene. Tabell 2 viser imidlertid at en ikke får økt presisjonen så mye at en kan publisere flere estimater ved bruk av denne metoden enn ved bruk av metoden med "nedbryting" ved folketallsandel. Vi ser at både standardavviket til estimatorene og skjevheten er av samme størrelsesorden. Metoden med syntetisk estimering og med den delklasseinndeling vi har valgt, gir derfor ikke vesentlige bidrag til reduksjon av bruttostandardavvik. En mulig forklaring på dette kan være at delklasseinndelingen etter at dataene er samlet inn, og "nedbrytingen" med kjente fylkesvekter har liten effekt ut over den effekten forhåndsstratifiseringen har.

På grunnlag av resultatene kan en altså konkludere med at en for regioner kan publisere sysselsettingstall som er større enn 7 000. En kan også lage kvartalsvise tall for sysselsettingen i fylker på grunnlag av "nedbryting" ved folketall, men disse bør ikke offentliggjøres.

4. FORMELL BESKRIVELSE AV BYRÅETS STANDARD UTVALGSPLAN OG DE FORESLÅTTE ESTIMATORENE

I dette hovedavsnittet vil vi først gi en formell beskrivelse av Byråets standard utvalgsplan og studere estimatorer for totalverdier for bestanden innen strata og totalverdier for hele undersøkelsesbestanden. Deretter vil vi studere de statistiske egenskapene til de foreslåtte estimatorene for totalverdier innen fylker.

4.1. Formell beskrivelse av Byråets standard utvalgsplan

Som nevnt innledningsvis er bestanden inndelt i primære utvalgsområder. Utvalgsområdene er stratifisert. Vi antar at det er M_i utvalgsområder i stratum nr. i. Det j-te av disse har $N_i(j)$ trekkeenheter, og den k-te trekkeenheten har verdien $a_i(j,k)$ på det vi måler. Vi lar

$$N_i = \sum_j N_i(j),$$

$$N = \sum_i N_i,$$

$$a_i(j) = \sum_k a_i(j,k),$$

$$\bar{a}_i(j) = a_i(j)/N_i(j),$$

$$a_i = \sum_j a_i(j),$$

$$\bar{a}_i = a_i/M_i,$$

$$\text{og } a = \sum_i a_i = \sum_i \sum_j \sum_k a_i(j,k).$$

I Byråets utvalgsundersøkelser ønsker en vanligvis å estimere totalverdien a. I dette prosjektet er vi også interessert i å estimere totalverdien innen geografiske regioner som består av fylker eller par av fylker.

Anta at vi i stratum i trekker ut m_i av de M_i utvalgsområdene. La $\pi_i(j)$ være sannsynligheten for at utvalgsområde j i stratum i blir trukket ut, og la $\pi_i(j,k)$ være sannsynligheten for at både utvalgsområde j og k i stratum i blir trukket ut.

La $J_{i1}, J_{i2}, \ldots, J_{im_i}$ være numrene på de utvalgsområdene som blir trukket ut i stratum i, og la $\underset{\sim}{J} = (J_{11}, \ldots, J_{1m_1}, J_{21}, \ldots)$ være numrene på alle de uttrukne utvalgsområdene. Fra hvert uttrukket utvalgsområde trekkes et gitt antall trekkeenheter rent lotterisk. La $n(\underset{\sim}{J})$ og $n_{ij}(\underset{\sim}{J})$ være totalt antall i henholdsvis hele utvalget og i utvalget i j-te utvalgsområde i stratum i.

Numrene på de enhetene som blir trukket ut i utvalgsområde j i stratum i, betegner vi med $K_{ij1}, K_{ij2}, \ldots$, og vi lar

$$X_{ijs} = a_i(j, K_{ijs}),$$

og

$$\bar{X}_{ij} = \sum_s X_{ijs}/n_{ij}(\underset{\sim}{J}).$$

Vi innfører indeksvariabelen

$$I_{ij} = \begin{cases} 1 & \text{dersom utvalgsområde j i stratum i er i utvalget,} \\ 0 & \text{ellers,} \end{cases}$$

og lar $\bar{X}_{ij}$ være gjennomsnittet for de $n_{ij}(\underset{\sim}{J})$ uttrukne trekkeenhetene i utvalgsområde j i stratum i. Dersom $I_{ij} = 0$, er $n_{ij}(\underset{\sim}{J}) = 0$, og vi definerer $\bar{X}_{ij} = 0$.

En estimator for totalen i stratum i er da

$$\hat{a}_i = \sum_j \{I_{ij} N_i(j) \bar{X}_{ij} / \pi_i(j)\}, \tag{4.1.1}$$

og

$$\hat{a} = \sum_i \hat{a}_i$$

er en estimator for estimanden a. Laake (1974, s. 3) har da vist at $\hat{a}_i$ er forventningsrett estimator for a_i, og $\hat{a}$ er forventningsrett for a.

La nå

$$\sigma_i^2(j) = \frac{1}{N_i(j) - 1} \sum_k \{a_i(j,k) - \bar{a}_i(j)\}^2,$$

og

$$\tau_{ij}^2(\underset{\sim}{J}) = \frac{\sigma_i^2(j)}{n_{ij}(\underset{\sim}{J})} \cdot \frac{N_i(j) - n_{ij}(\underset{\sim}{J})}{N_i(j)}. \tag{4.1.2}$$

Dersom $I_{ij} = 0$, er $n_{ij}(\underset{\sim}{J}) = 0$, og $\tau_{ij}^2(\underset{\sim}{J})$ er udefinert. La nå $\hat{a}_i$ være definert ved (4.1.1). Da har Laake (1974, s. 4) vist at

$$\text{var}\, \hat{a}_i = \sum_j \sum_k \{\frac{\pi_i(j,k) - \pi_i(j)\, \pi_i(k)}{\pi_i(j)\, \pi_i(k)} a_i(j) a_i(k)\} +$$

$$+ \sum_j \{\eta_i(j)/\pi_i(j)\}, \tag{4.1.3}$$

hvor

$$\eta_i(j) = E\{N_i^2(j)\, \tau_{ij}^2(\underset{\sim}{J}) \mid I_{ij} = 1\}. \tag{4.1.4}$$

Vi sier at utvalget er selvveiende dersom estimatoren (4.1.1) kan skrives på formen

$$\hat{a} = \frac{1}{b(\underset{\sim}{J})} \sum_i \sum_j I_{ij} \sum_s X_{ijs}. \tag{4.1.5}$$

Vi lar

$$n = \sum_i \sum_j n_{ij}(\underset{\sim}{J}) \tag{4.1.6}$$

være gitt. I Byråets utvalgsplan er utvalget selvveiende. Dette oppnår vi dersom

$$b(\underset{\sim}{J}) = \pi_i(j)\, n_{ij}(\underset{\sim}{J}) / N_i(j),$$

$$b_i(\underset{\sim}{J}) = b(\underset{\sim}{J}) / \pi_i(j),$$

og

$$n_{ij}(\underset{\sim}{J}) = b_i(\underset{\sim}{J})\, N_i(j) \text{ for alle } j \text{ slik at } I_{ij} = 1.$$

I Byråets utvalgsplan trekkes det ett utvalgsområde i hvert stratum. Vi lar sannsynligheten for å trekke et utvalgsområde være proporsjonal med innbyggerantallet i området. Da er

$$\pi_i(j) = N_i(j)/N_i \text{ og } \pi_i(j,k) = 0 \text{ for } j \neq k, \tag{4.1.7}$$

og det kan lett vises at (4.1.3) da reduseres til

$$\text{var } \hat{a}_i = \sum_j \pi_i(j) \left\{\frac{a_i(j)}{\pi_i(j)} - a_i\right\}^2 + \sum_j \{\eta_i(j)/\pi_i(j)\}, \tag{4.1.8}$$

der $\eta_i(j)$ er definert ved (4.1.4).

I Byråets standard utvalgsplan er byer med flere enn 30 000 innbyggere tatt ut som egne strata. I disse strataene trekker vi ut enhetene rent lotterisk. Disse strataene krever en separat behandling, og vi vil nå studere nærmere en estimator for totalverdien innen disse strataene.

Anta at det er N_{oh} trekkeenheter i stratum h. Den k-te enheten i stratumet har verdien $a_{oh}(k)$ på det vi måler. Tilsvarende avsnittet foran innfører vi

$$a_{oh} = \sum_k a_{oh}(k),$$

og

$$\bar{a}_{oh} = a_{oh}/N_{oh}.$$

Av de N_{oh} enhetene trekkes et utvalg på

$$n_{oh}(\underset{\sim}{J}) = b(\underset{\sim}{J})N_{oh}$$

trekkeenheter. De $n_{oh}(\underset{\sim}{J})$ enhetene har numrene K_{oh1}, K_{oh2}, ...
La da

$$X_{ohs} = a_{oh}(K_{ohs}),$$

og

$$\bar{X}_{oh} = \sum_s X_{ohs}/n_{oh}(\underset{\sim}{J}).$$

Da blir altså

$$\hat{a}_{oh} = N_{oh}\bar{X}_{oh} \tag{4.1.9}$$

en forventningsrett estimator for totalen i stratum h. Vi definerer videre

$$\sigma^2_{oh} = \frac{1}{N_{oh} - 1} \sum_k \{a_{oh}(k) - \bar{a}_{oh}\}^2,$$

og da er

$$\text{var } \hat{a}_{oh} = \frac{\sigma^2_{oh}}{n} N_{oh}(N-n). \tag{4.1.10}$$

Vi har nå funnet forventningsrette estimatorer for totalverdien i undersøkelsesbestanden både i strataene der vi trekker utvalget i to trinn og i de strataene der vi trekker utvalget rent lotterisk. Som nevnt innledningsvis består hver region av en samling strata. La I(p) være samlingen av de strata som består av flere enn ett utvalgsområde og H(p) samlingen av strata med ett utvalgsområde. Siden $\hat{a}_i$ og $\hat{a}_{oh}$ er forventningsrette estimatorer for totaler innen de respektive strataene, er

$$\hat{a}(p) = \sum_{i\varepsilon\, I(p)} \hat{a}_i + \sum_{h\varepsilon\, H(p)} \hat{a}_{oh} \tag{4.1.11}$$

en forventningsrett estimator for totalen i region p. Siden $\text{cov}(\hat{a}_i,\hat{a}_j)=0$ for $i \neq j$ og $\text{cov}(\hat{a}_{oh},\hat{a}_{og}) = 0$ for $h \neq g$, er

$$\text{var } \hat{a}(p) = \sum_{i\varepsilon\, I(p)} \text{var } \hat{a}_i + \sum_{h\varepsilon\, H(p)} \text{var } \hat{a}_{oh}, \tag{4.1.12}$$

der var $\hat{a}_i$ og var $\hat{a}_{oh}$ er gitt ved henholdsvis (4.1.8) og (4.1.10).

