

Arbeidsnotater

S T A T I S T I S K S E N T R A L B Y R Å

Dronningensgt. 16, Oslo-Dep., Oslo 1. Tlf. 41 38 20

IO 76/4

17. februar 1976

Estimering og prediksjon ved bruk av metoder fra tidserieanalysen
i Byråets arbeidskraftundersøkelser

av

John Dagsvik^{*)}

INNHold

	Side
1. Innledning	1
2. Lineære stokastiskemodeller med et endelig antall parametre	3
3. Identifikasjon av modeller for arbeidskraftvariable	5
4. Autoregressiv representasjon og endring av tidsenhet	5
5. Estimering og prediksjon på grunnlag av løpende utvalgsundersøkelser	7
6. Konklusjon	12
7. Den matematiske teori for prediksjon av en tidserie ved hjelp av en annen	13
8. Spesielle modeller	18
 Referanser	 22
 Tabeller og diagrammer	 23

*) Jeg er takknemlig for kommentarer fra Sevaldson og Vannebo som har bidratt til en vesentlig forbedring av fremstillingen.

Sammendrag

Et klassisk problem i teorien for tidserier er filtrering av signal som er belagt med støy. Betrakter vi samlet antall sysselsatte i en næring som "signalet" observeres ikke dette direkte, men estimeres ved løpende utvalgsundersøkelser. "Støyen" er altså i dette tilfelle utvalgsfeilen.

Ved å anvende denne teorien har vi utledet estimatorer, prediktorer og deres forventede kvadratavvik med henblikk på estimering av arbeidskraftvariable. Som datagrunnlag for identifikasjon av autokorrelasjonsstrukturen for disse variablene har vi brukt Sysselsettingstatistikken fra perioden 1955 - 1970.

1. Innledning

For mange variable som studeres i økonomi og samfunnsfag er det en vesentlig korrelasjon mellom variabelverdiene på ulike tidspunkter. Vi vet for eksempel at for de fleste personer er tilknytningen til arbeidsmarkedet i en viss grad stabil; med andre ord er totalt antall sysselsatte fordelt på arbeidsmarkedstatus ved ulike tidspunkter korrelerte. I tillegg vet vi at personers forhold til arbeidsmarkedet ofte er sesongavhengig, noe som betyr at det er spesielt høy korrelasjon mellom arbeidsmarkedstatus en måned ett år og samme måned det påfølgende år.

I tidserieanalysen studerer en systematisk hvordan tidserier for autokorrelerte variable kan utnyttes til å konstruere modeller som tar vare på et underliggende korrelasjonsmønster. I tidligere anvendelser har det vært vanlig å dekomponere tidserien i en langtidstrend, en eventuell periodisk komponent og et stokastisk restledd. Prognoser ble laget ved å tilpasse og ekstrapolere et polynom eller en annen valgt funksjon av tiden for trendkomponenten og så justere dette med en "sesongkomponent" som for eksempel kunne være additiv eller multiplikativ. En annen framgangsmåte er å bygge analysen på en eller annen form for treghets- eller lag-modell. Dersom $\{x_t\}$ er den aktuelle prosessen (med måned som tidsenhet) vet vi at x_t er korrelert med x_{t-1} , x_{t-2} , ..., altså er det nærliggende å anta en regresjonsmodell av typen

$$x_t = \sum_{j=1}^{\infty} \gamma_j x_{t-j} + a_t$$

hvor restleddene a_t , a_{t-1} , ... er uavhengige identisk fordelte og a_t er uavhengig av strukturen. Våre høyresidevariable er tidligere verdier av den aktuelle prosessen, altså er dette en autoregresjonsmodell. Denne modellen har et uendelig antall regresjonskoeffisienter men en vil i praksis anta at tilnærmelsen er god med et begrenset antall regresjonsledd. Imidlertid vil en i mange situasjoner være nødt til å ta med et stort antall regresjonsledd for å få god tilnærmelse. Vi skal derfor uttrykke modellen på en alternativ form som medfører at vi kan redusere antall parametre vesentlig.

Lar vi x_t være antall sysselsatte i en eller annen hovednæring, er tidserien for x_t nettopp karakterisert ved trender som har vekslende stigning fra periode til periode. (Se diagrammene på sidene 28 til 35). Vi vil skrive modellen for x_t på følgende form

$$x_t = x_{t-1} + s_t$$

Komponenten s_t uttrykker endringen fra måned $t-1$ til måned t som avhenger både av trenden og sesongsvingningene.

Vi har forutsatt s_t lik

$$s_t = x_{t-12} - x_{t-13} + e_t$$

hvor e_t er et stokastisk restledd med forventning null.

Siden restleddene ved ulike tidspunkt kan være korrelerte er det hensiktsmessig å uttrykke e_t som en lineærkombinasjon av uavhengige identisk fordelte "sjokk"-variable a_t, a_{t-1}, \dots . Vi ser at dersom restleddene er null vil trendens stigning være konstant mens negative (positive) restleddverdier reduseres (øker) dens stigning. Bestemmende for hvor fort trenden kan skifte retning er variansen og autokorrelasjonstrukturen til restleddet. Opplegget kan beskrives som en transformasjon av variable hvor $d_t = x_t - x_{t-12}$ og $u_t = d_t - d_{t-1}$. Vi studerer altså "differensvariablen" u_t gitt ved likningen $u_t = e_t$ i stedet for den "opprinnelige" variable x_t . De dataene vi har bekrefter at modellen ovenfor er god. Vi har spesielt identifisert en struktur på restleddet av typen

$$e_t = a_t - \beta a_{t-12}$$

hvilket betyr at det bare er e_t og e_{t-12} som er korrelerte.

Modellen er altså et forsøk på å beskrive "tregheten" i prosessens bevegelser uten at vi har bundet oss til noen bestemt funksjonsform for x_t som funksjon av t . Modellen kan transformeres til autoregressiv form av typen ovenfor, hvor regresjonskoeffisientene γ_j kan uttrykkes ved β .

Den framgangsmåten vi har brukt har altså gitt en autoregresjonsmodell der regresjonskoeffisientene avhenger bare av en parameter.

Som vi ser er modellen mekanisk i den forstand at den ikke inneholder forklaringsvariable. Vårt utgangspunkt er at det eksisterer en eller annen regelmessighet i prosessens bevegelser som vi ønsker å identifisere uten nødvendigvis å "forklare" den ut over sammenhengen med et årsmønster i tidsforløpet. Teoretisk er det imidlertid fullt mulig å bygge inn forklaringsvariable i modellene selv om det i så fall kan by på tekniske problemer i anvendelsene.

Prognoser for tidspunkt $t + 1$ kan nå lages ved å ta forventningen av x_{t+1} gitt x_t, x_{t-1}, \dots . Det kan vises at a_{t+1} er uavhengig av x_t, x_{t-1}, \dots slik at prediktoren kan uttrykkes umiddelbart ved den autoregressive representasjonen ovenfor.

