

Kopieringsundersøkelse i Den norske kirke 2003-04

En undersøkelse utført av MMI på oppdrag fra
Kirkens Arbeidsgiverorganisasjon og Kopinor

av Steinar Tamsfoss



FORORD

I denne rapporten presenteres resultatene av en undersøkelse av kopieringen i Den norske kirke. Undersøkelsen er gjennomført av Markeds- og Mediatitutttet as på oppdrag fra Kirkens Arbeidsgiverorganisasjon (KA) og Kopinor, og er et ledd i kopieringsavtalen som er inngått mellom de to partene. Datainnsamlingen ble gjennomført i tre puljer fra desember 2003 til april 2004.

Undersøkelsen er ledet av en styringsgruppe med 3 representanter oppnevnt av KA og 3 representanter fra Kopinor.

Prosjektleder for undersøkelsen har vært cand.real. Steinar Tamsfoss, MMI.

MMI takker alle ansatte ved tjenestestedene som har levert datamateriale i undersøkelsen, for deres positive medvirkning. En spesiell takk bør de ca 120 kontaktpersonene som påtok seg ansvaret for å organisere den lokale gjennomføringen av datainnsamlingen, ha.

MMI, Oslo, september 2004.

INNHold

FORORD	2
TABELLREGISTER	4
1. FORMÅLET MED UNDERSØKELSEN	5
2. GJENNOMFØRING AV UNDERSØKELSEN	5
2.1 STRATIFIKASJON, UTVALG OG OPPSLUTNING	5
2.2 OPPLEGG FOR GJENNOMFØRINGEN VED DET ENKELTE TJENESTESTED	8
2.3 KODING AV DET INNSENDETE MATERIALET	8
2.4 BEREGNINGSOPPLEGG	10
3. RESULTATER	13
4. FEILKILDER - FEILMARGINER	15
4.1 SAMPLINGFEIL	15
4.2 SYSTEMATISKE FEIL - SKJEVHET	16
4.3 BORTFALL AV TJENESTESTEDER	16
4.4 MÅLEFEIL - LOKAL OPPSLUTNING OM UNDERSØKELSEN	17
4.5 KODING OG REGISTRERING AV DATA - FEILMULIGHETER OG KONTROLLER	17
5. TABELLER	18
6. ESTIMERINGSOPPLEGG	18
6.1 ESTIMATORER	18
6.2 ESTIMERING AV SAMPLINGFEIL (VARIANS)	21
VEDLEGG	26

TABELLREGISTER

Tekstavsnittene

Tabell A	<i>Organisatoriske enheter</i> i bestanden og bruttoutvalget etter opprinnelig stratum. Antall enheter og tilhørende medlemstall i soknene
Tabell B	Fysiske <i>tenestesteder</i> i bruttoutvalget etter utvalgt organisatorisk enhets opprinnelige stratum og revidert stratum for tjenestestedet.
Tabell C	Nettoutvalget av tjenestesteder fordelt etter revidert stratuminnndeling og pulje for medvirkning. Antall tjenestesteder og antall medlemmer disse representerer.
Liste A	Volumbegrep og telleenhet
Liste B	Kategorier for stofftype for beskyttet materiale
Liste C	Kildekategorier (publikasjonstyper) for beskyttet materiale
Tabell D	Beregning av <i>Totalt årsvolum</i> for hele bestanden.
Tabell E	Hovedresultater fra undersøkelsen: Totalt volum av fotokopier og utskrifter, beskyttet volum og rate beskyttet for de enkelte perioder og hele året. Oppblåste tall for hele bestanden.
Tabell F	Totalt årsvolum og beskyttet årsvolum etter produksjonssted og kildetype.
Tabell G	Hovedresultater fra undersøkelsen med tilhørende 95% konfidensintervaller.

Tabellseksjonen

(OBS: ikke tatt med i nettutgaven – brukes primært av Kopinor til beregning av fordeling av vederlag til rettighetshavere (KA/red. anm.))

Tabell 1	Beskyttet materiale i alt fordelt etter type kilde og type stoff. Prosent
Tabell 2.1	Beskyttet materiale i ulike kategorier for opprinnelse fordelt etter kilde- og stofftype. Norske originalutgivelser. Prosent.
Tabell 2.2	Beskyttet materiale i ulike kategorier for opprinnelse fordelt etter kilde- og stofftype. Bearbeidet /arrangert til norsk. Prosent.
Tabell 2.3	Beskyttet materiale i ulike kategorier for opprinnelse fordelt etter kilde- og stofftype. Oversatt til norsk. Prosent.
Tabell 2.4	Beskyttet materiale i ulike kategorier for opprinnelse fordelt etter kilde- og stofftype. Utenlandsk opphavsperson, norsk utgivelse, ikke bearbeidet/oversatt. Prosent.
(Tabell 2.5)	Beskyttet materiale i ulike kategorier for opprinnelse fordelt etter kilde- og stofftype. Norsk opphavsperson, utenlandsk utgivelse. Prosent. (Denne kategorien finnes ikke i materialet)
Tabell 2.6	Beskyttet materiale i ulike kategorier for opprinnelse fordelt etter kilde- og stofftype. Utenlandsk. Prosent.
Tabell 3	Beskyttet materiale fordelt etter type kilde og kategori for opprinnelse. Prosent.
Tabell 4	Beskyttet materiale fordelt etter type stoff og kategori for opprinnelse. Prosent.
Tabell 5	Beskyttet materiale fordelt etter utgivelsesland og kategori for opprinnelse. Prosent.
Tabell 6	Beskyttet materiale fordelt etter opphavspersonens bostedsland og kategori for opprinnelse. Prosent.

1. FORMÅLET MED UNDERSØKELSEN

Kopieringsundersøkelsen har som formål å

- anslå totalt årsvolum av kopier av opphavsrettslig beskyttet (vederlagspliktig) materiale i Den norske kirke.
- anslå hvordan det opphavsrettslig beskyttede materialet fordeler seg på følgende kategorier:
 - kilde- og stofftype
 - opphavspersonen(e)s bostedsland
 - verkets utgivelsesland
 - eventuell oversettelse eller bearbeidelse til norsk

Under begrepet ”kopi” faller også utskrifter fra datamaskin. Undersøkelsen omfatter derfor både fotokopimaskiner og skrivere.

I flere kommuner har Kirken inngått avtale med kommunen om bruk av deres kopimaskin. Kirkens kopiering på kommunale maskiner inngår i undersøkelsen. Ved kopieringsundersøkelse i kommunesektoren blir Kirkens kopiering holdt utenom.

2. GJENNOMFØRING AV UNDERSØKELSEN

Bestanden undersøkelsen gjelder for, er alle lokale tjenestesteder innen Den norske kirke. Det ideelle register for utvalgstrekkning fra en slik bestand ville ha vært et som listet hvert fysiske tjenestested én og bare én gang. Trekkeregisteret som er benyttet, er et register levert av KA, og omfatter samtlige *organisatoriske enheter* (sokn og fellesråd). Bestandsregisteret teller i alt 1585 organisatoriske enheter. I praksis har nokså mange av disse enhetene kontorfellesskap, slik at det egentlige antallet fysiske *tjenestesteder* (kontorer) er langt mindre. Bortsett fra de enheter som tilfeldigvis er trukket ut for å delta i undersøkelsen, lar ikke de fysiske tjenestedene for bestanden som helhet seg identifisere. Det nøyaktige antallet fysiske tjenestesteder er derfor vanskelig å anslå.

2.1 Stratifikasjon, utvalg og oppslutning

Registerbestanden av organisatoriske enheter ble i utgangspunktet stratumindelt etter følgende kjennetegn:

- Fellesrådskontor i kommuner med ett sokn
- Fellesrådskontor i kommuner med flere sokn
- Sokn med færre enn 2000 medlemmer
- Sokn med 2000- 5000 medlemmer
- Sokn med 5000 - 8000 medlemmer
- Sokn med mer enn 8000 medlemmer

Innen hvert stratum ble et rent lotterisk utvalg¹ av enheter trukket for medvirkning i undersøkelsen. Hvert stratum kan oppfattes som en separat bestand, og stratumutvalgene som uavhengige. En slik

¹ Rent lotterisk utvalg er et utvalg der alle enheter har samme sannsynlighet for å bli trukket ut (jf *Lotto*)

inndeling kan ha gunstige statistiske effekter i form av redusert usikkerhet i resultatene. Bruttoutvalget teller i alt 200 organisatoriske enheter.