4.2. En estimator for totalverdien i fylker

I dette avsnittet vil vi studere egenskapene til en estimator for totalverdien innen fylker. Estimatoren er forventningsrett dersom det relative forholdet mellom totalverdien i fylke q og totalverdien i regionen p er lik det relative forholdet mellom innbyggertallene.

La nå N(q,p) og N(p) være antall innbyggere i henholdsvis fylke q i region p og hele regionen. En estimator for totalverdien a(q,p) i fylke q i region p er da

$$\hat{a}^*(q,p) = \frac{N(q,p)}{N(p)} \hat{a}(p). \tag{4.2.1}$$

Siden $\hat{a}(p)$ er forventningsrett for a(p), er skjevheten til estimatoren (4.2.1)

$$E\hat{a}^*(q,p) - a(q,p) = \frac{N(q,p)}{N(p)} a(p) - a(q,p). \tag{4.2.2}$$

Dersom nå

$$\frac{a(q,p)}{a(p)} = \frac{N(q,p)}{N(p)},$$

er altså estimatoren forventningsrett. Videre er bruttovariansen til estimatoren av formen

$$B \text{ var } \hat{a}^*(q,p) = \left\{\frac{N(q,p)}{N(p)}\right\}^2 \text{var } \hat{a}(p) + \left\{\frac{N(q,p)}{N(p)} a(p) - a(q,p)\right\}^2 \tag{4.2.3}$$

som altså består av to positive komponenter, en varianskomponent og en skjevhetskomponent.

4.3. Syntetiske estimatorer

Estimatoren (4.2.1) bruker informasjon om folketallet i fylket og den tilsvarende regionen til å lage estimatorer for totalverdien i bestanden i fylker. I dette avsnittet vil vi studere estimatorer som benytter mer informasjon om fylkene enn bare informasjon om folketallet. For å gjøre dette, inndeles enhetene i utvalget i delklasser. Estimatorene for totalverdiene i delklassene veies da sammen ved hjelp av vekter som er kjente fra fylkene og den tilsvarende regionen. Formålet med denne metoden er å etterkomme ønsket om en reduksjon av bruttovariansen til estimatoren for totalverdien i fylker.

Vi antar nå at enhetene i utvalget inndeles i L delklasser etter at dataene er samlet inn. Vi innfører da indeksvariabelen $I_i^{(\ell)}(j,k)$ som antar verdien 1 dersom enhet nr. (i,j,k) tilhører delklasse nr. ℓ og 0 ellers. Vi definerer da

$$a_i^{(\ell)}(j,k) = I_i^{(\ell)}(j,k)\ a_i(j,k)$$

som gir at

$$a_i(j,k) = \sum_{\ell} a_i^{(\ell)}(j,k).$$

For hver ℓ for vi altså en teori som er analog med den som er utviklet i avsnitt 4.1.

Vi lar altså

$$a_i^{(\ell)}(j) = \sum_k a_i^{(\ell)}(j,k),$$

$$\bar{a}_i^{(\ell)}(j) = a_i^{(\ell)}(j)/N_i(j),$$

$$a_i^{(\ell)} = \sum_j a_i^{(\ell)}(j),$$

$$a^{(\ell)} = \sum_i a_i^{(\ell)},$$

$$\sigma_i^2(j,\ell) = \frac{1}{N_i(j) - 1} \sum_k \{a_i^{(\ell)}(j,k) - a_i^{(\ell)}(j)\}^2,$$

$$\tau_{ij}^2(\underset{\sim}{J},\ell) = \frac{\sigma_i^2(j,\ell)}{n_{ij}(\underset{\sim}{J})}\ \frac{N_i(j) - n_{ij}(\underset{\sim}{J})}{N_i(j)},$$

og

$$\eta_i^{(\ell)}(j) = E\,\{N_{ij}^2(\underset{\sim}{J})\ \tau_{ij}^2(\underset{\sim}{J},\ell)\ |\ I_{ij} = 1\}.$$

Analogt med tidligere notasjon innfører vi videre

$$X_{ijs}^{(\ell)} = I_i^{(\ell)}(j,K_{ijs})\ a_i(j,K_{ijs}),$$

og

$$\bar{X}_{ij}^{(\ell)} = \sum_s X_{ijs}^{(\ell)}/n_{ij}(\underset{\sim}{J}).$$

Vi ønsker å estimere totalverdien $a_i^{(\ell)}$. Analogt med teorien i avsnitt 4.1. er nå

$$\hat{a}_i^{(\ell)} = \sum_j \{I_{ij}N_i(j)\bar{X}_{ij}^{(\ell)}/\pi_i(j)\} \tag{4.3.1}$$

en forventningsrett estimator for $a_i^{(\ell)}$, og

$$\text{var}\ \hat{a}_i^{(\ell)} = \sum_j \pi_i(j)\ \{\frac{a_i^{(\ell)}(j)}{\pi_i(j)} - a_i^{(\ell)}\}^2 + \sum_j\{\eta_i^{(\ell)}(j)/\pi_i(j)\}. \tag{4.3.2}$$

Siden $\bar{X}_{ij}^{(\ell)}$ og $\bar{X}_{ij}^{(\ell')}$ ikke er stokatisk uavhengige for $\ell \neq \ell'$, må vi finne et uttrykk for cov $\{\hat{a}_i^{(\ell)}, \hat{a}_i^{(\ell')}\}$.

Vi innfører da

$$\sigma_i(j,\ell,\ell') = \frac{1}{N_i(j) - 1}\ \sum_k\{a_i^{(\ell)}(j,k) - \bar{a}_i^{(\ell)}(j)\}\{a_i^{(\ell')}(j,k) - \bar{a}_i^{(\ell')}(j)\},$$

$$\tau_{ij}(\underset{\sim}{J},\ell,\ell') = \frac{\sigma_i(j,\ell,\ell')}{n_{ij}(\underset{\sim}{J})}\ \frac{N_i(j) - n_{ij}(\underset{\sim}{J})}{N_i(j)},$$

og

$$\eta_i(j,\ell,\ell') = E\,\{N_i^2(j)\ \tau_{ij}(\underset{\sim}{J},\ell,\ell')\ |\ I_{ij} = 1\}.$$

Ved å bruke

$$\text{cov}\ \{\hat{a}_i(\ell),\ \hat{a}_i(\ell')\} = \frac{1}{2}\left[\text{var}\{\hat{a}_i(\ell) + \hat{a}_i(\ell')\} - \text{var}\ \hat{a}_i(\ell) - \text{var}\ \hat{a}_i(\ell')\right],$$

finner vi at

$$\text{cov}\{\hat{a}_i(\ell),\ \hat{a}_i(\ell')\} = \sum_j \pi_i(j)\ \{\frac{a_i^{(\ell)}(j)}{\pi_i(j)} - a_i^{(\ell)}\}\{\frac{a_i^{(\ell')}(j)}{\pi_i(j)} - a_i^{(\ell')}\} + \sum_j\{\eta_i(j,\ell,\ell')/\pi_i(j)\}. \tag{4.3.3}$$

I avsnittet 4.1. ble de strataene som består av byer med flere enn 30 000 innbyggere, behandlet separat. Helt analogt med teorien for de strataene som består av flere enn ett utvalgsområde, innfører vi nå størrelsene

$$X_{ohs}^{(\ell)} = I_o^{(\ell)}(K_{ohs})\ a_{oh}(K_{ohs}),$$

$$\bar{X}_{oh}^{(\ell)} = \sum_s X_{ohs}^{(\ell)} / n_{oh}(\underset{\sim}{J}),$$

$$\sigma_{oh}^2(\ell) = \frac{1}{N_{oh} - 1} \sum_k \{a_{oh}^{(\ell)}(k) - \bar{a}_{oh}^{(\ell)}\}^2,$$

og

$$\sigma_{oh}(\ell,\ell') = \frac{1}{N_{oh} - 1} \sum_k \{a_{oh}^{(\ell)}(k) - \bar{a}_{oh}^{(\ell)}\}\ \{a_{oh}^{(\ell')}(k) - \bar{a}_{oh}^{(\ell')}\}.$$

Da er

$$\bar{a}_{oh}^{(\ell)} = N_{oh}\ \bar{X}_{oh}^{(\ell)} \tag{4.3.4}$$

en forventningsrett estimator for $a_{oh}^{(\ell)}$, og

$$\text{var}\ \hat{a}_{oh}^{(\ell)} = \frac{\sigma_{oh}^2(\ell)}{n} N_{oh}(N - n), \tag{4.3.5}$$

og

$$\text{cov}\ \{\hat{a}_{oh}^{(\ell)}, \hat{a}_{oh}^{(\ell')}\} = \frac{\sigma_{oh}(\ell,\ell')}{n} N_{oh}(N - n). \tag{4.3.6}$$

Vi finner altså en forventningsrett estimator for $a_i^{(\ell)}(p)$ ved å sette inn ℓ der det trengs i (4.1.11). Tilsvarende finner vi variansen til estimatoren ved å sette inn ℓ i uttrykket (4.1.12).