En slik punktprediksjon alene gir ikke nødvendigvis et bedre resultat enn ved mer tradisjonelle metoder. Den gir bare en av flere mulige x_{t+1} -verdier, nemlig midtpunktet i et prediksjonsintervall. Isteden for en enkel punktprediksjon gir vår metode et prediksjonsintervall som er lite dersom prosessen beveger seg tregt og stort i motsatt tilfelle. Prediksjonsintervallet dekker altså mengden av mulige x_{t+1} -verdier når "treghetsmekanismen" til prosessen er gitt.

Vanligvis benyttes metodene fra tidserieanalysen på data som er direkte observasjoner av de variable. Vi er imidlertid først og fremst interessert i å behandle situasjoner hvor verdiene av de variable ikke observeres direkte, men estimeres. Eksempler på dette er variablene i

Byråets arbeidskraftundersøkelser (AKU). I AKU estimeres de aktuelle variablene ved å benytte observasjoner fra et utvalg. Ved disse estimatene er utvalgsfeilen av betydelig størrelse og vi skal her studere hvordan tidseriemodeller kan benyttes til å redusere utvalgsfeilen for nivå- og endringstall på det aktuelle tidspunkt samt å predikere disse på grunnlag av de respektive estimatene.

Dette er nettopp en situasjon hvor teorien for filtrering av signal deformert av støy kan benyttes. Vårt "signal" er eksempelvis nivåtallene for en bestemt næring x_t og "støyen" som signalet er belagt med er utvalgsfeilen v_t . Vi observerer altså y_t , der

$$y_t = x_t + v_t$$

hvor x_t og v_t antas stokastisk uavhengig. Ved hjelp av modeller for x_t og v_t kan vi finne et estimat $\hat{x}_t(0)$ for x_t som minimerer forventet kvadratavvik, $E\{\hat{x}_t(0) - x_t\}^2$ og vi kan lage prognoser på tilsvarende måte. Siden de prosessene vi studerer har både trender og sesongvariasjoner blir matematikken komplisert og forutsetter kjennskap til kompleks funksjonsteori. De eksemplene som er behandlet i lærebøkene er vesentlig enklere enn våre anvendelser, slik at det har vært nødvendig å utlede nye resultater. Framstillingen er derfor hovedsakelig delt i to - en ikke-matematisk del hvor resultatene presenteres og diskuteres og en matematisk del hvor resultatene bevises og behandles i detalj.

I et tidligere arbeid (Dagsvik 1975) har jeg studert effekten av å kombinere estimater ved flere tidspunkter. (Sammensatt estimering.) I det arbeidet var det imidlertid ikke forutsatt noe om sammenhengen mellom påfølgende verdier av de populasjonstotaler en ønsker å estimere.

Testing og estimering av modellene behandles ikke i dette notatet, men slike problemer vil bli tatt opp i et senere arbeid.

2. Lineærestokastiske modeller med et endelig antall parametre

En tidserie er et sett av observasjoner for et gitt kjennetegn ordnet kronologisk. Vi skal betrakte en slik tidserie som én av mange mulige realisasjoner av en stokastisk prosess $\{x_t\}$. Antall sysselsatte fordelt på næringer tolket som tilfeldige realisasjoner blant mengden av mulige realisasjoner er eksempler på slike prosesser. En stokastisk prosess sies å være strengt stasjonær dersom dens sannsynlighetsfordeling forblir uendret når tiden øker. En strengt stasjonær prosess svinger altså tilfeldig omkring et konstant nivå. En stokastisk prosess sies å være stasjonær dersom (Wide sense stationary process) kovariansfunksjonen

$$E\{(x_{t+k} - E(x_{t+k})) (x_t - E(x_t))\}, k = 0, 1, 2, \dots$$

er uavhengig av t . Dersom prosessen er normalfordelt og stasjonær vil den altså svinge omkring en eller annen deterministisk funksjon av tiden og $\{x_t - E(x_t)\}$ vil være strengt stasjonær. Befolkningens fordeling på tilstander i arbeidsmarkedet er eksempler på ikke-stasjonære prosesser, da disse har både sesongvariasjoner og trender med vekslende stigning. Det første som må gjøres er derfor å transformere prosessen til en stasjonær prosess. Et vanlig kriterium for å avgjøre om en prosess er stasjonær er å studere hvor raskt den estimerte autokorrelasjonsfunksjonen avtar med økende argument. Dersom dette skjer "forholdsvis raskt" vil en si at en har stasjonærhet. Det viser seg at en ofte kan oppnå stasjonærhet ved å ta mange nok differenser av typen $x_t - x_{t-s}$ for bestemte verdier av s , eller ved å ta tilsvarende differenser av prosessen $\log X_t$. Siden de prosessene vi betrakter har sesongsvingninger med periode 12 når måned er tidsenhet, er det naturlig å prøve med differensoperasjonen $\Delta_{12} x_t = x_t - x_{t-12}$. Dersom

vi fremdeles har ikke-stasjonæritet prøvervidere etter om differensoperasjonen $\Delta_1(x_t - x_{t-12}) = \Delta_1 \Delta_{12} x_t = x_t - x_{t-12} - (x_{t-1} - x_{t-13})$ er tilstrekkelig. Eventuelt må vi ta flere differenser. I våre anvendelser viser det seg (avsnitt 3) at det er nok å foreta transformasjonene $\Delta_1 \Delta_{12} x_t$. Vi skal derfor anta at prosessen $\{e_t\}$ definert ved

$$(2.1) \quad e_t = \Delta_1 \Delta_{12} x_t = x_t - x_{t-1} - x_{t-12} + x_{t-13}$$

er stasjonær. På grunnlag av autokorrelasjonsfunksjonen

$$(2.2) \quad \rho_r = \text{Korr}\{e_t, e_{t-r}\}, \quad r = 1, 2, \dots$$

ønsker vi å bestemme en modell for $\{e_t\}$ av typen

$$(2.3) \quad e_t = \sum_{j=1}^n \alpha_j e_{t-j} + \sum_{j=0}^q \phi_j a_{t-j}$$

hvor a_t, a_{t-1}, \dots er uavhengige identisk fordelte variable med forventning null (ikke observerbare). En modell av denne typen kalles en blandet autoregressiv - glidende gjennomsnittsmodell (ARMA). Denne typen modeller har vist seg å være ganske fleksibel med lite antall parametre. Modellen (2.3) kan uttrykkes som en lineærkombinasjon i a_t, a_{t-1}, \dots

$$(2.4) \quad e_t = \sum_{j=0}^{\infty} \beta_j a_{t-j}$$

Den kalles da den glidende gjennomsnittsrepresentasjonen av $\{e_t\}$. En slik representasjon eksisterer imidlertid bare for en stasjonær prosess.

Endelig eksisterer det en autoregressiv representasjon definert ved

$$(2.5) \quad e_t = a_t + \sum_{j=1}^{\infty} \gamma_j e_{t-j}$$

Siden a_t, a_{t-1}, \dots er uavhengige følger det av (2.4) at a_t er uavhengig av e_{t-j} for $j \geq 1$. Altså er restleddet i den autoregressive representasjonen uavhengig av strukturdelen. Lar vi $L(\cdot)$ være det glidende gjennomsnitt (lineære filter) definert ved

$$L(e_t) = e_t - \sum_{j \geq 1} \gamma_j e_{t-j}$$

vil altså operasjonen L transformere prosessen til en prosess med autokorrelasjonsfunksjon lik null.