Det ble ved planleggingen av undersøkelsen forutsett at noen av de utvalgte enhetene kunne ha kontorfellesskap med andre enheter. De utvalgte enhetene ble således bedt om å oppgi hvilke andre enheter som inngikk i fellesskapet. Undersøkelsen skulle omfatte også disse uansett om disse var valgt ut eller ikke. Det skulle vise seg at kontorfellesskap på tvers av stratuminndelingen forekommer i så stor grad at utvalgsdesignen og beregningsopplegg i etterkant måtte revideres fullstendig. (Mer om dette i avsnitt 2.4 Beregningsopplegg.)

Tabell A viser den opprinnelige stratuminndeling av bestanden og bruttoutvalget av organisatoriske enheter.

Tabell A

Organisatoriske enheter i bestanden og bruttoutvalget etter opprinnelig stratum. Antall enheter og tilhørende medlemstall i soknene (fellesråd i flersoknskommuner er tildelt medlemstall 0 for å unngå dobbelttelling).

Opprinnelig stratum	Bestand		Bruttoutvalg	
	Antall enheter	Antall medlemmer	Antall enheter	Antall medlemmer
Fellesrådsctr – kommuner med ett sokn	148	396 644	42	108 631
Fellesrådsctr – kommuner med flere sokn	287	0	34	0
Sokn med mindre enn 2000 medl	625	524 969	32	26 702
Sokn med 2000 – 5000 medl	279	924 613	28	96 516
Sokn med 5000 – 8000 medl	142	893 835	25	162 501
Sokn med mer enn 8000 medl	103	1 056 444	38	386 031
TOTALT	1 584	3 796 505	199	780 381

Tabell B viser hvordan de *fysiske tjenestestedene* i bruttoutvalget fordeler seg mht både den opprinnelige stratumtilhørigheten for den organisatoriske enheten som ble valgt ut, og den reviderte stratuminndelingen av de fysiske tjenestestedene der den opprinnelig utvalgte enheten inngår. Det gjøres oppmerksom på at i tilfeller der flere organisatoriske enheter fra samme kommune inngår i bruttoutvalget, og har kontorfellesskap, er bare én av disse ”beholdt” som tjenestestedsdefinisjon – og vilkårlig hvilken. Dette medfører tilsynelatende konflikter mellom Tabell A og Tabell B – for eksempel at ”Antall tjenestesteder” under ”I alt” i Tabell B for Opprinnelig stratum ”Fellesrådskontor – kommuner med flere sokn” er mindre enn den tilsvarende størrelsen i Tabell A (hhv 31 og 34). Avviket skyldes altså at vi i 3 tilfeller har ”flyttet” et utvalgt fellesrådskontor til et av de utvalgte soknene i kommunen slik at fellesrådsenheten opptre i det opprinnelige stratumet der tilflyttingssoknet hører hjemme.

Tabell B

Fysiske tjenestesteder i bruttoutvalget etter utvalgt organisatorisk enhets opprinnelige stratum og revidert stratum for tjenestestedet.

Antall tjenestesteder og antall sokn disse omfatter (fellesrådsenheten teller ikke med som sokn).

Opprinnelig stratum	Revidert stratum										I alt	
	Kun fellesråd		Fellesråd og alle sokn		Fellesråd og enkelte sokn		Ett sokn alene		Flere sokn sammen			
	Tj.-steder	Sokn	Tj.-steder	Sokn	Tj.-steder	Sokn	Tj.-steder	Sokn	Tj.-steder	Sokn	Tj.-steder	Sokn
F.rådsctr – komm. med ett sokn	-	-	42	43 *)	-	-	-	-	-	-	42	43
F.rådsctr – komm. med flere sokn	12	0	15	32	4	14	-	-	-	-	31	46
Sokn med mindre enn 2000 medl	-	-	6	20	3	3	13	13	2	6	24	42
Sokn med 2000 – 5000 medl	-	-	3	8	4	4	16	16	1	2	24	30
Sokn med 5000 – 8000 medl	-	-	3	12	3	5	16	16	1	2	23	35
Sokn med mer enn 8000 medl	-	-	4	10	7	9	25	25	-	-	36	44
TOTALT	12	0	73	125	21	35	70	70	4	10	180	240

*) Herav ett sokn fra en annen kommune (felleskontor på tvers av kommunegrense)

Kopieringen innen kirken har enkelte mer intense perioder der det kopieres noe mer og til dels annet stoff enn ellers i året. Spesielt er tiden før jul og påske av en slik art. Utvalget ble derfor inndelt i 3 puljer med ulike registreringsperioder - hver på 2 uker. De respektive puljene representerer imidlertid perioder av ulik varighet i løpet av året. Styringsgruppen for undersøkelsen vurderte varigheten av disse periodene til å være som angitt nedenfor (det er antatt at "kopieringsåret" har en varighet på 47 uker):

	<u>Registreringsperiode</u>	<u>Periode puljen representerer</u>
Pulje 1:	8. - 21. desember 2003	jul – 4 uker
Pulje 2:	1. – 15. februar 2004	året forøvrig – 41 uker
Pulje 3:	29. mars - 11. april 2004	påske – 2 uker

Periodenes varighet tas hensyn til ved beregning av resultater (se avsnitt 2.4).

Enkelte tjenestesteder ble noe forsinket med gjennomføringen i forhold til den oppsatte tidsplanen – spesielt tjenestesteder som skulle delta i pulje 1. Forsinkede tjenestesteder ble automatisk overført til først påfølgende pulje. Tabell C viser hvordan nettoutvalgets tjenestesteder fordeler seg på de ulike puljene og reviderte strata:

Tabell C

Nettutvalget av tjenestesteder fordelt etter revidert stratuminddeling og pulje for medvirkning. Antall tjenestesteder og antall medlemmer disse representerer. (Fellesrådsenheten tilordnes 0 medlemmer).

Reviderte strata	Pulje for medvirkning						Alle puljer	
	Pulje 1 (jul)		Pulje 2 (året for øvrig)		Pulje 3 (påske)		Tj.-steder	Medlemmer
	Tj.-steder	Medlemmer	Tj.-steder	Medlemmer	Tj.-steder	Medlemmer		
Kun fellesråd	3	0	1	0	5	0	9	0
Fellesråd og alle sokn	15	79 186	23	123 049	20	116 882	58	319 117
Fellesråd og enkelte sokn	4	32 786	8	30 885	4	34 061	16	97 732
Ett sokn alene	9	67 548	13	97 639	13	100 267	35	265 454
Flere sokn sammen	1	12 482	1	13 929	1	10 880	3	37 291
Hele utvalget	32	192 002	46	265 502	43	262 090	121	719 594

Vi vil senere i rapporten slå sammen noen av de reviderte strataene ved beregning av totale årsvolum. I sammenslått versjon blir Tabell C som følger:

Reviderte strata – hovedgrupper	Pulje for medvirkning						Alle puljer	
	Pulje 1 (jul)		Pulje 2 (året for øvrig)		Pulje 3 (påske)		Tj.-steder	Medlemmer
	Tj.-steder	Medlemmer	Tj.-steder	Medlemmer	Tj.-steder	Medlemmer		
Kun fellesråd	3	0	1	0	5	0	9	0
Fellesråd og sokn sammen	19	111 972	31	153 934	24	150 943	74	416 849
Kun sokn	10	80 030	14	111 568	14	111 147	38	302 745
Hele utvalget	32	192 002	46	265 502	43	262 090	121	719 594

Pulje 1 ble noe redusert i forhold til det planlagte, mens Pulje 2 ble noe større spesielt fordi undersøkelse i denne puljen kunne gjennomføres innenfor er videre tidsramme enn de to øvrige. Oppslutningen i hver av puljene må imidlertid kunne karakteriseres som tilfredsstillende. Totalt sett har altså 121 av 180 tjenestesteder deltatt i undersøkelsen, dvs en responsrate for tjenestesteder på vel 67 prosent. Ved den tilsvarende undersøkelsen i 2001 var responsraten vel 78%.