Anta nå at utvalget er delt inn i L klasser. La $N^{(\ell)}(q,p)$ og $N^{(\ell)}(p)$ være antallet i delklasse ℓ i henholdsvis fylke q i region p og hele regionen p. Vi definerer da $r^{(\ell)}(q,p) = N^{(\ell)}(q,p)/N^{(\ell)}(p)$. Vi ønsker å estimere totalverdien a(q,p) og foreslår da estimatoren

$$\hat{a}^{**}(q,p) = \sum_\ell r^{(\ell)}(q,p)\ \hat{a}^{(\ell)}(p). \tag{4.3.7}$$

Da er skjevheten til estimatoren gitt ved at

$$E\ \hat{a}^{**}(q,p) - a(q,p) = \sum_\ell r^{(\ell)}(q,p)\ a^{(\ell)}(p) - a(q,p). \tag{4.3.8}$$

Videre er

$$\text{var}\ \hat{a}^{**}(q,p) = \sum_{\ell} \{r^{(\ell)}(q,p)\}^2 \sum_i \text{var}\ \hat{a}_i^{(\ell)} + \tag{4.3.9}$$

$$+ \sum_{\ell \neq \ell'} r^{(\ell)}(q,p)\ r^{(\ell')}(q,p) \sum_i \text{cov}\{\hat{a}_i^{(\ell)}, \hat{a}_i^{(\ell')}\}.$$

der var $\hat{a}_i^{(\ell)}$ er gitt ved (4.3.2) og (4.3.5) og cov $\{\hat{a}_i^{(\ell)}, \hat{a}_i^{(\ell')}\}$ er gitt ved (4.3.3) og (4.3.6). I uttrykket (4.3.9) er nå $r^{(\ell)}(q,p)$ kjente tall for alle ℓ, mens varians- og kovariansleddene er tall som avhenger av de ukjente parametrene i undersøkelsesbestanden.

Anta at vi ønsker å estimere total sysselsetting i et fylke. Enhetene i utvalget inndeles nå i L delklasser etter sysselsettingen til enhetene i utvalget, og vi antar at forholdet mellom antall sysselsatte i delklasse ℓ i fylke q og region p, $r^{(\ell)}(q,p)$ er kjent for alle ℓ, q og p. Da er estimatoren (4.3.6) forventningsrett. Det innser vi på følgende måte:

Siden $r^{(\ell)}(q,p) = a^{(\ell)}(q,p)/a^{(\ell)}(p)$, vil nå

$$E\ \hat{a}^{**}(q,p) = \sum_{\ell} \frac{a^{(\ell)}(q,p)}{a^{(\ell)}(p)}\ a^{(\ell)}(p) = \sum_{\ell} a^{(\ell)}(q,p) = a(q,p).$$

Vi vil imidlertid sjelden kjenne forholdet mellom totalantallet i fylket og regionen for den variabelen vi studerer. Det er da naturlig å innføre en inndeling i delklasser etter andre variable. Eksempelvis kan man som man har gjort ved etterhåndsstratifiseringen i AKU, inndele i klasser etter variablene kjønn og alder. Da vil $r^{(\ell)}(q,p)$ være det relative forholdet mellom antall personer i kjønns- og aldersklasse nr. ℓ i fylke q og fylkespar p. Hvis vi nå antar at

$$a^{(\ell)}(q,p) = \frac{N^{(\ell)}(q,p)}{N^{(\ell)}(p)}\ a^{(\ell)}(p),$$

vil estimatoren (4.3.7) være forventningsrett.

ENGLISH SUMMARY

This article presents a method for estimating the employment within small geographical areas.

The problem which is posed is the following: The standard sample survey of the Bureau is designed to produce statistics for the complete population of Norway. In the analysis of a sample survey one often desires, however, to produce statistics for smaller geographical areas. The fact that the survey is not simultaneously designed to make estimates for geographical regions and subdivisions of these complicates the analysis of the sample survey.

The stratification of the sample survey of the Bureau is carried out so that it is possible to construct unbiased estimates within regions which mainly consist of pair of counties. This fact has till now not been utilized in the production of statistics. In this article we have, however, studied estimates of employment within these geographical regions. The conclusion of our study is that it is possible to publish regional estimates within industries with more than 7 000 employees.

In order to make estimates for population characteristics within counties, we have constructed estimates which are synthesis of the regional estimates and a priori information of the counties. These estimates are called synthetic estimates and have earlier been studied by for instance Bureau of the Census.

In this article we have studied three classes of synthetic estimates. One set is based on a combination of regional estimates and information on the number of inhabitants in the counties, another is based on information on the age and sex distribution of the population of the counties and the third is based on information on the employment by industry.

Since the synthetic estimates are based, we have studied both the variance and the mean square error of the estimates. The numerical calculations are based on a population where the characteristics are known. The population is established from the Population and Housing Census 1970.

The numerical results indicate that there is no gain in the mean square error of the estimates by utilizing information on the sex and age distribution in addition to information on the number of inhabitants. This circumstance may be explained by the fact that the strata which from the regions are homogeneous with respect to the characteristics which are used as additional information for synthetic estimation. The essential conclusion of this part of the study is that within counties it is possible to publish estimates of employment within regional industries with more than 10 000 employees.

5. REFERANSER

Brenna, Svein & Jan M. Hoem: "Arbeidskraftsundersøkelsene bør ikke offentliggjøre oppblåste tall under ca. 10 000." Statistisk Sentralbyrå, Arbeidsnotat IO 73/6.

Cochran, William G. (1963): "Sampling Techniques", John Wiley & Sons, Inc., New York.

Dagsvik, John (1974): "Variansestimering for nivåtallestimater og endringstallestimater ved Byråets arbeidskraftsundersøkelser." Statistisk Sentralbyrå, Arbeidsnotat IO 74/50.

Dahmstrøm, Per & Karl-Erik Kristiansson (1976): "Nya urvals- og skattningsmetoder i arbetskraftsundersøkningarna (AKU)." Statistiska Centralbyrån, PM 1976-03-01.

Gonzalez, Maria E. & Joseph Waksberg (1973): "Estimation of the error of synthetic estimates." Notat presentert ved det 1. møte til International Association of Survey Statisticians, Wien, 18.-25. august 1975.

Laake, Petter (1974): "Estimering av totaler med en to-trinns utvalgsplan der de primære utvalgsområdene trekkes med ulik sannsynlighet i første trinn." Statistisk Sentralbyrå, Arbeidsnotat IO 74/49.

Laake, Petter, Lisa Forsén & Kjetil Sørlie (1974): "Orientering om bruk av Byråets program for estimering av gjennomsnitt og estimering av variansen til estimatoren i Byråets utvalgsundersøkelser." Maskinskrevet stensil (PL/TF, 24/4-1974).

National Center for Health Statistics (1968): "Synthetic estimates of disability." Public Health Service no. 1759.

Thomsen, Ib & Arne Rideng (1974): "Oversikt over arbeidet med ny utvalgsplan. Statistisk Sentralbyrå, Arbeidsnotat IO 74/25.

Vedleggstabeller

1. Sysselsetting etter næring i regioner. Beregningsbestand: 10 prosent av landets befolkning *Employment by industry within regions. Population: 10 per cent of the population of Norway*

Tabell 1.1. Sysselsetting etter næring. Region 1: Akershus *Employment by industry. Region 1: Akershus*

Næring *Industry*	Forventet syssel-setting *Expecta-tion of estimated employment*	Syssel-setting *Employ-ment*	Standard-avvik *Standard deviation*	Skjev-het *Bias*	Brutto-standard-avvik *Square root of mean square error*
Hele regionen *The whole region*	13 167	13 167	354	0	354
Jordbruk *Agriculture*	631	(631)	147	0	147
Skogbruk *Forestry* ...	105	(105)	49	0	49
Fiske og fangst *Hunting and fishing* ..	7	(7)	11	0	11
Bergverksdrift *Mining and quarrying* .	30	(30)	25	0	25
Industri *Manufac-turing*	3 175	3 175	280	0	280
Kraft- og vannfor-syning *Electricity, gas and water supply* .	136	(136)	53	0	53
Bygge- og anleggs-virksomhet *Construc-tion*	1 082	1 082	159	0	159
Varehandel, hotell- og restaurantvirksomhet *Wholesale and retail trade, restaurants and hotels*	2 373	2 373	221	0	221
Transport, lagring, post- og telekommu-nikasjoner *Transport, storage and commu-nication*	1 380	1 380	168	0	168
Bank- og finansier-ingsvirksomhet, for-sikringsvirksomhet, eiendomsdrift og for-retningsmessig tje-nesteyting *Financing, insurance, real estate and business services*	1 006	1 006	147	0	147
Offentlig, sosial og privat tjenesteyting *Community, social and personal services*	3 242	3 242	281	0	281

Tabell 1.2. Sysselsetting etter næring. Region 2: Oslo *Employment by industry. Region 2: Oslo*

Næring[1)]	Forventet sysselsetting	Sysselsetting	Standardavvik	Skjevhet	Bruttostandardavvik
Hele regionen	21 291	21 291	525	0	525
Jordbruk	47	(47)	37	0	37
Skogbruk	9	(9)	16	0	16
Fiske og fangst	5	(5)	12	0	12
Bergverksdrift	12	(12)	18	0	18
Industri	5 200	5 200	365	0	365
Kraft- og vannforsyning	236	(236)	83	0	83
Bygge- og anleggsvirksomhet ...	1 225	1 225	187	0	187
Varehandel, hotell- og restaurantvirksomhet	4 997	4 997	359	0	359
Transport, lagring, post- og telekommunikasjoner	2 233	2 233	250	0	250
Bank- og finansieringsvirksomhet, forsikringsvirksomhet, eiendomsdrift og forretningsmessig tjenesteyting	1 956	1 956	235	0	235
Offentlig, sosial og privat tjenesteyting	5 371	5 371	370	0	370

1) See English translation, table 1.1.