Den autoregressive representasjonen eksisterer også selv om prosessen er ikke-stasjonær. Alle disse representasjonene er ekvivalente (dersom de eksisterer), men det er som regel mest hensiktsmessig å arbeide med (2.3) under identifikasjonsprosedyren.

3. Identifikasjon av modeller for arbeidskraftvariable

Som kjent var Sysselsettingstatistikken før 1971 basert på fullstendige tellinger. Disse dataene gir derfor mulighet til blant annet å identifisere modeller for sysselsatte fordelt på næring. I figur 1 til 7 er tidseriene for noen næringer i tidsrommet 1955 - 1970 framstilt grafisk med måned som tidsenhet. Vi merker oss at sesongvariasjonene er meget regelmessige fra år til år samt at "langtidstrenden" er stigende i noen - og avtagende i andre perioder.

Autokorrelasjonsberegninger vi har gjort for noen næringer (tabell 1 - 4) bekrefter at vi må ta differensene $x_t - x_{t-1} - x_{t-12} + x_{t-13}$ for å oppnå stasjonaritet. Beregningene viser forøvrig at estimatene $\hat{\rho}_r$, $r \geq 1$ definert ved (2.2) er små sammenliknet med $\hat{\rho}_{12}$. Dette tyder på at $\rho_r = 0$ for $r \neq 12$. En slik hypotese kan testes, men vi skal la det ligge her. Dersom $\rho_r = 0$ for $r \neq 12$ ser vi umiddelbart at e_t har glidende gjennomsnittsrepresentasjon,

$$(3.1) \quad e_t = a_t - \beta a_{t-12}, \quad |\beta| < 1$$

og en har sammenhengen

$$(3.2) \quad \rho_{12} = -\beta / (1 + \beta^{12})$$

Variabelen a_t kalles ofte det tilfeldige "sjokk" ved tidspunkt t og den kan tolkes som avviket fra forventet x_t -verdi gitt x_{t-1}, x_{t-2}, \dots . Er x_t antall sysselsatte i måned t , er altså a_t avviket fra den "normale" økningen (nedgang) fra forrige måned. Det følger videre fra (3.1) at et sjokk som påvirket prosessen en bestemt måned ett år vil påvirke prosessen samme måned neste år dempet med faktoren β . En har altså en bølgemekanisme med 12 måneders periode som karakteriseres ved at dersom et enkelt sjokk påvirker prosessen ved et bestemt tidspunkt tar det 25 måneder før prosessen har stabilisert seg med den nye trend.

For å illustrere hvordan en slik modell fungerer, har vi laget en kunstig tidserie ved å begynne med fem startverdier, og deretter la prosessen påvirkes av sjokkverdier med forskjellig størrelse og ulik avstand mellom sjokkene. Denne "tidserien" er framstilt i figur 8.

4. Autoregressiv representasjon og endring av tidsenhet.

Vi skal i det følgende gå ut fra at prosessen $\{x_t\}$ genereres ved modellen

$$(4.1) \quad x_t - x_{t-1} - x_{t-k} + x_{t-k-1} = e_t = a_t - \beta a_{t-k}$$

hvor k er perioden. Dersom tidsenheten er måned, kvartal eller år er k henholdsvis 12, 4 og 1. Det er nå enkelt å innse at dersom (4.1) gjelder med måned som tidsenhet vil også (4.1) gjelde om tidsenheten er kvartal eller år, med samme parameter β . Dersom $k = 12$ er altså

$$x_t - x_{t-12} = x_{t-1} - x_{t-13} + e_t$$

som medfører at

$$\begin{aligned} x_t - x_{t-12} &= x_{t-3} - x_{t-15} + e_t + e_{t-1} + e_{t-2} \\ &= x_{t-3} - x_{t-15} + a_t + a_{t-1} + a_{t-2} - \beta(a_{t-12} + a_{t-13} + a_{t-14}), \end{aligned}$$

Vi ser at vi kan gå over til kvartalsenheten ved å innføre tiden målt i kvartaler idet vi setter

$$\begin{array}{l} \text{for } t = \begin{array}{|c|c|c|c|c|c|} \hline t & t-3 & t-6 & t-9 & t-12 & t-15 \\ \hline \end{array} \\ \text{er } \begin{array}{|c|c|c|c|c|c|} \hline T & T-1 & T-2 & T-3 & T-4 & T-5 \\ \hline \end{array} \end{array}$$

og

$$a_T = a_t + a_{t-1} + a_{t-2} .$$

Innsetting gir altså

$$x_T - x_{T-4} = x_{T-1} - x_{T-5} + a_T - \beta a_{T-4} .$$

Altså er økningen i x_t -verdi et år lik økningen i løpet av ett år fra fem kvartaler bakover opptil forrige kvartal pluss restleddet for kvartalet. Restleddet er sammensatt av kvartalsjokk som er summen av månedsjokkene i kvartalet. Helt analogt vises det at (4.1) gjelder når tidsenheten er år hvor det tilfeldige sjokk er summen av månedsjokkene i løpet av året.

For prognoseformål er det ofte hensiktsmessig å ha modellen på autoregressiv form. Denne er

$$(4.2) \quad x_t = a_t + x_{t-1} + (1-\beta) \sum_{r=1}^{\infty} \beta^{r-1} (x_{t-kr} - x_{t-kr-1}) .$$

La $\tilde{x}_t(\ell)$ være prediktoren for $x_{t+\ell}$ ved tidspunkt t definert ved

$$(4.3) \quad \tilde{x}_t(\ell) = E\{x_{t+\ell} | x_t, x_{t-1}, \dots\} .$$

Siden $a_{t+\ell}$ har forventning null og er uavhengig av x_t, x_{t-1}, \dots blir

$$E\{a_{t+\ell} | x_t, x_{t-1}, \dots\} = 0 \text{ for } \ell \geq 1 .$$

Vi har videre

$$E\{x_{t+\ell-kr} | x_t, x_{t-1}, \dots\} = \tilde{x}_t(\ell-kr) \text{ når } kr < \ell$$

og

$$E\{x_{t+\ell-kr} | x_t, x_{t-1}, \dots\} = x_{t+\ell-kr} \text{ når } kr \geq \ell .$$

Lar vi

$$x_t^{*(m)} = \begin{cases} x_{t+m} & \text{når } m \leq 0 \\ \tilde{x}_t(m) & \text{når } m > 0 \end{cases}$$

følger det direkte av (4.2) at

$$(4.4) \quad \hat{x}_t^{(l)} = \hat{x}_t^{(l-1)} + (1-\beta) \sum_{r=1}^R \beta^{r-1} (\hat{x}_t^{*(l-kr)} - \hat{x}_t^{*(l-kr-1)}) + M(R)$$

hvor $M(R) \rightarrow 0$ når R øker. (Husk at $|\beta| < 1$). Det kan videre vises at $\hat{x}_t^{(l)}$ definert ved (4.3) minimerer forventet kvadratavvik $E\{\hat{x}_t^{(l)} - x_{t+l}\}^2$. Observer at vi ikke har forutsatt noe om fordelingen til a_t .