2.2 Opplegg for gjennomføringen ved det enkelte tjenestested

Ved hvert av de utvalgte tjenestestedene skulle alle kopier og utskrifter som ble produsert i løpet av undersøkelsesukene registreres. Registreringen skulle foretas av hver enkelt som tok kopier/utskrifter, etter følgende retningslinjer:

- Dersom dokumentet var hentet fra et *utgitt verk*, skulle registreringen skje ved at det ble tatt en *ekstra kopi* som skulle stiftes til et skjema utfyllt med opplysninger om bl.a. opplagstall (antall kopier), forfatterens bostedsland og verkets utgivelsesland.
- Dersom dokumentet var *fortrolig og/eller 100% egenprodusert*, skulle man bare notere volumet (antall kopiesider) av slikt materiale på et særskilt skjema. Fortrolig materiale regnes i denne sammenhengen som ikke beskyttet materiale (ikke utgitt).

Det ble altså benyttet to ulike registreringsskjema, og de samme skjemaene ble brukt for både fotokopier og utskrifter uten markering av hvilken av disse produksjonen gjaldt.

Ved hvert tjenestested som medvirket ble det før undersøkelsen oppnevnt en eller flere kontaktpersoner som skulle legge undersøkelsen til rette lokalt. Kontaktpersonen hadde følgende oppgaver:

- fungere som kontaktledd for MMI
- sørge for utdeling av informasjon om undersøkelsen ved tjenestestedet
- sørge for at nødvendig utstyr og materiell fantes ved kopimaskinene/skriverne gjennom hele undersøkelsesperioden
- sende inn observasjonsmaterialet til MMI

I tillegg til denne registreringen ble det innhentet telleverksoppgaver for samtlige kopimaskiner/skrivere ved de medvirkende tjenestestedene i den hensikt å anslå de totale årsvolum ved tjenestestedene. I praksis er alle fotokopimaskiner og kombinerte kopimaskiner/skrivere utstyrt med avlesbare telleverk, mens rene skrivere som har slikt – ofte vanskelige å avlese uten særskilt instruksjon. De anslåtte årsvolumene vil derfor i noen tilfeller være basert på volumet av materiale som ble levert for toukersperioden. Siden dette ofte kan være mindre enn det faktisk produserte materialet (underrapportering), vil det sannsynligvis foreligge en systematisk underestimering av årsvolumene som skal bestemmes i undersøkelsen. Hvor stor denne kan være, er imidlertid uråd å anslå. (Se også avsnitt 2.4.)

2.3 Koding av det innsendte materialet

Hos MMI ble det innsendte materialet kodet på et særskilt *kodeskjema* i henhold til en fastsatt kodeinstruks. Instruksjonen, som er tatt med som vedlegg sammen med kodeskjema, gir også forklaring på de sentrale begrepene i undersøkelsen. Under kodingen ble det benyttet følgende telleenheter i forbindelse med de ulike volumbegrepene:

Liste A: Volumbegrep og telleenhet

Volumbegrep	Telleenhet
Totalt kopivolum	A 4-kopiesider
Volum opphavsrettslig beskyttet materiale	Originalsider (i det kopierte verkets format)

Vi benytter her begrepet *beskyttet materiale* om materiale som er berørt av Lov om opphavsrett til åndsverk, og som har vært gjenstand for *utgivelse*. Et unntak her er utgitte verk som tjenestestedet selv har opphavsretten til. Intern kopiering av slike er blitt kodet som ”ikke beskyttet” (egentlig ”ikke vederlagspliktig”).

I det etterfølgende vil vi for enkelhets skyld benytte benevnelsen ”kopier” eller ”kopivolum” for både fotokopier og utskrifter – med mindre annet er presisert.

De ulike kategoriene av stofftype og kildetype som er benyttet for klassifisering av det beskyttede materialet, er vist i de følgende lister.

Liste B: Kategorier for stofftype for beskyttet materiale.

Drama/skuespill	Brosjyre- og annonsetekster, tekst fra off. publikasjon o.l.
Dikt, sangtekster	Foto av kunsthåndverk
Skjønnlitterær prosa	Foto av billedkunst
Lære-/fag-/sakboktekst, vitenskapelig tekst	Fagtegning/fagillustrasjon
Lederartikkel, nyhetskommentar, nyhet og reportasje	Annen illustrasjon
Kronikker, artikler o.l.	Fotografi
Anmeldelser av billedkunst, litteratur, film, teater og musikk	Noter
	Tabell/tallmessig framstilling

Liste C: Kildekategorier (publikasjonstyper) for beskyttet materiale

Læreverk	Tidsskrift med allmennkulturelt innhold
Fagbok/sakelitteratur	Fagpresse
Leksikon/oppslagsbok	Uken og månedsblad, magasin, tegneserieblad
Skjønnlitterært originalverk	Andre tidsskrifter og blad
Skjønnlitterært samleverk/antologi	Offentlig publikasjon, utredning o.l.
Noteblad/notesamling/notabok	Annet trykt/forlagsprodusert materiale
Sangbok/visebok	Kirkens publikasjoner *:
Avis	- Noter
Vitenskapelig tidsskrift	- Annet

* Verk utgitt på Verbum forlag eller av Det Norske Bibelselskap i hht kataloger.

I observasjonsmaterialet som er innsendt foreligger enkelte dokumenter som ikke opplagt lar seg klassifisere som ”beskyttet” etter kodeinstruksjonen og kategoriene ovenfor. Bl.a. gjelder dette utskrifter fra Internett eller andre elektroniske lagre. Disse er kodet som ”beskyttet” bare dersom verket også er *utgitt* i papirform. Som kildetype er valgt den korresponderende papirutgave. I tilfeller der det ikke foreligger en papirutgave, er utskriften klassifisert som ”ikke beskyttet” (ikke vederlagspliktig). For slike utskrifter er det imidlertid ikke alltid lett å finne ut om det foreligger noen papirutgave.

Også offentlige publikasjoner kan være vanskelige å kode. Her fremgår det ikke alltid klart av ekstrakopien som er levert, om dokumentet gjelder direkte myndighetsutøvelse/forvaltning (”ikke vederlagspliktig”) eller er av annen karakter. Bl.a. vil stortingsdokumenter, lover, forskrifter og andre retningslinjer og utredninger for offentlig myndighetsutøvelse – bli klassifisert som ”ikke vederlagspliktig”, mens publikasjoner fra for eksempel Statistisk sentralbyrå, FN, EU, Verdensbanken, og lignende kan bli klassifisert som ”fagbøker”. Publikasjoner fra internasjonale organisasjoner har dessuten et tilleggsproblem: Hvilket land er opphavspersonen bosatt i? Med mindre forfatterens hjemland fremgår av kopien, vil publikasjoner fra organer der Norge er medlem, bli definert som ”norske”.

I tvilstilfeller er den generelle regelen at klassifiseringen blir drøftet med prosjektansvarlig eller partene i kopieringsavtalen. Alle varianter er brukt i denne undersøkelsen.

2.4 Beregningsopplegg

Vi vil her gi en verbal beskrivelse av beregningsopplegget som er valgt. En mer detaljert og matematisk beskrivelse er gitt i kapittel 6. Målet for undersøkelsen er, som nevnt, å anslå bl.a. samlet årlig volum av kopier fra beskyttede verk i hele bestanden av tjenestesteder/organisatoriske enheter. Det skal med andre ord generaliseres fra et utvalg av tjenestesteder til hele bestanden, og fra en registreringsperiode på to uker til et helt år. Den tilfeldige trekkingen av tjenestesteder gir metodisk grunnlag for å generalisere til hele bestanden, og samtidig anslå feilmarginer for den usikkerheten som skyldes at bare et tilfeldig utvalg - og ikke hele bestanden - er undersøkt (samplingfeil - se vedlegg om estimering). Generaliseringen i tid er gjort på basis av puljeinndelingen og varigheten av de perioder hver pulje representerer.

Undersøkelsen gir estimater for:

- årlig totalt kopivolum (fotokopier og utskrifter)
- årlig volum av beskyttet materiale
- raten for beskyttet materiale i forhold til totalt kopivolum
- prosentvis fordeling av det beskyttede materialet etter opprinnelse og etter de ulike stoff- og kildekategoriene

For å bestemme det beskyttede årsvolumet ved *det enkelte tjenestested* er følgende prinsipp fulgt:

$$\text{Beskyttet årsvolum} = [\text{Observert beskyttet volum} / \text{Observert totalvolum}] \times \text{Totalt årsvolum}$$

Alle størrelser på høyre side må estimeres på grunnlag av data innhentet i undersøkelsen. "Observert beskyttet volum" og "Observert totalvolum" bestemmes på grunnlag av de innrapporterte kopier/utskrifter, mens *Totalt årsvolum* må estimeres særskilt. Ettersom dette volumet har en avgjørende innflytelse på resultatene, skal vi gå mer detaljert inn på bestemmelsen av dette i det følgende.