Tabell 1.3. Sysselsetting etter næring. Region 3: Hedmark/Oppland *Employment by industry. Region 3: Hedmark/Oppland*

Næring[1)]	Forventet sysselsetting	Sysselsetting	Standardavvik	Skjevhet	Bruttostandardavvik
Hele regionen	13 124	13 124	406	0	406
Jordbruk	2 407	2 407	271	0	271
Skogbruk	468	(468)	134	0	134
Fiske og fangst	3	(3)	8	0	8
Bergverksdrift	72	(72)	70	0	70
Industri	3 241	3 241	309	0	309
Kraft- og vannforsyning	146	(146)	61	0	61
Bygge- og anleggsvirksomhet ...	1 515	1 515	192	0	192
Varehandel, hotell- og restaurantvirksomhet	1 699	1 699	204	0	204
Transport, lagring, post- og telekommunikasjoner	1 012	1 012	170	0	170
Bank- og finansieringsvirksomhet, forsikringsvirksomhet, eiendomsdrift og forretningsmessig tjenesteyting	281	(281)	87	0	87
Offentlig, sosial og privat tjenesteyting	2 280	2 280	234	0	234

1) See English translation, table 1.1.

Tabell 1.4. Sysselsetting etter næring. Region 4: Buskerud/Telemark
Employment by industry. Region 4: Buskerud/Telemark

Næring[1)]	Forventet antall	Sysselsetting	Standardavvik	Skjevhet	Bruttostandardavvik
Hele regionen	13 102	13 102	494	0	494
Jordbruk	1 086	1 086	183	0	183
Skogbruk	208	(208)	77	0	77
Fiske og fangst	12	(12)	18	0	18
Bergverksdrift	54	(54)	36	0	36
Industri	4 758	4 758	315	0	315
Kraft- og vannforsyning	177	(177)	64	0	64
Bygge- og anleggsvirksomhet ...	1 254	1 254	170	0	170
Varehandel, hotell- og restaurantvirksomhet	1 806	1 806	201	0	201
Transport, lagring, post- og telekommunikasjoner	1 167	1 167	162	0	162
Bank- og finansieringsvirksomhet, forsikringsvirksomhet, eiendomsdrift og forretningsmessig tjenesteyting	387	(387)	93	0	93
Offentlig, sosial og privat tjenesteyting	2 193	2 193	227	0	227

1) See English translation, table 1.1.

Tabell 1.5. Sysselsetting etter næring. Region 5: Østfold/Vestfold
Employment by industry. Region 5: Østfold/Vestfold

Næring[1)]	Forventet antall	Sysselsetting	Standardavvik	Skjevhet	Bruttostandardavvik
Hele regionen	14 707	14 707	422	0	422
Jordbruk	924	924	174	0	174
Skogbruk	89	(89)	51	0	51
Fiske og fangst	47	(47)	37	0	37
Bergverksdrift	68	(68)	44	0	44
Industri	5 853	5 853	358	0	358
Kraft- og vannforsyning	159	(159)	60	0	60
Bygge- og anleggsvirksomhet ...	1 147	1 147	166	0	166
Varehandel, hotell- og restaurantvirksomhet	2 016	2 016	215	0	215
Transport, lagring, post- og telekommunikasjoner	1 443	1 443	176	0	176
Bank- og finansieringsvirksomhet, forsikringsvirksomhet, eiendomsdrift og forretningsmessig tjenesteyting	428	(428)	100	0	100
Offentlig, sosial og privat tjenesteyting	2 533	2 533	239	0	239

1) See English translation, table 1.1.

Tabell 1.6. Sysselsetting etter næring. Region 6: Agder/Rogaland
Employment by industry. Region 6: Agder/Rogaland

Næring[1)]	Forventet sysselsetting	Sysselsetting	Standardavvik	Skjevhet	Bruttostandardavvik
Hele regionen	16 774	16 774	437	0	437
Jordbruk	1 689	1 689	310	0	310
Skogbruk	94	(94)	72	0	72
Fiske og fangst	200	(200)	94	0	94
Bergverksdrift	130	(130)	110	0	110
Industri	4 980	4 980	369	0	369
Kraft- og vannforsyning	213	(213)	71	0	71
Bygge- og anleggsvirksomhet ...	1 568	1 568	201	0	201
Varehandel, hotell- og restaurantvirksomhet	2 433	2 433	225	0	225
Transport, lagring, post- og telekommunikasjoner	1 950	1 950	199	0	199
Bank- og finansieringsvirksomhet, forsikringsvirksomhet, eiendomsdrift og forretningsmessig tjenesteyting	516	(516)	105	0	105
Offentlig, sosial og privat tjenesteyting	3 001	3 001	258	0	258

1) See English translation, table 1.1.

Tabell 1.7. Sysselsetting etter næring. Region 7: Hordaland/Sogn og Fjordane *Employment by industry. Region 7: Hordaland/Sogn og Fjordane*

Næring[1)]	Forventet sysselsetting	Sysselsetting	Standardavvik	Skjevhet	Bruttostandardavvik
Hele regionen	17 229	17 229	450	0	450
Jordbruk	1 771	1 771	235	0	235
Skogbruk	12	(12)	16	0	16
Fiske og fangst	388	(388)	136	0	136
Bergverksdrift	23	(23)	23	0	23
Industri	4 505	4 505	374	0	374
Kraft- og vannforsyning	181	(181)	65	0	65
Bygge- og anleggsvirksomhet ...	1 786	1 786	195	0	195
Varehandel, hotell- og restaurantvirksomhet	2 523	2 523	235	0	235
Transport, lagring, post- og telekommunikasjoner	2 040	2 040	216	0	216
Bank- og finansieringsvirksomhet, forsikringsvirksomhet, eiendomsdrift og forretningsmessig tjenesteyting	619	619	115	0	115
Offentlig, sosial og privat tjenesteyting	3 381	3 381	273	0	273

1) See English translation, table 1.1.

Tabell 1.8. Sysselsetting etter næring. Region 8: Møre og Romsdal
Employment by industry. Region 8: Møre og Romsdal

Næring[1)]	Forventet sysselsetting	Sysselsetting	Standardavvik	Skjevhet	Bruttostandardavvik
Hele regionen	7 976	7 976	296	0	296
Jordbruk	976	976	175	0	175
Skogbruk	7	(7)	13	0	13
Fiske og fangst	454	(454)	129	0	129
Bergverksdrift	43	(43)	52	0	52
Industri	2 382	2 382	231	0	231
Kraft- og vannforsyning	80	(80)	42	0	42
Bygge- og anleggsvirksomhet ...	690	690	123	0	123
Varehandel, hotell- og restaurantvirksomhet	1 011	1 011	153	0	153
Transport, lagring, post- og telekommunikasjoner	913	913	153	0	153
Bank- og finansieringsvirksomhet, forsikringsvirksomhet, eiendomsdrift og forretningsmessig tjenesteyting	189	189	65	0	65
Offentlig, sosial og privat tjenesteyting	1 231	1 231	206	0	206

1) See English translation, table 1.1.

Tabell 1.9. Sysselsetting etter næring. Region 9: Trøndelag
Employment by industry. Region 9: Trøndelag

Næring[1)]	Forventet sysselsetting	Sysselsetting	Standardavvik	Skjevhet	Bruttostandardavvik
Hele regionen	12 592	12 592	374	0	374
Jordbruk	1 853	1 853	269	0	269
Skogbruk	122	(122)	72	0	72
Fiske og fangst	205	(205)	138	0	138
Bergverksdrift	150	(150)	134	0	134
Industri	2 580	2 580	279	0	279
Kraft- og vannforsyning	141	(141)	62	0	62
Bygge- og anleggsvirksomhet ...	1 108	1 108	169	0	169
Varehandel, hotell- og restaurantvirksomhet	1 972	1 972	217	0	217
Transport, lagring, post- og telekommunikasjoner	1 349	1 349	176	0	176
Bank- og finansieringsvirksomhet, forsikringsvirksomhet, eiendomsdrift og forretningsmessig tjenesteyting	384	(384)	92	0	92
Offentlig, sosial og privat tjenesteyting	2 728	2 728	252	0	252

1) See English translation, table 1.1.

Tabell 1.10. Sysselsetting etter næring. Region 10: Nordland
Employment by industry. Region 10: Nordland

Næring[1]	Forventet sysselsetting	Sysselsetting	Standardavvik	Skjevhet	Bruttostandardavvik
Hele regionen	8 491	8 491	459	0	459
Jordbruk	984	(984)	210	0	210
Skogbruk	38	(38)	51	0	51
Fiske og fangst	664	(664)	193	0	193
Bergverksdrift	161	(161)	109	0	109
Industri	1 521	(1 521)	253	0	253
Kraft- og vannforsyning	80	(80)	49	0	49
Bygge- og anleggsvirksomhet ...	819	819	157	0	157
Varehandel, hotell- og restaurantvirksomhet	1 040	1 040	165	0	165
Transport, lagring, post- og telekommunikasjoner	1 214	1 214	210	0	210
Bank- og finansieringsvirksomhet, forsikringsvirksomhet, eiendomsdrift og forretningsmessig tjenesteyting	164	(164)	61	0	61
Offentlig, sosial og privat tjenesteyting	1 806	1 806	223	0	223

1) See English translation, table 1.1.