5. Estimering og prediksjon på grunnlag av løpende utvalgsundersøkelser

Som før nevnt observeres ikke alltid verdiene av prosessen $\{x_t\}$ direkte, men estimeres for eksempel på grunnlag av observasjoner fra et utvalg. Byråets AKU er et eksempel på en slik situasjon. Vi har derfor

$$(5.1) \quad y_t = x_t + v_t$$

hvor y_t er den estimerte x_t -verdi og v_t er utvalgsfeilen. Vi ønsker å utnytte kjennskapet til prosessene $\{x_t\}$ og $\{v_t\}$ for å forbedre det aktuelle estimatet y_t , samt å lage kortsiktige prognoser. De fundamentale forutsetninger vi gjør er at prosessen $\{v_t\}$ er stasjonær med forventning null og at x_t og v_t er uavhengige.

I Byråets utvalgsplan er utvalgsfeilene ved ulike tidspunkter korrelerte. La oss anta at vi kjenner autokovariansen til $\{v_t\}$. Siden x_t og v_t er uavhengige og autokovariansen til $\{\Delta_1 \Delta_4 x_t\}$ definert ved

$$\Delta_1 \Delta_4 x_t = x_t - x_{t-1} - x_{t-4} + x_{t-5}$$

er kjent kan vi finne autokovariansen til $\{\Delta_1 \Delta_4 y_t\}$ ved å benytte (5.1). Vi kan følgelig bestemme en modell for $\{\Delta_1 \Delta_4 y_t\}$ av typen (2.3). På grunnlag av modellene for $\{x_t\}$, $\{v_t\}$ og $\{y_t\}$ ønsker vi først og fremst å finne den beste lineære estimator

$$(5.2) \quad \hat{x}_t^{(o)} = \sum_{j \geq 0} \zeta_j y_{t-j}$$

for den aktuelle x_t -verdien. Dette gjøres ved å minimere forventet kvadratavvik til $\hat{x}_t(0)$. For å belyse prinsippene i metoden la oss for enkelhets skyld anta at $\{x_t\}$ er stasjonær og at $E\{x_t\} = 0$ som medfører at $E\{v_t\} = 0$. La $\{c_t\}$ være prosessen av tilfeldige sjokk i den glidende gjennomsnittsrepresentasjonen for $\{y_t\}$ dvs.

$$y_t = \sum_{j=0}^{\infty} \kappa_j c_{t-j}$$

der c_t, c_{t-1}, \dots er uavhengige identisk fordelte (se 2.4). Siden $\hat{x}_t(0)$ er uttrykt ved y_t, y_{t-1}, \dots , og y_t kan uttrykkes ved c_t, c_{t-1}, \dots , kan derfor $\hat{x}_t(0)$ uttrykkes ved c_t, c_{t-1}, \dots ,

$$(5.3) \quad \hat{x}_t(0) = \sum_{j \geq 0} \kappa_j c_{t-j} \quad .$$

Definer

$$d_j = \text{Cov}\{x_t, c_{t-j}\} / \text{Var}\{c_t\} = E\{x_t c_{t-j}\} / \text{Var}\{c_t\}$$

som medfører at

$$(5.4) \quad x_t = \sum_{j=-\infty}^{\infty} d_j c_{t-j} \quad .$$

Det følger nå umiddelbart at

$$\begin{aligned} E\{\hat{x}_t(0) - x_t\}^2 &= E\{x_t^2\} - 2E\{x_t \hat{x}_t(0)\} + E\{\hat{x}_t(0)\}^2 \\ &= \text{Var}\{x_t\} - 2 \sum_{j \geq 0} \kappa_j E\{x_t c_{t-j}\} + E\{\sum_{j \geq 0} \kappa_j c_{t-j}\}^2 \quad . \end{aligned}$$

Siden c_t, c_{t-1}, \dots , er uavhengige er

$$E\{\sum_{j \geq 0} \kappa_j c_{t-j}\}^2 = \sum_{j \geq 0} \kappa_j^2 \tau^2$$

hvor $\tau^2 = \text{Var}\{c_t\}$. Altså er

$$\begin{aligned} (5.5) \quad E\{x_t - \hat{x}_t(0)\}^2 &= \text{Var}\{x_t\} - 2 \sum_{j \geq 0} \kappa_j d_j \tau^2 + \tau^2 \sum_{j \geq 0} \kappa_j^2 \\ &= \text{Var}\{x_t\} - \sum_{j \geq 0} d_j^2 \tau^2 + \sum_{j \geq 0} (\kappa_j - d_j)^2 \tau^2 \quad . \end{aligned}$$

Dette uttrykket minimeres for $\kappa_j = d_j$, $j = 0, 1, \dots$. Den beste lineære estimator er derfor

$$(5.6) \quad \hat{x}_t(0) = \sum_{j \geq 0} d_j c_{t-j}$$

med forventet kvadratavvik

$$(5.7) \quad E\{x_t - \hat{x}_t(0)\}^2 = \text{Var}\{x_t\} - \sum_{j \geq 0} d_j^2 \tau^2.$$

Nå kan c_t -verdiene uttrykkes ved y_t -verdiene ved den autoregressive representasjonen for $\{y_t\}$ -prosessen og vektene ζ_j i (5.3) er dermed i prinsippet bestemt. La

$$y_t = (y_t, y_{t-1}, \dots).$$

Vi observerer at for $j < 0$ er c_{t-j} uavhengig av y_t . Videre er c_{t-j} bestemt ved y_t når $j \geq 0$ slik at

$$E\{c_{t-j} | y_t\} = 0 \text{ når } j < 0$$

og

$$E\{c_{t-j} | y_t\} = c_{t-j} \text{ når } j \geq 0.$$

Fra (5.4) og (5.6) får vi derfor at

$$\hat{x}_t(0) = E\{x_t | y_t\}$$

hvilket intuitivt er naturlig. Videre er

$$\begin{aligned} E\{\hat{x}_t(0) - x_t\}^2 &= E\{\hat{x}_t - E\{\hat{x}_t | y_t\}\}^2 \\ &= EE\{[x_t - E\{\hat{x}_t | y_t\}]^2 | y_t\} = E\text{Var}\{x_t | y_t\}. \end{aligned}$$

Forventet kvadratavvik for $\hat{x}_t(0)$ er altså lik forventet varians gitt observasjonene y_t .

Dersom vi ønsker å predikere $x_{t+\ell}$ ved tidspunkt t er utledningen av den optimale prediktor helt analog.