Totalt årsvolum

For kopimaskiner og skrivere med telleravlesninger kan årsvolumet bestemmes ut fra disse. Undersøkelsen omfatter imidlertid en del maskiner – spesielt skrivere – som ikke er telleravlest (en del skrivere har ikke telleverk, mens noen med telleverk kan være vanskelige å avlese). Årsvolum fra telleravleste maskiner alene vil derfor undervurdere det totale volumet.

I de tilfellene årsvolum eller telleravlesninger ikke har vært oppgitt, har vi for hvert tjenestested foretatt en sammenligning mellom telleravlest volum i registreringsperioden (2 uker) og volumet av innrapporterte ekstrakopier. Dersom sistnevnte volum har vært større enn det telleravleste, er uoppgitte avlesninger erstattet med differensen mellom de to størrelsene (noe som kan være en undervurdering av det faktiske volumet). Dersom det innrapporterte volumet har vært mindre enn det telleravleste (altså stor grad av underrapportering), har vi som erstatning for uoppgitte årsvolum satt inn årsvolumet for maskiner av tilsvarende type ved tjenestesteder som er sammenlignbare (type tjenestested og antallet medlemmer i soknet/kommunen). De estimerte årsvolumene for de ulike tjenestestedene er derfor noe skjønnsmessig bestemt, men etter hva vi vil anse som grundige vurderinger - og heller med forsiktige "gjetninger" enn det motsatte.

Som nevnt i et tidligere avsnitt, medfører den hyppige forekomsten av felleskontorer mellom organisatoriske enheter på tvers av de opprinnelige strata at denne stratuminndelingen ikke kan legges til grunn ved resultatberegningene. Heller ikke de reviderte strataene er brukbare til dette formålet

ettersom denne inndelingen kun gjelder utvalget og ikke bestanden – begge deler er en forutsetning for at inndelingen skal kunne benyttes. Det har derfor vært nødvendig å finne alternative metoder for estimeringen.

Analyser av kopivolumet (årsvolum) innen Kirken foretatt i forbindelse med undersøkelsen i 2001, viste en meget sterk sammenheng kommunevis mellom årlig kopivolum og antallet medlemmer av Kirken. Det er neppe grunn til å tro at dette forholdet har endret seg vesentlig på 3-4 år, og vil derfor kunne utnyttes for å kunne bestemme det totale årsvolum innen Kirken. Et problem i denne sammenhengen er imidlertid at fellesråd ikke har egne medlemmer. For fellesråd er det eneste medlemstallet som er aktuelt, summen av medlemmer i kommunens sokn. I det følgende vil vi derfor splitte beregningen av *Totalt årsvolum* i to komponenter:

$$\text{Totalt årsvolum} = \text{Fellesrådsenhetenes årsvolum} + \text{soknenhetenes årsvolum}$$

Splittingen er aktuell å utføre bare for de reviderte strata der tjenestestedene består av både fellesråd og sokn, og er gjort på følgende måte:

For tjenestestedene som kun omfatter fellesrådsenheter har vi beregnet det gjennomsnittlige årsvolumet pr medlem i de tilhørende kommuner til 1,846. For de øvrige tjenestesteder der fellesråd inngår, er fellesrådsvolumet anslått ved å multiplisere dette gjennomsnittet med antallet medlemmer i disse kommunene. Sokneenhetenes årsvolum ved hvert av disse tjenestestedene er så beregnet som differensen mellom det observerte årsvolumet og det anslåtte fellesrådsvolum. I noen få tilfeller viser det seg at denne differensen blir negativ. I disse tilfellene har vi satt fellesrådets del til 0, og latt hele årsvolumet falle på soknene (soknene har vanligvis større volum enn fellesrådet).

Tabell D viser de aktuelle størrelsene for tjenestestedene og deres kommuner i nettoutvalget. Beregningene som er angitt i tabellhodet, gjelder kun de to kategoriene der både fellesråd og sokn inngår. De beregnede størrelsene er satt i kursiv. Den grunnleggende ide bak disse beregningene er at fellesrådet betjener hele kommunen, mens hvert sokn betjener kun sine egne medlemmer.

Tabell D
Beregning av *Totalt årsvolum* for hele bestanden.

Type tjenestested i utvalget	Observerte data for utvalget			Beregnete data forutsatt at fellesråd produserer 1,85 sider pr år pr medlem i kommunen *)		
	Årsvolum (A)	Medlemmer i observerte sokn (B)	Medlemmer i tilhørende kommuner (C)	Fellesrådets del av årsvolumet (D=1,846xC)	Soknenes del av årsvolumet (A-D)	Soknenes årsvolum pr medlem i sokn ((A-D)/B)
Kun fellesråd	178 630	0	96 766	178 630	0	0
Fellesråd og sokn sammen	2 641 457	416 849	642 225	<i>1 185 547</i>	<i>1 455 910</i>	3,49
Kun sokn	2 294 185	302 745	1 327 010	0	2 294 185	7,58
Totalt	-	719 594	-	-	3 750 095	5,211

*) Tallet 1,846 er fremkommet ved å beregne det gjennomsnittlige årsvolum pr medlem i fellesrådenes kommuner i utvalget (A/C for kategorien *Kun fellesråd* i tabellen).

Ut fra dette kan vi etablere følgende regnestykke som gir oss det totale årsvolumet for bestanden som helhet:

	Gjennomsnittlig årsvolum pr medlem	Totalt medlemstall i Kirken	Bidrag til totalt årsvolum (avrundet)
Fellesrådsenheter	1,846	3 796 505	7 008 000
Sokn	5,211	3 796 505	19 784 000
Totalt årsvolum			26 792 000

Sammenvekting av puljene

Etter at det totale årsvolumet er bestemt, gjenstår to hovedoperasjoner i beregningene:

- å gjøre observasjonene fra hver av de tre puljene "representative" for hele bestanden
- å sette de tre puljene sammen i henhold til lengden av de perioder de representerer

Begge operasjonene utføres ved passende vekting av observasjonene. Den første gjøres ved å blåse opp observasjonene fra hver av puljene slik at det totale årsvolumet estimert for hver av dem blir det samme for alle puljer, nemlig lik totalvolumet vi har beregnet ovenfor. Vi betrakter med andre ord observasjonene fra hver av puljene som separate undersøkelser og "later som" om den enkelte pulje var de eneste data vi hadde. Denne vekten vil være lik:

$$\text{Volumvekt pulje } p = \text{Totalt årsvolum for bestanden} / \text{Sum årsvolum tjenesteder i utvalget pulje } p$$

Med denne vekten gir resultatene fra hver av puljene et bilde av kopieringen over hele året *dersom* årsproduksjonen hadde vært som i den perioden vedkommende pulje representerer. Dette er naturligvis ikke tilfelle. Puljeresultatene må derfor krympes slik at de gir uttrykk for produksjonen bare i de tilhørende perioder. Som nevnt foran er det stipulert at produksjonsåret har totalt 47 uker. De tre puljene representerer perioder med varighet hhv 4 uker (jul), 2 uker (påske) og 41 uker (resten av året). De tilhørende periodevekter blir dermed hhv 4/47 (pulje 1), 41/47 (pulje 2) og 2/47 (pulje 3), og slik at summen av dem blir eksakt 1. De endelige vektene oppnås ved å multiplisere volumvektene med de tilhørende periodevekter.

Observasjonsmaterialet lar seg ikke splitte på fellesråd og sokn slik som vi har gjort for årsvolumene. En slik operasjon ville kreve at vi ved blandede tjeneststeder (fellesråd og sokn) måtte kunne identifisere hvilke kopier som er produsert av hvilken type enhet, noe som ikke lar seg gjøre i det valgte undersøkelsesopplegget. Alle observasjoner innen samme pulje gis derfor samme vekt i beregningene.