Tabell 1.11. Sysselsetting etter næring. Region 11: Troms/Finnmark
Employment by industry. Region 11: Troms/Finnmark

Næring[1]	Forventet sysselsetting	Sysselsetting	Standardavvik	Skjevhet	Bruttostandardavvik
Hele regionen	7 702	7 702	299	0	299
Jordbruk	759	(759)	213	0	213
Skogbruk	16	(16)	21	0	21
Fiske og fangst	816	(816)	188	0	188
Bergverksdrift	210	(210)	94	0	94
Industri	1 132	1 132	220	0	220
Kraft- og vannforsyning	71	(71)	40	0	40
Bygge- og anleggsvirksomhet ...	791	(791)	166	0	166
Varehandel, hotell- og restaurantvirksomhet	1 046	1 046	167	0	167
Transport, lagring, post- og telekommunikasjoner	945	945	143	0	143
Bank- og finansieringsvirksomhet, forsikringsvirksomhet, eiendomsdrift og forretningsmessig tjenesteyting	172	(172)	61	0	61
Offentlig, sosial og privat tjenesteyting	1 744	1 744	265	0	265

1) See English translation, table 1.1.

2. Sysselsetting etter næring i fylkene, "nedbrutt"[1] etter folketall. Beregningsbestand: 10 prosent av landets befolkning *Employment by industry within counties, based on a priori information on the number of inhabitants. Population: 10 per cent of the population of Norway*

Tabell 2.1. Sysselsetting etter næring. Østfold *Employment by industry. Østfold*

Næring[2]	Forventet sysselsetting	Sysselsetting	Standardavvik	Skjevhet	Bruttostandardavvik
Hele fylket *The whole county* .	8 217	8 440	235	-223	323
Jordbruk	516	(587)	97	-71	120
Skogbruk	49	(54)	28	-5	29
Fiske og fangst	26	(32)	20	-6	21
Bergverksdrift	38	(18)	24	-20	31
Industri	3 270	3 555	200	-285	347
Kraft- og vannforsyning	88	(114)	33	-26	42
Bygge- og anleggsvirksomhet ...	640	665	93	-25	96
Varehandel, hotell- og restaurantvirksomhet	1 126	1 144	120	-18	121
Transport, lagring, post- og telekommunikasjoner	806	(682)	98	124	158
Bank- og finansieringsvirksomhet, forsikringsvirksomhet, eiendomsdrift og forretningsmessig tjenesteyting	239	(243)	56	-4	56
Offentlig, sosial og privat tjenesteyting	1 415	1 347	133	68	150

1) For en nærmere presisering av "nedbrutt", se avsnitt 2.
1) For a definition of "breakdown", see section 2. 2) See English translation, table 1.1.

Tabell 2.2. Sysselsetting etter næring. Hedmark *Employment by industry. Hedmark*

Næring[1]	Forventet sysselsetting	Sysselsetting	Standardavvik	Skjevhet	Bruttostandardavvik
Hele fylket	6 706	6 696	207	10	208
Jordbruk	1 229	1 077	138	152	206
Skogbruk	239	(348)	68	-108	128
Fiske og fangst	1	(1)	4	0	4
Bergverksdrift	36	(33)	36	3	37
Industri	1 656	1 687	158	-31	161
Kraft- og vannforsyning	74	(53)	31	21	38
Bygge- og anleggsvirksomhet	774	784	98	-10	99
Varehandel, hotell- og restaurantvirksomhet	868	865	104	3	105
Transport, lagring, post- og telekommunikasjoner	517	528	87	-11	88
Bank- og finansieringsvirksomhet, forsikringsvirksomhet, eiendomsdrift og forretningsmessig tjenesteyting	143	(154)	44	-11	46
Offentlig, sosial og privat tjenesteyting	1 165	1 166	119	-1	119

1) See English translation, table 1.1.

Tabell 2.3. Sysselsetting etter næring. Oppland *Employment by industry. Oppland*

Næring[1]	Forventet sysselsetting	Sysselsetting	Standardavvik	Skjevhet	Bruttostandardavvik
Hele fylket	6 418	6 427	198	-10	199
Jordbruk	1 177	1 330	132	-153	202
Skogbruk	228	(120)	65	108	127
Fiske og fangst	1	(2)	4	-1	4
Bergverksdrift	35	(139)	34	-4	35
Industri	1 584	1 554	151	30	154
Kraft- og vannforsyning	71	(93)	30	-22	37
Bygge- og anleggsvirksomhet	740	731	94	9	95
Varehandel, hotell- og restaurantvirksomhet	830	834	100	-4	100
Transport, lagring, post- og telekommunikasjoner	494	484	83	10	84
Bank- og finansieringsvirksomhet, forsikringsvirksomhet, eiendomsdrift og forretningsmessig tjenesteyting	137	(127)	42	10	44
Offentlig, sosial og privat tjenesteyting	1 115	1 114	114	1	115

1) See English translation, table 1.1.

Tabell 2.4. Sysselsetting etter næring. Buskerud *Employment by industry. Buskerud*

Næring[1]	Forventet sysselsetting	Sysselsetting	Standardavvik	Skjevhet	Bruttostandardavvik
Hele fylket	7 337	7 520	276	-183	331
Jordbruk	608	630	102	-22	105
Skogbruk	116	(91)	43	25	50
Fiske og fangst	6	(4)	10	2	10
Bergverksdrift	30	(24)	20	6	21
Industri	2 664	2 745	176	-81	194
Kraft- og vannforsyning	99	(103)	36	-4	36
Bygge- og anleggsvirksomhet	702	672	95	30	100
Varehandel, hotell- og restaurantvirksomhet	1 011	1 105	112	-94	146
Transport, lagring, post- og telekommunikasjoner	653	661	90	8	91
Bank- og finansieringsvirksomhet, forsikringsvirksomhet, eiendomsdrift og forretningsmessig tjenesteyting .	216	(227)	52	-11	53
Offentlig, sosial og privat tjenesteyting	1 228	1 258	127	-30	130

1) See English translation, table 1.1.

Tabell 2.5. Sysselsetting etter næring. Vestfold *Employment by industry. Vestfold*

Næring[1]	Forventet sysselsetting	Sysselsetting	Standardavvik	Skjevhet	Bruttostandardavvik
Hele fylket	6 489	6 266	185	223	291
Jordbruk	407	(337)	77	70	105
Skogbruk	39	(35)	22	4	23
Fiske og fangst	20	(15)	16	5	17
Bergverksdrift	30	(50)	19	-20	28
Industri	2 582	2 298	158	284	326
Kraft- og vannforsyning	70	(45)	26	25	37
Bygge- og anleggsvirksomhet	506	482	73	24	77
Varehandel, hotell- og restaurantvirksomhet	889	872	94	17	97
Transport, lagring, post- og telekommunikasjoner	636	761	78	-125	146
Bank- og finansieringsvirksomhet, forsikringsvirksomhet, eiendomsdrift og forretningsmessig tjenesteyting .	188	(185)	44	3	44
Offentlig, sosial og privat tjenesteyting	1 117	1 186	105	-69	126

1) See English translation, table 1.1.

Tabell 2.6. Sysselsetting etter næring. Telemark *Employment by industry. Telemark*

Næring[1)]	Forventet sysselsetting	Sysselsetting	Standardavvik	Skjevhet	Bruttostandardavvik
Hele fylket *The whole county* ..	5 764	5 582	217	182	283
Jordbruk	477	456	80	21	83
Skogbruk	91	(117)	34	-26	43
Fiske og fangst	5	(8)	8	-3	9
Bergverksdrift	23	(30)	15	-7	17
Industri	2 093	2 013	138	80	160
Kraft-og vannforsyning	77	(74)	28	3	28
Bygge- og anleggsvirksomhet	551	582	74	-31	81
Varehandel, hotell- og restaurantvirksomhet	794	701	88	93	129
Transport, lagring, post- og telekommunikasjoner	513	506	71	7	71
Bank- og finansieringsvirksomhet, forsikringsvirksomhet, eiendomsdrift og forretningsmessig tjenesteyting .	170	(160)	41	10	42
Offentlig, sosial og privat tjenesteyting	964	935	100	29	104

1) See English translation, table 1.1.

Tabell 2.7. Sysselsetting etter næring. Aust-Agder *Employment by industry. Aust-Agder*

Næring[1)]	Forventet sysselsetting	Sysselsetting	Standardavvik	Skjevhet	Bruttostandardavvik
Hele fylket	2 904	2 802	75	102	127
Jordbruk	292	(235)	53	57	79
Skogbruk	16	(72)	12	56	57
Fiske og fangst	34	(9)	16	25	30
Bergverksdrift	22	(16)	19	6	20
Industri	862	742	63	120	136
Kraft-og vannforsyning	36	(43)	12	-7	14
Bygge- og anleggsvirksomhet	271	306	34	-35	49
Varehandel, hotell- og restaurantvirksomhet	421	375	39	46	61
Transport, lagring, post- og telekommunikasjoner	337	418	34	-81	88
Bank- og finansieringsvirksomhet, forsikringsvirksomhet, eiendomsdrift og forretningsmessig tjenesteyting	89	(98)	18	-9	20
Offentlig, sosial og privat tjenesteyting	519	488	44	31	55

1) See English translation, table 1.1.

Tabell 2.8. Sysselsetting etter næring. Vest-Agder *Employment by industry. Vest-Agder*

Næring[1]	Forventet sysselsetting	Sysselsetting	Standardavvik	Skjevhet	Bruttostandardavvik
Hele fylket	4 392	4 370	114	22	117
Jordbruk	442	(318)	81	124	148
Skogbruk	24	(18)	19	6	20
Fiske og fangst	52	(42)	24	10	27
Bergverksdrift	34	(20)	28	6	32
Industri	1 304	1 324	96	-20	99
Kraft- og vannforsyning	55	(63)	18	-8	20
Bygge- og anleggsvirksomhet	410	500	52	-90	104
Varehandel, hotell- og restaurantvirksomhet	637	640	59	-3	59
Transport, lagring, post- og telekommunikasjoner	510	491	52	19	56
Bank- og finansieringsvirksomhet, forsikringsvirksomhet, eiendomsdrift og forretningsmessig tjenesteyting	135	(126)	27	9	29
Offentlig, sosial og privat tjenesteyting	785	828	67	-43	80

1) See English translation, table 1.1.