Vi har analogt

$$(5.8) \quad \hat{x}_t(\ell) = E\{x_{t+\ell} | y_t\}$$

og

$$(5.9) \quad E\{\hat{x}_t(\ell) - x_{t+\ell}\}^2 = E \text{Var}\{x_{t+\ell} | y_t\}.$$

Dersom vi kjenner $\hat{x}_t(0)$ og $\hat{x}_t(4)$ er det nå enkelt å finne $\hat{x}_t(4\ell)$, dvs. prediktoren ℓ år framover. Siden x_t er uavhengig av v_t , er $\{a_t\}$ uavhengig av $\{v_t\}$. Videre er a_{t+j} , $j > 0$, uavhengig av x_t, x_{t-1}, \dots , slik at a_{t+j} er uavhengig av y_t . Følgelig får vi at

$$E\{a_{t+j} | y_t\} = 0 \text{ når } j > 0.$$

Likning (4.1) gir derfor (når $k = 4$)

$$\hat{x}_t(n+5) - \hat{x}_t(n+1) = \hat{x}_t(n+4) - \hat{x}_t(n), \quad n = 0, 1, \dots$$

som medfører at

$$\hat{x}_t(4\ell) - x_t(4\ell-4) = \hat{x}_t(4) - \hat{x}_t(0), \ell = 1, 2, \dots$$

Økningen (nedgangen) i løpet av et år er altså konstant hvilket medfører at

$$(5.10) \quad \hat{x}_t(4\ell) = \hat{x}_t(0) + \ell(\hat{x}_t(4) - \hat{x}_t(0)),$$

I vår situasjon er $\{x_t\}$ og dermed $\{y_t\}$ ikke-stasjonære prosesser slik at det ikke eksisterer en representasjon av typen (5.3). Dette problemet kan imidlertid omgås ved å la modellen for $\{x_t\}$ være

$$(5.8) \quad x_t - \theta x_{t-1} - \theta^k x_{t-k} + \theta^{k+1} x_{t-k-1} = e_t$$

istedet for (4.1). Denne prosessen er nemlig stasjonær dersom $|\theta| < 1$. Når $\hat{x}_t(\ell)$ er uttrykt ved $\{y_t\}$ lar vi (5.8) nærme seg den ikke-stasjonære prosessen ved å la $\theta \rightarrow 1$. Som før nevnt er matematikken komplisert og forutsetter kjennskap til kompleks funksjonsteori. Den er derfor skilt ut som et eget kapittel.

Vi skal her se i detalj på modeller for utvalgsfeilen som er tilpasset Byråets spesielle utvalgsplan. La x_{it} være den aktuelle variabel knyttet til person nr. i ved tidspunkt t og la $\mathbf{x}_t = (x_{1t}, x_{2t}, \dots)$. Vi har sammenhengen

$$\begin{aligned} \text{cov}\{v_t, v_{t-k}\} &= E \text{cov}\{y_t - x_t, y_{t-k} - x_{t-k} | x_t, x_{t-k}\} \\ &+ \text{cov}\{E(v_t | x_t, x_{t-k}), E(v_{t-k} | x_t, x_{t-k})\} \\ &= E \text{cov}\{y_t, y_{t-k} | x_t, x_{t-k}\} = E \text{cov}\{y_t, y_{t-k} | \mathbf{x}_t, \mathbf{x}_{t-k}\}. \end{aligned}$$

Nå er

$$\text{cov}\{y_t, y_{t-k} | \mathbf{x}_t, \mathbf{x}_{t-k}\}$$

kovariansen mellom estimatorene y_t, y_{t-k} slik den er definert i klassisk utvalgsteori, dvs. når populasjonsverdiene ikke betraktes som realisasjoner av en stokastisk prosess. I vår teori er altså autokovariansen for utvalgsfeilen forventet verdi av autokovariansen for $\{y_t\}$ gitt populasjonsverdiene. Autokovariansfunksjonen for $\{v_t\}$ kan derfor estimeres ved å ta gjennomsnittet av estimatene for

$$\text{cov}\{y_t, y_{t-k} | \mathbf{x}_t, \mathbf{x}_{t-k}\}$$

over flere tidspunkter.

Det som kjennetegner utvalgsplanen for AKU er at utvalget trekkes i to trinn der første trinns enheter er de samme fra undersøkelse til undersøkelse. Utvalget innen annen trinns enheter er roterende slik at halvparten av utvalget er felles ett kvartal og det påfølgende, og ett kvartal og samme kvartal året etter. En fjerdedel av utvalget er felles ett kvartal og tre kvartaler senere, og ett kvartal og fem kvartaler senere mens utvalgene ett kvartal, to kvartaler og seks kvartaler senere er ikke-overlappende. Rotasjonsplanen kan beskrives ved følgende

figur

	Kvartal	1	2	3	4	5	6	7
Pulje								
1		x						
2		x	x					
3			x	x				
4				x	x			
5		x			x	x		
6		x	x			x	x	
7			x	x			x	x
8				x	x			x
9					x	x		
10						x	x	
							x	x
								x

Den totale autokovariansen for utvalgsfeilen kan som kjent dekomponeres i summen av autokovariansen mellom - og innen utvalgsområdene.

I følge rotasjonsplanen er altså autokovariansen innen utvalgsområdene null bortsett fra for variable som er ett, tre, fire eller fem kvartaler fra hverandre i tid.

Siden første trinns enheter er de samme hele tiden vil utvalgsfeilene mellom utvalgsområdene være autokorrelerte uendelig mange tidsenheter bakover.

Vi vil anta at utvalgsfeilen $\{v_{1t}\}$ mellom utvalgsområdene tilfredsstiller en modell av typen

$$(5.11) \quad v_{1t} = \lambda v_{1,t-1} + \mu v_{1,t-4} - \lambda\mu v_{1,t-5} + b_{1t}$$

hvor b_{1t} , $b_{1,t-1}$, .. er uavhengige identisk fordelte. Vi kan videre uttrykke autokorrelasjonsfunksjonen ved parametrene λ , μ og $\text{Var}\{v_{1t}\}$ (og vice versa). Når λ og μ samt $\text{var}\{v_{1t}\}$ er estimert kan vi følgelig finne autokorrelasjonsfunksjonen til den totale utvalgsfeilen v_t . Ved hjelp av autokorrelasjonsfunksjonen til $\{v_t\}$ og til $\{e_t\}$ kan vi bestemme modellen for den observerte prosessen $\{y_t\}$. Denne får formen

$$(5.12) \quad \Delta_1 \Delta_4 y_t = \lambda \Delta_1 \Delta_4 y_{t-1} + \mu \Delta_1 \Delta_4 y_{t-4} - \lambda \mu \Delta_1 \Delta_4 y_{t-5}$$

$$+ \sum_{j=0}^{15} v_j^c y_{t-j}$$

Prediktoren for $x_{t+\ell}$, $\ell \geq 0$ får formen

$$(5.13) \quad \hat{x}_t^{(\ell)} = \sum_{j=1}^{15} v_j \hat{x}_{t-j}^{(\ell)} + \sum_{j=0}^9 \eta_j^{(\ell)} y_{t-j}$$

hvorkoeffisientene $\eta_j^{(\ell)}$ samt forventet kvadratavvik er funnet i avsnitt 8.

La oss til slutt betrakte situasjonen når tidsenheten er år og vi antar at utvalgsfeilen mellom utvalgsområdene tilfredsstillere modellen

$$v_{1t} = \lambda v_{1t-1} + b_{1t}$$

Vi får da følgende modell for $\{y_t\}$

$$(5.14) \quad \Delta_1^2 y_t = \lambda \Delta_1^2 y_{t-1} + v_1 c_t + v_2 c_{t-1} + v_3 c_{t-3}$$

Spesielt blir estimatoren for x_t lik

$$(5.15) \quad \hat{x}_t(0) = v_1 \hat{x}_{t-1}(0) + v_2 \hat{x}_{t-2}(0) + v_3 \hat{x}_{t-3}(0) \\ + \eta_0 y_t + \eta_1 y_{t-1} + \eta_2 y_{t-2}$$

hvor koeffisientene η_j er gitt ved setning 7.2 (ii) og forventet kvadratavvik er gitt ved (8.9).