3. RESULTATER

Hovedresultatene fra undersøkelsen er gjengitt i Tabell E nedenfor. Det bemerkes at *Rate beskyttet* er beregnet som forholdstallet mellom volumet av beskyttede *originalsider* og det totale volum av *kopisider*. Volum av ikke opphavsrettslig beskyttet materiale angis i A4-kopisider, mens opphavsrettslig beskyttet materiale derimot telles både i kopisider (når det er tale om "kopivolum") og i originalsider ("beskyttet volum"). Med en originalside menes sideformatet på den opprinnelige kilden. Det forekommer at flere originalsider blir kopiert på samme kopiside, slik at det kan bli flere originalsider enn kopisider. For aviser er situasjonen spesiell ettersom en originalside utgjør mer enn en A4-side. Her er en originalside regnet som en A4-side, slik at artikler som vi, uten forminsking eller forstørring, må bruke f.eks. to A4-sider for å kopiere (tabloidformat), regnes som to originalsider. Salmer og sanger som kopieres opp for utdeling, foreligger ofte i avskrifts form. For slike regnes vanligvis hver sang som én originalside med mindre den beviselig overstiger dette. Dette innebærer at prosentandelen av beskyttet materiale (beskyttetraten) i teorien kan overstige 100%.

Tabell E

Hovedresultater fra undersøkelsen: Totalt volum av fotokopier og utskrifter, beskyttet volum og rate beskyttet for de enkelte perioder og hele året. Oppblåste tall for hele bestanden.

Periode	Totalvolum (maskinsider)	Beskyttet volum (originalsider)	Herav transparenter	Rate beskyttet (%)
Jul (4 uker)	2 280 170	2 003 834	1 537	87,9
Påske (2 uker)	1 140 085	734 053	57	64,4
Året for øvrig (41 uker)	23 371 745	10 028 724	6 600	42,9
Hele året 2004	26 792 000	12 766 612	8 194	47,7
Hele året 2001	28 425 076	6 813 930	(Ikke målt)	24,0

For hele bestanden under ett er det årlige kopivolumet anslått til nær 27 mill maskinsider, som er litt mindre enn volumet anslått ved undersøkelsen i 2001. Volumet av beskyttede verk er anslått til i underkant av 13 mill originalsidekopier, som svarer til en samlet rate av beskyttet materiale på nær 48% (48 beskyttede originalsider pr 100 maskinsider). Dette er nesten det dobbelte av hva som ble målt i 2001. Kopiering av beskyttet materiale er spesielt stor – relativt sett - ved jule- og påsketider.

Kirken benytter flere steder kommunale maskiner for kopiering. I noen tilfeller er det tale om en fast leieordning, mens det i andre tilfeller skjer ved større kopieringsoppdrag. Kopiering på kommunale maskiner er inkludert i resultatene ovenfor. Ut fra det innsamlede materialet (der innsendte kopier ble merket med "kommunal maskin"), er årsvolumet av denne kopieringen anslått til totalt ca 417 500 maskinsider, mens antallet beskyttede originalsider utgjør ca 692 000. Her er det altså tale om produksjon av bl.a. flere salmetekster på samme side (det aller meste klassifisert under kildetypen "Kirkelige publikasjoner - annet materiale").

Av det beskyttede materialet totalt stammer to tredeler (66%) fra Kirkelige publikasjoner (forlagene Verbum og Det norske bibelselskap). Dette utgjør til sammen ca 8,4 mill originalsider på årsbasis. Av disse er 171 000 originalsider klassifisert som "noter" (kildetype "Noteblad, notesamling, notebok").

Tabell F viser volumfordelingen etter både type kilde og produksjonssted.

Tabell F**Totalt årsvolum og beskyttet årsvolum etter produksjonssted og kildetype.**

Type kilde	I ALT		Produksjonssted			
	Årlig totalvolum (mask.sider)	Årlig beskyttet volum (orig.sider)	Internt		Eksternt	
			Årlig totalvolum (mask.sider)	Årlig beskyttet volum (orig.sider)	Årlig totalvolum (mask.sider)	Årlig beskyttet volum (orig.sider)
Læreverk	192 493	221 337	192 493	221 337	-	-
Fagbok	981 516	1 016 033	981 516	1 016 033	-	-
Leksikon	273 537	383 801	273 537	383 801	-	-
Skjønnlitterært originalverk	1 671	2 089	1 671	2 089	-	-
Noteblad/-bok	335 581	340 218	335 581	340 218	-	-
Sang-/visebok	765 792	892 928	765 792	892 928	-	-
Avis	28 463	52 092	28 463	52 092	-	-
Fagpresse	125 024	131 857	125 024	131 857	-	-
Ukeblad/tegneserie	29 919	579 029	29 919	579 029	-	-
Off. publikasjon/utredning	15 408	880	15 408	880	-	-
Annet forlagsprod. materiale	463 291	735 467	442 998	694 882	20 293	40 586
Kirkens publ.: Noter	113 800	171 031	113 800	171 031	-	-
Kirkens publ.: Annet	4 978 325	8 239 849	4 652 930	7 588 284	325 395	651 565
Ikke forlagsprodusert/utgitt	18 487 180	0	18 415 363	0	71 816	0
Alle kilder	26 792 000	12 766 612	26 374 496	12 074 460	417 504	692 151

Tabelldelen av rapporten viser ulike prosentfordelinger av det beskyttede materialet for hele bestanden under ett, og er basert på oppblåste volumdata for hele året.

4. FEILKILDER - FEILMARGINER

Alle resultater som presenteres i denne rapporten, er beheftet med usikkerhet eller feil. Feilkildene kan grupperes i fire kategorier:

- Samplingfeil som skyldes at vi har observert bare et utvalg av tjenestesteder - og ikke hele bestanden ("romlige" feil).
- Feil på grunn av bortfall av utvalgte tjenestesteder. Fører dels til at datagrunnlaget får noe mindre omfang enn planlagt, og dels til mulige skjevheter dersom bortfalte enheter har et kopieringsmønster som avviker fra de medvirkendes.
- Feil som skyldes at kopieringen ved hvert tjenestested er observert i kun to uker av året som resultatene gjelder for. Kopieringen i den valgte observasjonsperioden kan avvike i en bestemt retning fra den årstypiske ("tidsvariasjon"). Inndelingen av utvalget i 3 puljer med registreringsperioder som hver representerer hhv jul, påske og resten av året, fanger imidlertid opp de største sesongvariasjonene. Sammenvektingen av periodene skal således gi et bilde som i prinsippet er representativt for hele årsproduksjonen.
- Målefeil som kan oppstå ved kopiregistreringen - f.eks. feil i opplagstall, feilklassifisering eller manglende registrering av kopier og utskrifter som blir tatt i observasjonsperioden, eller mangler/feil ved oppgavene/anslagene for totale årsvolum.

Av de fire feiltypene er det bare den første som lar seg *tallfeste* i denne undersøkelsen, noe vi skal se nærmere på i neste avsnitt. Vi har foran konstatert at det er et visst, men ikke spesielt stort, bortfall av tjenestesteder i alle puljer, noe som *kan* gi rom for skjevheter dersom kopieringsmønsteret blant de bortfalte er vesentlig forskjellig fra de medvirkendes. Dette har vi imidlertid for lite informasjon til å kunne vurdere.

4.1 Samplingfeil

Samplingfeilen til de estimerte resultater er et statistisk mål for hvor mye disse estimatene vil kunne variere dersom undersøkelsen ble gjentatt (samtidig) etter det samme opplegget *uendelig* mange ganger - altså ved at vi trakk stadig nye utvalg (med samme størrelse og etter samme design) av tjenestesteder som ble observert. Formlene for å anslå disse feilene er gjengitt i vedlegget. "Feilmarginene" som er angitt i Tabell G nedenfor, er såkalte 95% *konfidensintervaller* (estimert resultat $\pm 2 \times$ samplingfeilen), og er å fortolke som 95% *sannsynlige områder* for de "sanne" verdiene.

Denne sannsynligheten - *konfidenskoeffisienten* (som her er valgt lik 95%) - kan fastsettes vilkårlig. Velges lavere konfidenskoeffisient, får vi kortere intervaller. Dette betyr imidlertid ikke at resultatet blir sikrere, men at det kortere intervallet er et mindre sannsynlig område for lokaliseringen av den ukjente størrelsen. Utrekningen på konfidensintervallet forteller hvor presist vårt estimat er når konfidenskoeffisienten holdes fast.