Tabell 2.9. Sysselsetting etter næring. Rogaland *Employment by industry. Rogaland*

Næring[1]	Forventet sysselsetting	Sysselsetting	Standardavvik	Skjevhet	Bruttostandardavvik
Hele fylket	9 477	9 602	247	-125	277
Jordbruk	954	(1 136)	175	-182	252
Skogbruk	53	(4)	41	49	64
Fiske og fangst	113	(149)	53	-36	65
Bergverksdrift	73	(94)	62	-19	66
Industri	2 813	2 914	208	-101	231
Kraft- og vannforsyning	120	(107)	40	13	43
Bygge- og anleggsvirksomhet	886	(762)	113	124	168
Varehandel, hotell- og restaurantvirksomhet	1 374	1 418	127	-44	135
Transport, lagring, post- og telekommunikasjoner	1 101	1 041	112	60	128
Bank- og finansieringsvirksomhet, forsikringsvirksomhet, eiendomsdrift og forretningsmessig tjenesteyting	291	292	59	-1	60
Offentlig, sosial og privat tjenesteyting	1 695	1 685	146	10	147

1) See English translation, table 1.1.

Tabell 2.10. Sysselsetting etter næring. Hordaland *Employment by industry. Hordaland*

Næring[1]	Forventet sysselsetting	Sysselsetting	Standardavvik	Skjevhet	Bruttostandardavvik
Hele fylket	13 595	13 502	355	93	367
Jordbruk	1 397	(803)	185	594	623
Skogbruk	9	(9)	13	0	13
Fiske og fangst	306	(213)	107	93	143
Bergverksdrift	18	(18)	18	0	18
Industri	3 554	3 708	295	-154	332
Kraft- og vannforsyning	142	(147)	51	-5	52
Bygge- og anleggsvirksomhet	1 409	1 348	154	61	166
Varehandel, hotell- og restaurantvirksomhet	1 990	2 165	185	-175	255
Transport, lagring, post- og telekommunikasjoner	1 609	1 694	171	-85	191
Bank- og finansieringsvirksomhet, forsikringsvirksomhet, eiendomsdrift og forretningsmessig tjenesteyting	488	(568)	91	-80	121
Offentlig, sosial og privat tjenesteyting	2 667	2 829	215	-161	269

1) See English translation, table 1.1.

Tabell 2.11. Sysselsetting etter næring. Sogn og Fjordane *Employment by industry. Sogn og Fjordane*

Næring[1]	Forventet sysselsetting	Sysselsetting	Standardavvik	Skjevhet	Bruttostandardavvik
Hele fylket	3 633	3 727	95	-94	133
Jordbruk	373	(968)	49	-595	596
Skogbruk	2	(3)	3	-1	4
Fiske og fangst	81	(175)	28	-94	98
Bergverksdrift	4	(5)	5	-1	5
Industri	950	797	78	153	172
Kraft- og vannforsyning	38	(34)	13	4	15
Bygge- og anleggsvirksomhet	376	438	41	-62	74
Varehandel, hotell- og restaurantvirksomhet	532	(358)	49	174	181
Transport, lagring, post- og telekommunikasjoner	430	(346)	45	84	96
Bank- og finansieringsvirksomhet, forsikringsvirksomhet, eiendomsdrift og forretningsmessig tjenesteyting	130	(51)	24	79	83
Offentlig, sosial og privat tjenesteyting	713	(552)	57	161	171

1) See English translation, table 1.1.

Tabell 2.12. Sysselsetting etter næring. Sør-Trøndelag *Employment by industry. Sør-Trøndelag*

Næring[1]	Forventet sysselsetting	Sysselsetting	Standardavvik	Skjevhet	Bruttostandardavvik
Hele fylket	8 474	8 550	252	-76	263
Jordbruk	1 247	(930)	181	317	365
Skogbruk	82	(36)	48	46	67
Fiske og fangst	138	(154)	92	-16	94
Bergverksdrift	100	(77)	90	23	93
Industri	1 736	1 803	188	-67	199
Kraft- og vannforsyning	94	(86)	42	8	43
Bygge- og anleggsvirksomhet	745	729	113	16	115
Varehandel, hotell- og restaurantvirksomhet	1 327	1 502	146	-175	228
Transport, lagring, post- og telekommunikasjoner	907	976	118	-67	137
Bank- og finansieringsvirksomhet, forsikringsvirksomhet, eiendomsdrift og forretningsmessig tjenesteyting	258	(333)	62	-75	97
Offentlig, sosial og privat tjenesteyting	1 835	1 924	170	-89	191

1) See English translation, table 1.1.

Tabell 2.13. Sysselsetting etter næring. Nord-Trøndelag *Employment by industry. Nord-Trøndelag*

Næring[1]	Forventet sysselsetting	Sysselsetting	Standardavvik	Skjevhet	Bruttostandardavvik
Hele fylket	4 118	4 042	122	76	144
Jordbruk	606	(923)	88	-317	329
Skogbruk	39	(86)	23	-47	52
Fiske og fangst	67	(51)	45	16	48
Bergverksdrift	49	(73)	43	-24	50
Industri	843	777	91	66	113
Kraft- og vannforsyning	46	(55)	20	-9	22
Bygge- og anleggsvirksomhet	362	379	55	-17	58
Varehandel, hotell- og restaurantvirksomhet	644	(470)	71	174	189
Transport, lagring, post- og telekommunikasjoner	441	373	57	68	89
Bank- og finansieringsvirksomhet, forsikringsvirksomhet, eiendomsdrift og forretningsmessig tjenesteyting	125	(51)	30	74	80
Offentlig, sosial og privat tjenesteyting	892	804	82	88	121

1) See English translation, table 1.1.

Tabell 2.14. Sysselsetting etter næring. Troms *Employment by industry. Troms*

Næring[1]	Forventet syssel-setting	Syssel-setting	Stan-dard-avvik	Skjev-het	Brutto-stan-dard-avvik
Hele fylket	4 979	4 965	193	14	194
Jordbruk	490	(567)	138	-77	157
Skogbruk	10	(8)	13	2	13
Fiske og fangst	527	(479)	121	48	131
Bergverksdrift	135	(62)	61	73	95
Industri	731	(646)	142	85	166
Kraft- og vannforsyning	45	(41)	26	4	27
Bygge- og anleggsvirksomhet	511	(546)	107	-35	113
Varehandel, hotell- og restaurantvirksomhet	676	687	108	-11	108
Transport, lagring, post- og telekommunikasjoner	610	620	92	-10	93
Bank- og finansieringsvirksomhet, forsikringsvirksomhet, eiendomsdrift og forretningsmessig tjenesteyting	111	(130)	39	-19	44
Offentlig, sosial og privat tjenesteyting	1 127	1 179	171	-52	179

1) See English translation, table 1.1.

Tabell 2.15. Sysselsetting etter næring. Finnmark *Employment by industry. Finnmark*

Næring[1]	Forventet syssel-setting	Syssel-setting	Stan-dard-avvik	Skjev-het	Brutto-stan-dard-avvik
Hele fylket	2 722	2 737	105	-15	106
Jordbruk	268	(192)	75	76	107
Skogbruk	5	(8)	7	-3	8
Fiske og fangst	288	(337)	66	-49	82
Bergverksdrift	74	(148)	33	-74	81
Industri	400	(486)	78	-86	116
Kraft- og vannforsyning	25	(30)	14	-5	15
Bygge- og anleggsvirksomhet	279	(245)	58	34	68
Varehandel, hotell- og restaurantvirksomhet	369	359	59	10	60
Transport, lagring, post- og telekommunikasjoner	334	325	50	9	51
Bank- og finansieringsvirksomhet, forsikringsvirksomhet, eiendomsdrift og forretningsmessig tjenesteyting	60	(42)	21	18	28
Offentlig, sosial og privat tjenesteyting	616	(565)	94	51	107

1) See English translation, table 1.1.

3. Sysselsetting i fylkene, "nedbrutt"[1] etter andel i hver næring. Beregningsbestand: 10 prosent av landets befolkning *Employment within counties, based on a priori information on the employment by industry. Population: 10 per cent of the population of Norway*

Tabell 3.1. Sysselsetting i fylkene *Employment within counties*

Fylke *County*	Forventet sysselsetting *Expectation of estimated employment*	Sysselsetting *Employment*	Standardavvik *Standard deviation*	Skjevhet *Bias*	Bruttostandardavvik *Square root of mean square error*
Østfold	8 440	8 440	227	0	227
Akershus	13 167	13 167	354	0	354
Oslo	21 291	21 291	525	0	525
Hedmark	6 696	6 696	210	0	210
Oppland	6 428	6 428	204	0	204
Buskerud	7 520	7 520	214	0	214
Vestfold	6 266	6 266	167	0	167
Telemark	5 582	5 582	160	0	160
Aust-Agder	2 802	2 802	84	0	84
Vest-Agder	4 370	4 370	113	0	113
Rogaland	9 602	9 602	274	0	274
Hordaland	13 502	13 502	368	0	368
Sogn og Fjordane	3 727	3 727	126	0	126
Møre og Romsdal	7 976	7 976	296	0	296
Sør-Trøndelag	8 550	8 550	258	0	258
Nord-Trøndelag	4 042	4 042	141	0	141
Nordland	8 491	8 491	531	0	531
Troms	4 965	4 965	192	0	192
Finnmark	2 737	2 737	123	0	123

1) For en nærmere presisering av "nedbrutt", se avsnitt 2.
1) For a definition of "breakdown", see section 2.