6. Konklusjon

Vi har sett at prediksjons- og estimeringsproblemet metodisk kan løses eksplisitt både for estimatorene, prediktorer og deres respektive forventede kvadratavvik. Presisjonsgevinsten ved metoden vil være sterkt avhengig av variansen σ^2 til det tilfeldige sjokk a_t . Ser vi bort fra usikkerheten i parameterestimaterne vil presisjonsgevinsten være liten dersom σ^2 er stor, mens estimatoren for x_t på tidspunkt t er eksakt i det ekstreme tilfellet $\sigma^2 = 0$. Stor σ^2 betyr jo at x_t kan anta "ikke forventede" verdier med stor sannsynlighet. Det er derfor åpenbart viktig å foreta observasjoner med korte mellomrom. Om vi for eksempel observerer y_t -verdien bare en gang i året istedet for kvartalsvis vil σ^2 multipliseres med 4 siden det tilfeldige sjokk pr. år er summen av kvartalsjokkene.

For å identifisere modellen for $\{x_t\}$ har vi som tidligere nevnt benyttet data fra perioden 1955 - 1970. Estimering av parametrene i $\{x_t\}$ -modellen på grunnlag av disse dataene kan gjøres ved standard datamaskinprogrammer som foreligger. Nå er imidlertid variabeldefinisjonene noe forskjellige i dag enn for tidsrommet før 1970, slik at en bør vurdere påliteligheten av disse estimatene for bruk i framtidig AKU estimering.

For å få brukbar modellestimering er det nødvendig med tiserier på 30 - 40 observasjoner og siden tidseriene på grunnlag av AKU er svært korte er det følgelig ønskelig å benytte data fra før 1970. Dette er en av grunnene til at det er av stor interesse å undersøke hvor robust estimatoren $\hat{x}_t(0)$ er med hensyn på endringer i parametrene for $\{x_t\}$ -modellen og $\{v_t\}$ -modellen.

Estimering av parametrene i modellen for utvalgsfeilen kan gjøres ved hjelp av standard metoder fra utvalgsteorien på grunnlag av de individuelle observasjoner.

Når parametrene i modellene for $\{x_t\}$ og $\{v_t\}$ er estimert kan parametrene i modellen for $\{y_t\}$ finnes ved en såkalt Wiener-Hopf faktorisering. Det er laget en enkel algoritme for datamaskin som løser dette faktoreringsproblemet i praksis. (Wilson 1969)

7. Den matematiske teori for prediksjon av en tidserie ved hjelp av en annen

I dette kapitlet skal vi studere i detalj den matematiske analysen av prediksjons- og estimeringsproblemet. Teorien er delvis basert på Whittle (1962) men det er nødvendig å utlede prediktorer og forventet kvadratavvik spesielt i vårt tilfelle fordi de modellene som behandles som eksempler i litteraturen er vesentlig enklere enn de som er aktuelle for oss.

De prosessene vi studerer er $\{x_t\}$, $\{y_t\}$ og $\{v_t\}$ hvor $\{x_t\}$ og $\{v_t\}$ er forutsatt uavhengige og

$$(7.1) \quad y_t = x_t + v_t.$$

Vi definerer operatorene $\Delta_\theta, \Gamma, \phi$ og Ψ ved

$$\Delta_\theta = \Delta_\theta(B) = (1-\theta B)(1-\theta^k B^k), \quad \Delta_\theta = \Delta \text{ når } \theta=1,$$

$$\Gamma = \Gamma(B) = 1 - \Gamma_1 B - \Gamma_2 B^2 - \dots - \Gamma_p B^p,$$

$$\phi = \phi_k(B) = (1-\beta B^k),$$

og

$$\Psi = \Psi(B) = 1 - \psi_1 B - \psi_2 B^2 - \dots - \psi_q B^q$$

hvor B er den baklengse skiftoperator, $Bx_t = x_{t-1}$ og β, Γ_j, ψ_j er parametre. De modellene vi tar utgangspunkt i er

$$(7.2) \quad \Delta(B)x_t = \phi(B)a_t, \quad |\theta| < 1,$$

$$(7.3) \quad \Gamma(B)v_t = \psi(B)b_t$$

hvor $a_t, a_{t-1}, \dots, b_t, b_{t-1}, \dots$ er uavhengige identisk fordelte variable med forventning null og varianser $\sigma^2 = \text{Var}\{a_t\}, \text{Var}\{b_t\} = h^2$.

Prosessene er stasjonære dersom røttene i de karakteristiske polynomene $\Delta_\theta(z), \Gamma(z)$ og $\phi_k(z)$ ligger utenfor enhetsirkelen. I så fall er også de respektive operatorene invertible.

La $g_r = \text{cov}\{x_t, x_{t-r}\}$, $r = \dots -1, 0, 1, 2, \dots$ være autokovariansfunksjonen til en vilkårlig prosess $\{x_t\}$ og la

$$(7.4) \quad G_x(z) = \sum_r g_r z^r$$

være den tilhørende genererende funksjonen. Dersom

$$x_t = \gamma(B) a_t = \sum_{j \geq 0} \gamma_j a_{t-j}$$

er den glidende gjennomsnittsrepresentasjonen til $\{x_t\}$, gjelder generelt

$$(7.5) \quad G_x(z) = \sigma^2 \gamma(z) \gamma(z^{-1}) = \sigma^2 |\gamma(z)|^2$$

i et område som inneholder enhets sirkelen. Herav ser vi umiddelbart at dersom $A(B)$ er den autoregressive operatoren definert ved

$$a_t = A(B)x_t = \sum_{j \geq 0} A_j x_{t-j}$$

blir

$$(7.6) \quad G_x(z) = \sigma^2 / A(z) A(z^{-1}).$$

Dersom kovariansstrukturen er kjent kan altså modellen bestemmes ved en faktorisering av typen (7.5) og (7.6). En slik faktorisering kalles kanonisk Wiener-Hopf faktorisering og det kan vises at den er entydig under visse forutsetninger.