Tabell G viser slike konfidensintervaller for hovedresultatene i undersøkelsen. Vi ser bl.a. at det er heftet forholdsvis stor usikkerhet til resultatene for de enkelte strata, mens feilmarginen for alle strata under ett er noe mindre – relativt sett. Det er spesielt ett stratum som bidrar sterkt til den samlede usikkerheten, nemlig det som opptrer nederst i tabellen ("Annet organ, internt fotokopivolum 1999 uoppgitt"). I dette stratomet mangler vi altså volumdata for 1999, og har dermed ikke kunnet utnytte

denne informasjonen i estimeringen – noe som har medført at estimatet er vesentlig mer usikkert enn for de øvrige strata.

Tabell G
Hovedresultater fra undersøkelsen med tilhørende 95% konfidensintervaller.

	Estimert årsvolum	Nedre konfidensgrense	Øvre konfidensgrense
Totalt årsvolum	26 792 000	22 035 000	31 549 000
Beskyttet årsvolum	12 766 612	9 489 000	16 044 000

Det understrekes at feilmarginene i tabellen kun avspeiler *samplingfeil* – dvs usikkerhet som følger av at bare et utvalg av, og ikke hele tjenestestedsbestanden – er observert. Vi har i et foregående avsnitt anført at oppgavene over totale årsvolum, som influerer sterkt på sluttresultatene, kan være noe usikre for enkelte av de medvirkende tjenestestedene. Denne usikkerhetskomponenten er imidlertid ikke tatt hensyn til ved beregningen av feilmarginer. Heller ikke er ukjente effekter av bortfall inkorporert. Et stort bortfall vil alltid øke samplingfeilen (i forhold til det planlagte), og kan i visse tilfeller, påføre resultatene systematiske skjevheter (se avsnitt 4.3).

Når det gjelder fordelingstallene (tabellseksjonen), som vi ikke har beregnet feilmarginer for, kan observasjonsperiodens lengde ha en viss betydning for hvor stor tillit man kan feste til enkeltestimatene. Det viktigste momentet her er at antallet observasjoner av beskyttet materiale i en kort periode (to uker) blir lite – spesielt når dette materialet bare utgjør en liten andel av det totale volumet. Med få observasjoner vil den enkelte observasjon kunne slå kraftigere ut i resultatene enn om man har et større observasjonsmateriale. Effekten av slike tilfeldige utslag (få observasjoner) fanges i prinsippet opp i samplingfeilen.

4.2 Systematiske feil - skjevhet

Det sier seg selv av når datainnsamlingen pågår over en periode på bare to uker, og resultatene skal gi uttrykk for årsvolum, så kan valg av observasjonsperiode ha betydning for utfallet. Vi har i denne undersøkelsen inndelt året i tre ulike perioder (jul, påske og året for øvrig). Innen hver periode er ca en tredel av utvalget observert. Kopieringsaktiviteten i de ulike puljer må antas å være representative for den totale kopieringen i tilhørende periode. I beregningene er det tatt hensyn til periodenes lengde slik at skjevheter som følge av ulikheter mellom puljene i prinsippet skal være motvirket i årsvolumtallene.

4.3 Bortfall av tjenestesteder

Når utvalgte tjenestesteder av ulike grunner ikke leverer materiale i det hele tatt, taler vi om bortfall av enheter. Problemet med slikt bortfall er - i tillegg til at det reduserer omfanget av det planlagte datagrunnlaget - at man ikke vet hvor "tilfeldig" det er. Dersom bortfalte tjenestesteder representerer et betydelig årlig kopivolum og har et kopieringsmønster som avviker systematisk fra de medvirkende tjenestestedene, vil resultatene naturligvis kunne bli skjeve.

I denne undersøkelsen utgjør bortfallet av tjenestesteder ca 33% av bruttoutvalget (se avsnitt 2.1), noe som er forholdsvis mye i denne typen undersøkelser. Vi har imidlertid ikke noen data som kan gi grunnlag for å vurdere om de bortfalte tjenestestedene avviker vesentlig kopieringsmessig fra de medvirkende.

4.4 Målefeil - lokal oppslutning om undersøkelsen

Resultatene av kopieringsundersøkelsen er basert på tjenestestedenes innrapporterte kopieringer i undersøkelsesperioden. Kopiregistreringen har blitt gjort enten ved innsending av ekstrakopier eller ved at kopieringen er notert på et særskilt rapportskjema for fortrolig og egenprodusert materiale. I de tilfellene hvor det ble tatt en ekstrakopi, skulle kopitakeren også fylle ut en blankett med informasjon om antall kopiesider og opplag for hvert kopisett. Hvis kopisettet inneholdt opphavsrettslig beskyttet materiale, skulle de også oppgi hvilket verk dette var hentet fra. Denne prosedyren kan føre til:

- Underrapportering, fordi kopitaker unnlater å rapportere all kopiering
- Feilrapportering, fordi kopitaker oppgir feil opplagstall, eller at opphavsrettslig beskyttet materiale føres opp på rapportskjema for fortrolig og egenprodusert materiale.

En oversikt over rapporteringsgraden ved de medvirkende tjenestestedene (innrapportert volum i forhold til telleravlest volum – ikke vist i denne rapporten) gir et forholdsvis positivt bilde av oppfølgingen ved det enkelte sted. Rapporteringsgraden er tilfredsstillende stor ved de aller fleste stedene. Her må vi imidlertid skyte inn at mange maskiner (særlig skrivere) ikke er blitt telleravlest, og at tellervolumet for disse i registreringsperioden til dels er blitt anslått på basis av de innrapporterte kopier – og da nokså ”nøkternt” (dvs at det er forutsatt at tjenestestedet har rapportert så å si all produksjon i perioden).

Feil opplysninger om kilde og sidetall kan kontrolleres ved sammenligning av skjema for ekstrakopi og selve ekstrakopien der slik foreligger. Feil bruk av rapportskjema er det derimot umulig å oppdage. Feil i opplagstall for enkeltkopieringer kan ikke kontrolleres enkeltvis. Slike feil kan bare evalueres og eventuelt korrigeres dersom det viser seg at det samlede innrapporterte volumet avviker sterkt (spesielt “oppover”) fra telleverksavlesningene. I slike tilfeller er normal prosedyre at alt materiale gjennomgås på ny, og kopitaker kontaktes dersom det foreligger “mistenkkelige” enkeltkopieringer. I denne undersøkelsen har sistnevnte ikke vært nødvendig. Denne typen av kontroll kan imidlertid bare utføres for tjeneststeder som har telleravlest samtlige maskiner.

Hvis materialet som feil- eller underrapporteres ikke er systematisk forskjellig i sammensetning fra det øvrige materialet, vil den eneste effekten være et redusert datagrunnlag. Ellers kan systematiske feil- eller underrapporteringer føre til at andelen av opphavsrettslig beskyttet materiale i forhold til totalkopieringen blir feil, og/eller at enkelte kategorier av opphavsrettslig beskyttet materiale over- eller underestimeres.

4.5 Koding og registrering av data - feilmuligheter og kontroller

Når det innsendte materialet skal kodes og registreres kan det oppstå tekniske problemer, som f.eks. at koder plasseres i feil posisjon. I andre tilfeller kan det være tvil om hva som er riktig kode, f.eks. hvilket verk (type kilde) kopien er hentet fra, og skyldes i første rekke at kopitakerne bare unntaksvis fyller inn opplysninger om dette på blanketten som følger ekstrakopien. Det er gjennomført en rekke manuelle og maskinelle konsistens- og rimelighetskontroller av de registrerte kodingene for å redusere forekomsten av kodefeil så langt det er mulig.

5. TABELLER

(OBS: ikke tatt med i nettutgaven – brukes primært av Kopinor til beregning av fordeling av vederlag til rettighetshavere (KA/red. anm.))

6. ESTIMERINGSOPPLEGG

I dette avsnittet er alle formlene (*estimatorene*) som er brukt i beregningen av resultatene, dokumentert.