4. Sysselsetting etter næring i fylkene, "nedbrutt"[1] etter alders- og kjønnsklasser. Beregningsbestand: 10 prosent av landets befolkning
Employment by industry within counties, based on a priori information on sex and age distribution. Population: 10 per cent of the population of Norway

Tabell 4.1. Sysselsetting etter næring. Østfold *Employment by industry. Østfold*

Næring[2]	Forventet sysselsetting	Sysselsetting	Standardavvik	Skjevhet	Bruttostandardavvik
Hele fylket *The whole county* .	8 238	8 440	213	-202	294
Jordbruk	521	(587)	98	-66	118
Skogbruk	50	(54)	29	-4	29
Fiske og fangst	26	(32)	21	-6	22
Bergverksdrift	38	(18)	25	-20	32
Industri	3 278	3 555	195	-277	339
Kraft og vannforsyning	89	(114)	34	-25	42
Bygge- og anleggsvirksomhet ...	642	665	93	-23	96
Varehandel, hotell- og restaurantvirksomhet	1 128	1 114	119	-16	120
Transport, lagring, post- og telekommunikasjoner	806	(682)	98	124	158
Bank- og finansieringsvirksomhet, forsikringsvirksomhet, eiendomsdrift og forretningsmessig tjenesteyting	239	(243)	56	-4	56
Offentlig, sosial og privat tjenesteyting	1 416	1 347	133	69	150

1) For en nærmere presisering av "nedbrutt", se avsnitt 2.
1) For a definition of "breakdown", see section 2. 2) See English translation, table 1.1.

Tabell 4.2. Sysselsetting etter næring. Hedmark *Employment by industry. Hedmark*

Næring[1]	Forventet sysselsetting	Sysselsetting	Standardavvik	Skjevhet	Bruttostandardavvik
Hele fylket	6 714	6 696	208	18	209
Jordbruk	1 241	(1 077)	140	164	216
Skogbruk	240	(348)	69	-108	128
Fiske og fangst	1	(1)	4	0	4
Bergverksdrift	36	(33)	36	3	36
Industri	1 651	1 687	158	-36	162
Kraft- og vannforsyning	74	(53)	32	21	40
Bygge- og anleggsvirksomhet ...	774	784	99	-10	100
Varehandel, hotell- og restaurantvirksomhet	867	865	104	2	104
Transport, lagring, post- og telekommunikasjoner	516	528	87	-12	88
Bank- og finansieringsvirksomhet, forsikringsvirksomhet, eiendomsdrift og forretningsmessig tjenesteyting	143	(154)	44	-11	45
Offentlig, sosial og privat tjenesteyting	1 166	1 166	120	0	120

1) See English translation, table 1.1.

Tabell 4.3. Sysselsetting etter næring. Oppland *Employment by industry. Oppland*

Næring[1]	Forventet sysselsetting	Sysselsetting	Standardavvik	Skjevhet	Bruttostandardavvik
Hele fylket	6 409	6 428	198	-17	199
Jordbruk	1 165	1 330	132	-165	211
Skogbruk	227	(120)	66	107	126
Fiske og fangst	1	(2)	4	-1	4
Bergverksdrift	35	(39)	34	-4	34
Industri	1 589	1 554	151	35	155
Kraft- og vannforsyning	71	(93)	30	-22	37
Bygge- og anleggsvirksomhet ...	740	731	94	9	95
Varehandel, hotell- og restaurantvirksomhet	831	834	100	-3	100
Transport, lagring, post- og telekommunikasjoner	495	484	84	11	85
Bank- og finansieringsvirksomhet, forsikringsvirksomhet, eiendomsdrift og forretningsmessig tjenesteyting	137	(127)	43	10	44
Offentlig, sosial og privat tjenesteyting	1 113	1 114	114	-1	114

1) See English translation, table 1.1.

Tabell 4.4. Sysselsetting etter næring. Buskerud *Employment by industry. Buskerud*

Næring[1)]	Forventet sysselsetting	Sysselsetting	Standardavvik	Skjevhet	Bruttostandardavvik
Hele fylket	7 350	7 520	238	-170	292
Jordbruk	607	630	96	-23	99
Skogbruk	116	(91)	44	25	50
Fiske og fangst	6	(4)	10	2	10
Bergverksdrift	30	(24)	20	6	21
Industri	2 670	2 745	132	-75	152
Kraft- og vannforsyning	98	(103)	36	-5	36
Bygge- og anleggsvirksomhet ...	703	672	91	31	96
Varehandel, hotell- og restaurantvirksomhet	1 012	1 105	103	-93	139
Transport, lagring, post- og telekommunikasjoner	655	672	88	-17	90
Bank- og finansieringsvirksomhet, forsikringsvirksomhet, eiendomsdrift og forretningsmessig tjenesteyting	217	(227)	51	-10	52
Offentlig, sosial og privat tjenesteyting	1 230	1 258	99	-28	103

1) See English translation, table 1.1.

Tabell 4.5. Sysselsetting etter næring. Vestfold *Employment by industry. Vestfold*

Næring[1)]	Forventet sysselsetting	Sysselsetting	Standardavvik	Skjevhet	Bruttostandardavvik
Hele fylket	6 468	6 266	168	198	260
Jordbruk	402	(337)	76	65	100
Skogbruk	38	(35)	22	3	22
Fiske og fangst	20	(15)	16	5	16
Bergverksdrift	29	(50)	19	-21	28
Industri	2 574	2 298	153	276	316
Kraft- og vannforsyning	69	(45)	27	24	36
Bygge- og anleggsvirksomhet ...	504	482	73	22	77
Varehandel, hotell- og restaurantvirksomhet	887	872	94	15	95
Transport, lagring, post- og telekommunikasjoner	636	761	78	-125	147
Bank- og finansieringsvirksomhet, forsikringsvirksomhet, eiendomsdrift og forretningsmessig tjenesteyting	188	(185)	44	3	44
Offentlig, sosial og privat tjenesteyting	1 116	1 186	104	-70	125

1) See English translation, table 1.1.

Tabell 4.6. Sysselsetting etter næring. Telemark *Employment by industry. Telemark*

Næring[1]	Forventet sysselsetting	Sysselsetting	Standardavvik	Skjevhet	Bruttostandardavvik
Hele fylket	5 751	5 582	186	169	251
Jordbruk	478	456	76	32	83
Skogbruk	91	(117)	34	-26	43
Fiske og fangst	5	(8)	8	-3	10
Bergverksdrift	23	(30)	16	-7	17
Industri	2 087	2 013	103	74	127
Kraft- og vannforsyning	78	(74)	28	4	28
Bygge- og anleggsvirksomhet	550	582	71	-32	78
Varehandel, hotell- og restaurantvirksomhet	793	701	81	92	115
Transport, lagring, post- og telekommunikasjoner	511	506	68	5	68
Bank- og finansieringsvirksomhet, forsikringsvirksomhet, eiendomsdrift og forretningsmessig tjenesteyting	169	(160)	39	9	40
Offentlig, sosial og privat tjenesteyting	962	935	77	27	82

1) See English translation, table 1.1.

Tabell 4.7. Sysselsetting etter næring. Aust-Agder *Employment by industry. Aust-Agder*

Næring[1]	Forventet sysselsetting	Sysselsetting	Standardavvik	Skjevhet	Bruttostandardavvik
Hele fylket	2 885	2 802	75	83	112
Jordbruk	304	(235)	56	69	89
Skogbruk	16	(72)	13	-56	58
Fiske og fangst	35	(9)	17	26	31
Bergverksdrift	22	(16)	19	-6	20
Industri	858	742	64	116	132
Kraft- og vannforsyning	37	(43)	12	-6	13
Bygge- og anleggsvirksomhet	272	306	35	-34	49
Varehandel, hotell- og restaurantvirksomhet	409	375	38	34	51
Transport, lagring, post- og telekommunikasjoner	332	418	34	-86	92
Bank- og finansieringsvirksomhet, forsikringsvirksomhet, eiendomsdrift og forretningsmessig tjenesteyting	87	(98)	19	-11	22
Offentlig, sosial og privat tjenesteyting	509	488	44	21	49

1) See English translation, table 1.1.

Tabell 4.8. Sysselsetting etter næring. Vest-Agder *Employment by industry. Vest-Agder*

Næring[1]	Forventet sysselsetting	Sysselsetting	Standardavvik	Skjevhet	Bruttostandardavvik
Hele fylket	4 377	4 370	115	7	115
Jordbruk	443	(318)	81	125	148
Skogbruk	24	(18)	19	6	20
Fiske og fangst	52	(42)	25	10	27
Bergverksdrift	33	(20)	29	13	32
Industri	1 299	1 324	96	-25	99
Kraft- og vannforsyning	55	(63)	19	-8	21
Bygge- og anleggsvirksomhet	408	500	52	-92	106
Varehandel, hotell- og restaurantvirksomhet	636	640	59	-4	59
Transport, lagring, post- og telekommunikasjoner	507	491	52	16	54
Bank- og finansieringsvirksomhet, forsikringsvirksomhet, eiendomsdrift og forretningsmessig tjenesteyting	134	(126)	27	8	28
Offentlig, sosial og privat tjenesteyting	782	828	67	-46	81

1) See English translation, table 1.1.

Tabell 4.9. Sysselsetting etter næring. Rogaland *Employment by industry. Rogaland*

Næring[1]	Forventet sysselsetting	Sysselsetting	Standardavvik	Skjevhet	Bruttostandardavvik
Hele fylket	9 510	9 602	248	-92	264
Jordbruk	941	(1 136)	173	-195	260
Skogbruk	52	(4)	41	48	64
Fiske og fangst	112	(149)	53	-37	65
Bergverksdrift	73	(94)	62	21	66
Industri	2 822	2 914	209	-92	228
Kraft- og vannforsyning	120	(107)	40	13	42
Bygge- og anleggsvirksomhet	887	(762)	114	125	169
Varehandel, hotell- og restaurantvirksomhet	1 386	1 418	129	-32	133
Transport, lagring, post- og telekommunikasjoner	1 110	1 041	114	69	133
Bank- og finansieringsvirksomhet, forsikringsvirksomhet, eiendomsdrift og forretningsmessig tjenesteyting	294	(292)	60	2	60
Offentlig, sosial og privat tjenesteyting	709	1 685	147	24	149

1) See English translation, table 1.1.