Dette faktoreringsproblemet forenkles betydelig dersom $G_x(z)$ er en rasjonal funksjon. Siden enhver analytisk funksjon kan approksimeres ved rasjonale funksjoner er det ingen vesentlig innskrenkning å begrense seg til å studere $G_x(z)$ av typen

$$G_x(z) = P(z)/Q(z)$$

hvor P og Q er polynomer. En kan dermed faktorisere P og Q hver for seg. I vårt tilfelle finner vi spesielt

$$P(z) = \phi(z)\phi(z^{-1}) \quad \text{og} \quad Q(z) = \Delta_\theta(z) \Delta_\theta(z^{-1}).$$

I praksis går en gjerne den motsatte vei, dvs. en beregner $G_x(z)$ for en rekke modeller av typen (2.4) slik at identifikasjonsproblemet består i å velge de enkleste modellene som gir en autokorrelasjonsfunksjon med god tilpasning til data. Ved bruk av (7.5) kan vi nå bestemme modellen for $\{y_t\}$. Siden $\{x_t\}$ og $\{v_t\}$ er uavhengige er

$$G_y(z) = G_x(z) + G_v(z).$$

Herav følger det at $\{y_t\}$ genereres ved modellen

$$(7.7) \quad \Gamma(B) \Delta_\theta(B) y_t = \xi_\theta(B) c_t$$

hvor ξ_θ bestemmes ved

$$(7.8) \quad \tau^2 \xi_\theta(z) \xi_\theta(z^{-1}) = \sigma^2 \phi(z) \phi(z^{-1}) \Gamma(z) \Gamma(z^{-1}) + h^2 \psi(z) \psi(z^{-1}) \Delta_\theta(z) \Delta_\theta(z^{-1})$$

der $\tau^2 = \text{Var} \{c_t\}$.

Vi innfører nå følgende notasjoner. La $f(z)$ være en kompleks funksjon med Laurentrekkeutvikling

$$f(z) = \sum_{j=-\infty}^{\infty} f_j z^j$$

gyldig i et eller annet område. Vi definerer følgende operatører

$$(7.9) \quad f(z)_+ = \sum_{j \geq 0} f_j z^j, \quad f(z)_- = \sum_{j < 0} f_j z^j$$

og

$$(7.10) \quad \mathcal{A}f(z) = f_0.$$

De funksjonene vi kommer til å behandle vil alltid ha Laurentrekkeutvikling i et område som inneholder enhetssirkelen. Vi har derfor alltid

$$(7.11) \quad f(z)_+ = \frac{1}{2\pi i} \int_{|u|=1} \frac{f(u)du}{(u-z)}, \quad i^2 = -1, \quad |z| < 1$$

$$(7.12) \quad f(z)_- = \frac{1}{2\pi i} \int_{|u|=1} \frac{f(u)du}{(z-u)}, \quad |z| > 1$$

og

$$(7.13) \quad \mathcal{A}f(z) = \frac{1}{2\pi i} \int_{|u|=1} \frac{f(u)du}{u}.$$

Det er umiddelbart innlysende at en har følgende identiteter

$$(7.14) \quad \mathcal{A}|f(z)|^2 = \mathcal{A}\{f(z)f(z^{-1})\} = \sum_{-\infty}^{\infty} f_j^2,$$

$$(7.15) \quad \mathcal{A}|f(z)|^2 = \mathcal{A}|f(z)_+|^2 + \mathcal{A}|f(z)_-|^2$$

og

$$(7.16) \quad \mathcal{A}\left[\left[\frac{f(z)}{z^\ell}\right]_+\right]^2 = \mathcal{A}|f(z)_+|^2 - \sum_{r=0}^{\ell-1} \left[\mathcal{A}\left[\frac{f(z)}{z^r}\right]\right]^2 (1-\delta_{0\ell}) \quad (\text{Kronecker delta})$$

Setning 7.1: La

$$\hat{x}_t(\ell) = \zeta(\ell)y_t = \sum_{j \geq 0} \zeta_j(\ell)y_{t-j},$$

være den lineære prediktor for $x_{t+\ell}$ ved tidspunkt t som minimerer forventet kvadratavvik. Da gjelder

$$\zeta(\ell, z) = \sum_{j \geq 0} \zeta_j(\ell)z^j = \frac{\sigma^2 \Delta_\theta(z) \Gamma(z)}{\tau^2 \xi_\theta(z)} \left[\frac{\Gamma(z^{-1}) \phi(z) \phi(z^{-1})}{z^\ell \Delta_\theta(z) \xi_\theta(z^{-1})} \right]_+$$

og

$$E\{\hat{x}_t(\ell) - x_{t+\ell}\}^2 = \text{Var}\{x_t\} - \frac{\sigma^4}{\tau^2} \mathcal{A}\left[\frac{\Gamma(z^{-1}) \phi(z) \phi(z^{-1})}{\Delta_\theta(z) \xi_\theta(z^{-1}) z^\ell}\right]_+^2,$$

Disse resultatene følger direkte fra Whittle, 1963, side 68. Ved å ta utgangspunkt i setning 7.1 kan vi finne følgende eksplisitte uttrykk for prediktorene.

Setning 7.2: (i) Dersom perioden $k=4$ er

$$\begin{aligned} \zeta^*(\ell, z) &= \lim_{\theta \rightarrow 1} \zeta(\ell, z) = \frac{\Gamma(z)}{\xi(z)} \left[\frac{\{(7+\ell)\Gamma(1)+\Gamma'(1)\}\xi(1)-\Gamma(1)\xi'(1)}{4\Gamma(1)^2} (1-z^4) \right. \\ &+ \xi(1) \frac{z(1+z)(1+z^2)}{4\Gamma(1)} + \frac{(-1)^\ell \xi(-1)(1-z)^2(1+z^2)}{8\Gamma(1)} \\ &\left. + \frac{i^\ell(1-i)\xi(-i)(1+zi)(1-z)^2(1+z)}{8\Gamma(-i)} + \frac{(-i)^\ell(1+i)\xi(i)(1-iz)(1-z)^2(1+z)}{8\Gamma(i)} \right] \end{aligned}$$

hvor $\xi = \xi_1$.

(ii) Dersom $k=1$ er

$$\zeta^{\bullet}(\ell, z) = \lim_{\theta \uparrow 1} \zeta(\ell, z) = \frac{\Gamma(z)}{\xi(z)} \left[\frac{\xi(1)\Gamma(1) + (1-z)\{\xi(1)\Gamma'(1) - \xi'(1)\Gamma(1) + \ell\Gamma(1)\xi(1)\}}{\Gamma(1)^2} \right].$$

Bevis: Vi skal bare gjennomføre beviset for $k=4$ da tilfellet $k=1$ er helt analogt. Fra (7.11) finner vi ved å ta variabeltransformasjonen $u \rightarrow \frac{1}{u}$

$$\left[\frac{\Gamma(z^{-1})\phi(z)\phi(z^{-1})}{\Delta_{\theta}(z)\xi_{\theta}(z^{-1})z^{\ell}} \right]_{+} = \frac{1}{2\pi i} \int_{|u|=1} \frac{\Gamma(u)\phi(u)\phi(u^{-1})u^{\ell-1}}{(1-uz)\Delta_{\theta}(u^{-1})\xi_{\theta}(u)} du, \quad |z| < 1.$$

Integranden er analytisk i et området $|u| < 1 + \epsilon, \epsilon > 0$, bortsett fra punktene $\theta, -\theta, i\theta$ og $-i\theta$. Ved delbrøkkoppspalting får vi

$$1/u\Delta_{\theta}(u^{-1}) = \frac{7u-6\theta}{4(u-\theta)^2} + \frac{1}{8} \left[\frac{1}{u+\theta} + \frac{1-i}{u-i\theta} + \frac{1+i}{u+i\theta} \right].$$