Ved estimering i utvalgsundersøkelser må (bør) det alltid tas hensyn til designen som er benyttet ved trekking av utvalget (*samplingdesignen*), samt øvrig informasjon som kan utnyttes for om mulig å forbedre presisjonen i resultatene. De viktigste elementene i en samplingdesign er stratuminndelingen av og utvelgingssannsynlighetene for de enkelte enheter. Som vi har sett i forangående avsnitt, har imidlertid den opprinnelige stratuminndelingen brutt sammen som følge av samlokalisering av og felles rapportering fra enheter i ulike strata. Det har derfor vært nødvendig å revidere opplegget for beregningene. De informasjonsbærende enhetene er altså ikke lenger de opprinnelig utvalgte *organisatoriske* enhetene, men de *tjenestestedene* ett eller flere av dem i praksis går inn under. Dette fører til en betydelig komplikasjon designmessig sett. Vi har derfor søkt å forenkle det hele ved å ta hensyn til det faktum (kjent fra foregående undersøkelse i Kirken) at det totale årsvolumet av kopier og utskrifter i en kommune er så godt som proporsjonalt med medlemstallet i Kirken i vedkommende kommune. Beregningene kan derfor baseres mer på en *modell* for estimeringen enn på selve *samplingdesignen*. Vi kan imidlertid ikke se helt bort fra designen, slik at opplegget må bli en kombinasjon av de to teoretiske prinsippene (modell vs design).

Fremstillingen i det følgende har først og fremst som formål å dokumentere prosedyrene som er anvendt for å utarbeide resultater. Teksten må derfor nødvendigvis preges av matematiske symboler og formler.

6.1 Estimatorer

Størrelsene som skal estimeres i undersøkelsen er følgende - der vi lar X stå for både det årlige volumet av en hvilken som helst kategori av beskyttet materiale, og som spesifisert benevnelse på selve kategorien:

X = Årsvolum for beskyttet materiale av kategori X

B = Årlig beskyttet volum totalt

K = Totalt årsvolum av fotokopier og utskrifter

I beregningene vil vi dessuten ha bruk for

M = Antall medlemmer av Kirken

Størrelser uten fotskrifter vil generelt gjelde for hele bestanden, mens fotskrifter på disse størrelsene vil angi hvilken kommune (i) eller pulje (p) størrelsen gjelder for. I estimeringen av de ovennevnte bestandsstørrelser vil vi således trenge bl.a. følgende notasjoner:

X_i = X -årsvolum i kommune i

K_i = Totalt årsvolum i kommune i

M_i = Totalt antall medlemmer i kommune i

Oppgaver over totalt årsvolum (K) og X -volum foreligger kun for tjenestesteder i utvalget, mens oppgaver over medlemstall i Kirken foreligger for samtlige enheter i både utvalget bestanden. På grunn av den sterke positive samvariasjonen mellom kommunenes totale årsvolum og medlemstallet i Kirken, vil en såkalt *rateestimator* med fordel kunne benyttes som estimator for det totale årsvolumet i hele bestanden. Rateestimatoren kombinerer data for kopivolum og medlemstall.

Totalt årsvolum

Det årlige totalvolumet i kommune i kan splittes i to komponenter – en for fellestrådets kopiering og en for soknens kopiering:

$$K_i = K_i^f + K_i^s$$

der toppskriftene indikerer hvilken komponent volumet gjelder for (f for fellestråd og s for sokn).

Siden fellestrådene ikke har egne medlemmer, må disse behandles særskilt i vår modell. Vi har således tatt utgangspunkt i de av de observerte tjenestestedene som kun er fellestrådskontor, og relatert deres årsvolum med medlemstallet i hele hjemkommunen for fellestrådet. Vi oppnår dermed et estimat for fellestrådets *årsvolum pr medlem* (for alle fellestråd):

$$\hat{K}^f = \frac{K_*^f}{M_*^f}$$

der K_*^f er summen av totale årsvolum ved alle de observerte tjenestestedene som kun omfatter fellestråd, og M_*^f er det summerte medlemstallet i de tilhørende *kommuner*.

For kommune i blir således fellestrådets årsvolum estimert ved å blåse opp gjennomsnittet pr medlem med kommunens medlemstall:

$$\hat{K}_i^f = M_i \hat{K}^f$$

Blåses gjennomsnittet opp med det totale antallet medlemmer i Kirken, får vi et estimat for det totale fellestrådsvolumet i alle kommuner:

$$\hat{K}^f = M \hat{K}^f$$

Estimering av soknekomponenten i de reviderte strata der tjenestestedene omfatter både fellestråd og sokn, er gjort i to trinn: Først anslås fellestrådets del av det samlede årsvolumet, og deretter beregnes soknens årsvolum som differensen mellom det observerte totalvolumet og det anslåtte fellestrådsvolumet:

$$\hat{K}_{ij}^s = K_{ij} - \hat{K}_i^f$$

som altså gjelder for de tjenestesteder (i,j) i kommuner der fellestråd og et eller flere sokn er samlokalisert. Vi benytter her fotskriften (i,j) for å indikere at dette gjelder tjenestested j i kommune i . I de tilfeller der tjenestestedene kun omfatter sokn, kan samme relasjon likevel benyttes idet fellestrådskomponenten jo vil være 0.

Soknekomponenten kan da estimeres på tilsvarende måte som fellesrådskomponenten, men med den forskjell at medlemstallene nå gjelder kun de sokn som inngår i de observerte tjenestestedene:

$$\hat{K}^s = M \frac{\sum_{i,j} \hat{K}_{ij}^s}{\sum_{i,j} M_{ij}}$$

Estimatet for totalt årlig kopivolum i hele bestanden fås ved å summere de to volumkomponentene:

$$\hat{K} = \hat{K}^f + \hat{K}^s$$

Beskyttet årsvolum

Den neste størrelsen som skal estimeres er årlig volum av en bestemt kategori av beskyttede kopier, X , samt totalt beskyttet volum, B . I estimeringen vil vi i tillegg til notasjonene ovenfor også bruke:

x_{ij}^p = observert X -volum ved tjenestested (i,j) i undersøkelsesperioden pulje p

k_{ij}^p = observert kopivolum ved tjenestested (i,j) i undersøkelsesperioden pulje p

Disse (aggregerte) størrelsene kan betraktes som våre enkeltobservasjoner idet alle enkeltkopieringer fra samme tjenestested (i,j) gis den samme beregningstekniske behandlingen.

I prinsippet foretas beregningene av årsvolum for hele bestanden separat for hver av puljene. Dette skjer ved å blåse opp alle enkeltobservasjoner fra puljen med følgende faktor:

$$w_p = \frac{\hat{K}}{\sum_{i,j} k_{ij}^p}$$

Benyttes denne vekten på for eksempel de observerte kopivolumene, får vi det foran anslåtte totale årsvolumet for hele bestanden:

$$\sum_{i,j} w_p k_{ij}^p = \sum_{i,j} \frac{\hat{K}}{\sum_{i,j} k_{ij}^p} k_{ij}^p = \hat{K}$$

Denne estimatoren er altså konsistent i betydningen at den gir det logisk "korrekte" resultat. Vi ser dessuten at vi ved estimering av det totale årsvolumet får samme resultat uansett hvilken pulje vi legger inn data fra.

X -årsvolumet estimeres på tilsvarende måte:

$$\hat{X}^p = \sum_{i,j} w_p x_{ij}^p$$

Setter vi her inn uttrykket for w , ser vi at estimatoren også kan skrives slik:

$$\hat{X}^p = \sum_{i,j} \frac{\hat{K}}{\sum_{i,j} k_{ij}^p} x_{ij}^p = \hat{K} \frac{\sum_{i,j} x_{ij}^p}{\sum_{i,j} k_{ij}^p} = \hat{K} \hat{P}^p$$

der vi har innført størrelsen \hat{P}^p for brøken som inneholder summerasjoner. Denne størrelsen er den observerte beskyttede rate i pulje p .

Dette estimatet gjelder for hele året, men vil avhenge av hvilken pulje vi håndterer. Vi må derfor veie sammen puljeresultatene med passende vekt. Hver pulje representerer bare en avgrenset del av kopieringsåret. Regnes hele året å bestå av 47 produksjonsuker, er periodeinndelingen og de tilhørende periodevektene slik:

	Periodelengde	Periodevekt (α_p)
Pulje 1	4 uker (jul)	4/47
Pulje 3	2 uker (påske)	2/47
Pulje 2	41 uker (resten av året)	41/47

Det sluttelige estimatet blir dermed:

$$\hat{X} = \sum_p \alpha_p \hat{X}^p$$

Siden X også kan stå for det totale beskyttede volumet (B), estimeres B ved å sette inn dette for X .