Tabell 4.10. Sysselsetting etter næring. Hordaland *Employment by industry. Hordaland*

Næring[1]	Forventet sysselsetting	Sysselsetting	Standardavvik	Skjevhet	Bruttostandardavvik
Hele fylket	13 538	13 502	355	36	357
Jordbruk	1 376	(803)	184	573	602
Skogbruk	9	(9)	13	0	13
Fiske og fangst	303	(213)	107	90	140
Bergverksdrift	18	(18)	19	0	19
Industri	3 535	3 708	294	-173	341
Kraft- og vannforsyning	141	(147)	51	-6	52
Bygge- og anleggsvirksomhet	1 398	1 348	153	50	161
Varehandel, hotell- og restaurantvirksomhet	1 991	2 165	186	-174	254
Transport, lagring, post- og telekommunikasjoner	1 602	1 694	171	-92	194
Bank- og finansieringsvirksomhet, forsikringsvirksomhet, eiendomsdrift og forretningsmessig tjenesteyting	490	(568)	92	-78	121
Offentlig, sosial og privat tjenesteyting	2 670	2 829	216	-159	268

1) See English translation, table 1.1.

Tabell 4.11. Sysselsetting etter næring. Sogn og Fjordane *Employment by industry. Sogn og Fjordane*

Næring[1]	Forventet sysselsetting	Sysselsetting	Standardavvik	Skjevhet	Bruttostandardavvik
Hele fylket	3 690	3 727	96	-37	103
Jordbruk	394	(968)	52	-574	576
Skogbruk	2	(3)	4	-1	4
Fiske og fangst	84	174	30	-91	96
Bergverksdrift	4	(5)	5	-1	5
Industri	969	(797)	80	181	198
Kraft- og vannforsyning	39	(34)	14	5	15
Bygge- og anleggsvirksomhet	387	438	42	-51	66
Varehandel, hotell- og restaurantvirksomhet	531	(358)	49	173	180
Transport, lagring, post- og telekommunikasjoner	437	(346)	46	91	102
Bank- og finansieringsvirksomhet, forsikringsvirksomhet, eiendomsdrift og forretningsmessig tjenesteyting	128	(51)	24	76	80
Offentlig, sosial og privat tjenesteyting	710	(552)	58	158	168

1) See English translation, table 1.1.

Tabell 4.12. Sysselsetting etter næring. Sør-Trøndelag *Employment by industry. Sør-Trøndelag*

Næring[1)]	Forventet sysselsetting	Sysselsetting	Standardavvik	Skjevhet	Bruttostandardavvik
Hele fylket	8 485	8 550	252	-65	260
Jordbruk	1 237	(930)	180	307	356
Skogbruk	81	(36)	48	45	66
Fiske og fangst	137	(154)	92	-17	94
Bergverksdrift	100	(77)	89	23	92
Industri	1 734	1 803	188	-69	200
Kraft- og vannforsyning	94	(86)	42	8	43
Bygge- og anleggsvirksomhet ...	743	729	113	14	114
Varehandel, hotell- og restaurantvirksomhet	1 334	1 502	147	-168	223
Transport, lagring, post- og telekommunikasjoner	909	976	119	-67	137
Bank- og finansieringsvirksomhet, forsikringsvirksomhet, eiendomsdrift og forretningsmessig tjenesteyting	260	(333)	63	-73	96
Offentlig, sosial og privat tjenesteyting	1 848	1 924	171	-76	187

1) See English translation, table 1.1.

Tabell 4.13. Sysselsetting etter næring. Nord-Trøndelag *Employment by industry. Nord-Trøndelag*

Næring[1)]	Forventet sysselsetting	Sysselsetting	Standardavvik	Skjevhet	Bruttostandardavvik
Hele fylket	4 107	4 042	123	65	139
Jordbruk	615	(923)	89	-308	320
Skogbruk	40	(86)	24	-46	52
Fiske og fangst	67	(51)	46	16	48
Bergverksdrift	49	(73)	44	-24	50
Industri	845	777	92	68	114
Kraft- og vannforsyning	46	(55)	20	-9	22
Bygge- og anleggsvirksomhet ...	346	379	56	-15	58
Varehandel, hotell- og restaurantvirksomhet	637	(470)	70	167	181
Transport, lagring, post- og telekommunikasjoner	439	373	58	66	88
Bank- og finansieringsvirksomhet, forsikringsvirksomhet, eiendomsdrift og forretningsmessig tjenesteyting	123	(51)	29	72	78
Offentlig, sosial og privat tjenesteyting	879	804	81	75	110

1) See English translation, table 1.1.

Tabell 4.14. Sysselsetting etter næring. Troms *Employment by industry. Troms*

Næring[1)]	Forventet sysselsetting	Sysselsetting	Standardavvik	Skjevhet	Bruttostandardavvik
Hele fylket	4 935	4 965	191	-30	193
Jordbruk	492	(567)	137	-75	156
Skogbruk	10	(8)	14	2	14
Fiske og fangst	521	(479)	120	42	127
Bergverksdrift	134	(62)	60	72	94
Industri	722	(646)	140	76	159
Kraft- og vannforsyning	45	(41)	26	4	27
Bygge- og anleggsvirksomhet ...	505	(546)	106	-41	114
Varehandel, hotell- og restaurantvirksomhet	669	687	107	-18	109
Transport, lagring, post- og telekommunikasjoner	604	620	92	-16	93
Bank- og finansieringsvirksomhet, forsikringsvirksomhet, eiendomsdrift og forretningsmessig tjenesteyting	109	(130)	39	-21	44
Offentlig, sosial og privat tjenesteyting	1 119	1 179	170	-60	180

1) See English translation, table 1.1.

Tabell 4.15. Sysselsetting etter næring. Finnmark *Employment by industry. Finnmark*

Næring[1)]	Forventet sysselsetting	Sysselsetting	Standardavvik	Skjevhet	Bruttostandardavvik
Hele fylket	2 766	2 737	109	29	110
Jordbruk	266	(192)	76	74	106
Skogbruk	5	(8)	7	-3	8
Fiske og fangst	294	(337)	68	43	80
Bergverksdrift	75	(148)	34	73	81
Industri	409	(486)	80	-77	111
Kraft- og vannforsyning	20	(30)	15	-10	18
Bygge- og anleggsvirksomhet ...	285	(245)	60	40	72
Varehandel, hotell- og restaurantvirksomhet	376	359	60	17	62
Transport, lagring, post- og telekommunikasjoner	340	325	52	15	54
Bank- og finansieringsvirksomhet, forsikringsvirksomhet, eiendomsdrift og forretningsmessig tjenesteyting	62	(42)	22	20	30
Offentlig, sosial og privat tjenesteyting	624	(565)	99	59	115

1) See English translation, table 1.1.

Utkommet i serien ART

Issued in the series Artikler fra Statistisk Sentralbyrå (ART)

Nr. 78 Inger Gabrielsen: Aktuelle skattetall *Current Tax Data* 1975 45 s. kr 8,00

" 79 Idar Møglestue: Befolkningens utdanningsbakgrunn En analyse av tall fra Folketelling 1970 *Educational Background of the Population An Analysis of Data from Population Census 1970* 1975 89 s. kr 8,00

" 80 Knut Ø. Sørensen: Statistisk Sentralbyrås befolkningsprognosemodell ved de regionale framskrivinger 1975 *The Population Projection Model of the Central Bureau of Statistics of Norway in the Regional Projections* 1976 48 s. kr 5,00

" 81 Petter Jakob Bjerve: Two Addresses on Statistical Co-operation *To talar om statistisk samarbeid* 1976 20 s. kr 5,00

" 82 Gerd Skoe Lettenstrøm: Fødselstall i ekteskapskohorter 1965-1972 En analyse på grunnlag av registerdata *Fertility of Marriage Cohorts An Analysis Based on Register Data* 1976 64 s. kr 8,00

" 83 Lorents Lorentsen og Tor Skoglund: MSG-3 En modell for analyse av den langsiktige økonomiske utvikling *MSG-3 A Model for Analysis of the Long Term Economic Development* 1976 46 s. kr 7,00

" 84 Petter Jakob Bjerve: Trends in Norwegian Planning 1945-1975 *Utviklingstendensar i norsk planlegging gjennom 30 år* 1976 42 s. kr 7,00

" 85 Idar Møglestue og Arild Jeber: Utdanning og yrke til lærerkandidatene fra 1965 *Education and Occupation of Graduates from Teachers' Training Colleges in 1965* 1976 55 s. kr 8,00

" 86 Jan M. Hoem, Erling Berge og Britta Holmbeck: Four Papers on the Analytic Graduation of Fertility Curves *Fire artikler om analytisk glatting av fruktbarhetskurver* 1976 57 s. kr 8,00

" 87 Vidar Ringstad: On the Estimation of Dynamic Relations from Combined Cross Section Time Series Data *Om estimering av dynamiske relasjonar frå tverrsnitts-tidsrekkedata* 1976 19 s. kr 5,00

Fullstendig oversikt over tidligere nummer av serien Artikler finnes i nr. 80.

Publikasjonen utgis i kommisjon hos
H. Aschehoug & Co., Oslo, og er til salgs hos alle bokhandlere
Pris kr 8,00

Omslag trykt hos Grøndahl & Søn Trykkeri A.s, Oslo

ISBN 82 - 537 - 0646 - 4