I fordelingstabellene i tabellseksjonen er alle resultater angitt i prosentvise fordelinger med det totale beskyttede volumet B som prosentueringsbasis. For å estimere prosentvis fordeling av det vederlagspliktige materialet på ulike kategorier, $R(X)$, kan alle X -estimer divideres med estimatet av B .

$$\hat{R}(X) = \frac{\hat{X}}{\hat{B}} \quad (\text{som multiplisert med 100\% gir prosenttall})$$

Raten av beskyttede kopier estimeres ved:

$$\hat{P} = \frac{\hat{B}}{\hat{K}} \quad \text{for bestanden som helhet.}$$

Multipliseres disse med 100 får vi beskyttede kopier målt i prosentenheter, eller antall beskyttede originalsider pr 100 kopiesider - et antall som i teorien kan overstige 100.

6.2 Estimering av samplingfeil (varians)

Som nevnt tidligere er samplingfeilen (variansen) til de estimerte resultatene et statistisk mål for hvor mye disse resultatene vil kunne variere dersom undersøkelsen ble gjentatt etter det samme opplegget

uendelig mange ganger (samtidig), men med de beregningsmessige modelltilpasninger vi har måttet foreta. Formlene vi har brukt for å beregne disse variansene gis her uten utledning.

De første størrelsene vi estimerte var totalt årsvolum for hele bestanden. Estimatoren for dette kan skrives slik:

$$\hat{K} = M\hat{K}^f + M\hat{K}^s = M \frac{\sum_i K_i^f}{\sum_i M_i} + M \frac{\sum_{i,j} K_{ij}^s}{\sum_{i,j} M_{ij}}$$

Variansen til denne størrelsen kan estimeres ved å estimere variansene til hvert av leddene i summen på høyre side (som kan anses som stokastisk uavhengige). Hvert av leddene er rateestimatorer for hhv det totale fellesrådsvolumet og det totale soknevolumet. Deres varianser estimeres ved følgende uttrykk – fellesrådsleddet slik:

$$S_f^2(K) = M^2 \left(\frac{\bar{K}_*^f}{\bar{M}_*^f} \right)^2 \left\{ \frac{S^2(\bar{K}_*^f)}{(\bar{K}_*^f)^2} + \frac{S^2(\bar{M}_*^f)}{(\bar{M}_*^f)^2} - 2 \frac{S(\bar{K}_*^f, \bar{M}_*^f)}{\bar{K}_*^f \bar{M}_*^f} \right\}$$

der de ulike (nye) størrelsene som inngår er:

$\bar{K}_*^f = \frac{1}{n'} \sum_i K_i^f$ = gjennomsnittlig totalt årsvolum pr *utvalgt tjenestested* som kun omfatter fellesråd (i alt n' slike i utvalget)

$\bar{M}_*^f = \frac{1}{n'} \sum_i M_i$ = gjennomsnittlig medlemstall i de n' kommunene der vi har tjenestesteder som kun omfatter fellesråd i utvalget

$$S^2(\bar{K}_*^f) = \frac{1}{n'(n'-1)} \sum_i [K_i^f - \bar{K}_*^f]^2$$

$$S^2(\bar{M}_*^f) = \frac{1}{n'(n'-1)} \sum_i [M_i - \bar{M}_*^f]^2$$

$$S(\bar{K}_*^f, \bar{M}_*^f) = \frac{1}{n'(n'-1)} \sum_i [K_i^f - \bar{K}_*^f] [M_i - \bar{M}_*^f]$$

Det andre leddet (sokneleddet) i summen estimeres ved:

$$S_s^2(K) = M^2 \left(\frac{\bar{K}_*^s}{\bar{M}_*^s} \right)^2 \left\{ \frac{S^2(\bar{K}_*^s)}{(\bar{K}_*^s)^2} + \frac{S^2(\bar{M}_*^s)}{(\bar{M}_*^s)^2} - 2 \frac{S(\bar{K}_*^s, \bar{M}_*^s)}{\bar{K}_*^s \bar{M}_*^s} \right\}$$

der de ulike størrelsene som inngår er:

$\bar{K}_*^s = \frac{1}{n''} \sum_{i,j} K_{ij}^s$ = gjennomsnittlig totalt årsvolum for sokn pr *utvalgt tjenestested* (n'' slike i utvalget)

$\bar{M}_*^s = \frac{1}{n''} \sum_{i,j} M_{ij}$ = gjennomsnittlig antall soknemedlemmer pr tjenestested som inngår i de n'' tjenestestedene

$$S^2(\bar{K}_*^s) = \frac{1}{n''(n''-1)} \sum_{i,j} [K_{ij}^s - \bar{K}_*^s]^2$$

$$S^2(\bar{M}_*^s) = \frac{1}{n''(n''-1)} \sum_{i,j} [M_{ij} - \bar{M}_*^s]^2$$

$$S(\bar{K}_*^s, \bar{M}_*^s) = \frac{1}{n''(n''-1)} \sum_{i,j} [K_{ij}^s - \bar{K}_*^s][M_{ij} - \bar{M}_*^s]$$

For å estimere variansen til det årlige X -volumet, vil vi gjøre en forenkling, nemlig å forutsette at det årlige totalvolumet er gitt. Variansene som estimeres, blir således betingede varianser – gitt totalt årsvolum. (Estimering av de ubetingede variansene vil medføre betydelige komplikasjoner.)

Variansen til årlig X -volum kan uttrykkes slik:

$$\text{var } \hat{X} = \sum_p \alpha_p^2 \text{var } \hat{X}^p = \hat{K}^2 \sum_p \alpha_p^2 \text{var } \hat{P}^p$$

De enkelte variansledd under summasjonen estimeres ved

$$S_p^2(\mathbf{X}) = \left(\frac{\mathbf{x}_*^p}{\mathbf{k}_*^p} \right)^2 \left\{ \frac{S^2(\bar{x}_*^p)}{(\bar{x}_*^p)^2} + \frac{S^2(\bar{k}_*^p)}{(\bar{k}_*^p)^2} - 2 \frac{S(\bar{x}_*^p, \bar{k}_*^p)}{\bar{x}_*^p \bar{k}_*^p} \right\}$$

der leddene på høyre side står for

$$\mathbf{x}_*^p = \sum_{i,j} \mathbf{x}_{ij}^p = \text{sum av observerte } X\text{-volum i undersøkelsesperioden for alle tjenestesteder } (n) \text{ i}$$

utvalget pulje p

$$\bar{x}_*^p = \frac{1}{n_p} \sum_{i,j} \mathbf{x}_{ij}^p = \text{gjennomsnittlig observert } X\text{-volum i undersøkelsesperioden pr tjenestested i}$$

utvalget pulje p

$$\mathbf{k}_*^p = \sum_{i,j} \mathbf{k}_{ij}^p = \text{sum av observerte kopivolum i undersøkelsesperioden for alle tjenestesteder } (n) \text{ i}$$

utvalget pulje p

$$\bar{k}_*^p = \frac{1}{n_p} \sum_{i,j} \mathbf{k}_{ij}^p = \text{gjennomsnittlig observert kopivolum i undersøkelsesperioden pr tjenestested i}$$

utvalget pulje p

$$S^2(\bar{x}_*^p) = \frac{1}{n_p(n_p-1)} \sum_{i,j} [x_{ij}^p - \bar{x}_*^p]^2$$

$$S^2(\bar{k}_*^p) = \frac{1}{n_p(n_p - 1)} \sum_{i,j} [k_{ij}^p - \bar{k}_*^p]^2$$

$$S(\bar{x}_*^p, \bar{k}_*^p) = \frac{1}{n_p(n_p - 1)} \sum_{i,j} [x_{ij}^p - \bar{x}_*^p] [k_{ij}^p - \bar{k}_*^p]$$

Variansestimatorene ovenfor gjelder volumstørrelsene \mathbf{K} og \mathbf{X} . Estimering av variansene til estimatene for raten \mathbf{P} og andelene $\mathbf{R}(\mathbf{X})$ kan skje ved å benytte tilsvarende fremgangsmåte som ovenfor. I Tabell G foran (avsnitt 4.1) er *konfidensintervaller* vist for enkelte hovedresultater i undersøkelsen. Konfidensintervallene er anslått som estimatverdier +/- 2 ganger kvadratroten av variansen til estimatet. Dette svarer til en konfidenskoeffisient på 95%.

VEDLEGG

- Brev fra MMI og KA til tjenestestedene vedr oppnevning av kontaktperson for undersøkelsen
- Materiell sendt kontaktpersonen til bruk i undersøkelsen
- Kodeveiledning
- Kodeskjema

(OBS: ikke tatt med i nettutgaven – kan fås ved henv. til KAs nettredaktør (KA/red. anm.))