La $g(u) = \Gamma(u)\phi(u)\phi(u^{-1})u^{\ell}$. Ved å anvende residu-satsen oppnår vi derfor

$$(7.17) \quad \left[\frac{\Gamma(z^{-1})\phi(z)\phi(z^{-1})}{\Delta_{\theta}(z)\xi_{\theta}(z^{-1})z^{\ell}} \right]_{+} = \frac{1}{4} \frac{d}{du} \left\{ \frac{(7u-6\theta)g(u)}{(1-uz)\xi_{\theta}(u)} \right\} \Big|_{u=\theta}$$

$$+ \frac{1}{8} \left[\frac{g(-\theta)}{(1+\theta z)\xi_{\theta}(-\theta)} + \frac{(1-i)g(i\theta)}{(1-i\theta z)\xi_{\theta}(i\theta)} + \frac{(1+i)g(-i\theta)}{(1+i\theta z)\xi_{\theta}(-i\theta)} \right].$$

For $\theta=1$ gir likning (7.8)

$$\xi(1)^2 = \Gamma(1)^2 \phi(1)^2 \sigma^2 / \tau^2,$$

$$\xi(i)\xi(-i) = \Gamma(i)\Gamma(-i)\phi(1)^2 \sigma^2 / \tau^2,$$

$$\xi(-1)^2 = \Gamma(-1)^2 \phi(1)^2 \sigma^2 / \tau^2$$

som innsatt i (7.17) gir

$$\frac{\sigma^2}{\tau^2} \left[\frac{\Gamma(z^{-1})\phi(z)\phi(z^{-1})}{\Delta(z)\xi(z^{-1})z^{\ell}} \right]_{+} = \frac{\{(7+\ell)\Gamma(1)+\Gamma'(1)\}\xi(1)-\Gamma(1)\xi'(1)}{4\Gamma(1)^2(1-z)}$$

$$+ \frac{z\xi(1)}{4(1-z)^2\Gamma(1)} + \frac{1}{8} \left[\frac{\xi(-1)(-1)^{\ell}}{(1+z)\Gamma(-1)} + \frac{(1-i)\xi(-i)(i)^{\ell}}{(1-iz)\Gamma(-i)} \right]$$

$$+ \frac{(1+i)\xi(i)(-i)^{\ell}}{(1+iz)\Gamma(i)} \quad \text{når } \theta=1.$$

Setningen følger dermed av setning 7.1.

Lemma 7.3. La $K(z) = \psi(z)\psi(z^{-1}) = G_{\nu}(z)/h^2$ og la $\lambda_j, j = 1, 2, \dots, p$ være røttene i $\Gamma(z^{-1})$. Dersom røttene er forskjellige er

$$\left[\frac{\Delta(z^{-1})K(z)}{\Gamma(z)(z^{-1})} \right]_{+} = \sum_{j=1}^p \frac{\lambda_j \Delta(\lambda_j) K(\lambda_j)}{(1-\lambda_j z) \xi(\lambda_j) \Gamma'(\lambda_j^{-1})} + R_{q-p}(z)$$

hvor

$$(q-p)! R_{q-p}(z) = \begin{cases} \frac{d^{(q-p)}}{du^{(q-p)}} \left\{ \frac{\Delta(u)K(u)u^q}{(1-uz)\Gamma(u^{-1})\xi(u)u^p} \right\} \Big|_{u=0}, & p \leq q \\ 0, & p > q \end{cases}$$

Bevis: Under forutsetningen om at røttene λ_j er forskjellige er

$$\frac{1}{u\Gamma(u^{-1})} = - \sum_{j=1}^p \frac{\lambda_j}{(u-\lambda_j)\Gamma'(\lambda_j^{-1})}.$$

Ved å ta variabeltransformasjonen $u \rightarrow u^{-1}$ får vi av (7.11)

$$\left[\frac{\Delta(z^{-1})K(z)}{\Gamma(z)\xi(z^{-1})} \right]_{+} = \frac{1}{2\pi i} \int_{|u|=1} \frac{\Delta(u)K(u)du}{(1-uz)u\Gamma(u^{-1})\xi(u)}.$$

Integranden har førsteordens poler i λ_j og en pol av orden $q+1-p$ i 0. Av delbrøkkoppspaltningen ovenfor ser vi at residuet i λ_j er lik

$$-\lambda_j \Delta(\lambda_j)K(\lambda_j) / \xi(\lambda_j)\Gamma'(\lambda_j)$$

og beviset er dermed fullført.

Lemma 7.4:

$$\begin{aligned} E\{\hat{x}_t^{(y)} - x_{t+\ell}\}^2 &= s^2 - \frac{h^4}{\tau^2} \mathcal{A} \left| \left[\frac{\Delta(z^{-1})K(z)}{\Gamma(z)\xi(z^{-1})} \right]_{+} \right|^2 \\ &+ \frac{\sigma^2}{\tau^2} \sum_{r=0}^{\ell-1} \left[\lim_{\theta \rightarrow 1} \mathcal{A} \left(\frac{\Gamma(z^{-1})\phi(z)\phi(z^{-1})}{z^r \xi_{\theta}(z)\xi_{\theta}(z^{-1})} \right) \right]^2 (1 - \delta_{\ell r}), \end{aligned}$$

hvor $s^2 = \text{Var}\{v_t\}$ og δ_{ij} er Kroneckers delta.

Bevis: Ved å bruke (7.15) og (7.8) får vi

$$\begin{aligned} (7.18) \quad \mathcal{A} \left| \left[\frac{\Gamma(z^{-1})\phi(z)\phi(z^{-1})}{\Delta_{\theta}(z)\xi_{\theta}(z^{-1})} \right]_{+} \right|^2 &= \frac{\tau^2}{\sigma^2} \mathcal{A} \left\{ \left| \frac{\phi(z)}{\Delta_{\theta}(z)} \right|^2 \right\} \\ &- \frac{h^2}{\sigma^2} \mathcal{A} \left\{ \left| \frac{\phi(z)^2 K(z)}{\xi_{\theta}(z)^2} \right| \right\} - \mathcal{A} \left| \left[\frac{\tau^2 \xi_{\theta}(z)}{\sigma^2 \Gamma(z)\Delta_{\theta}(z)} - \frac{h^2 \Delta_{\theta}(z^{-1}) K(z)}{\sigma^2 \Gamma(z)\xi_{\theta}(z)} \right]_{-} \right|^2. \end{aligned}$$

Siden $\left[\xi_{\theta}(z)/\Delta_{\theta}(z)\Gamma(z) \right]_{-} = 0$ blir det siste uttrykket lik

$$\begin{aligned} \frac{\tau^2}{\sigma^2} \mathcal{A} \left| \frac{\phi(z)}{\Delta_{\theta}(z)} \right|^2 - \frac{h^2}{\sigma^2} \mathcal{A} \left| \frac{\phi(z)^2 K(z)}{\xi_{\theta}(z)^2} \right| \\ - \frac{h^4}{\sigma^4} \mathcal{A} \left\{ \left| \frac{\Delta_{\theta}(z)^2 K(z)}{\Gamma(z)^2 \xi_{\theta}(z)^2} \right| \right\} + \frac{h^4}{\sigma^4} \mathcal{A} \left| \left[\frac{\Delta_{\theta}(z^{-1}) K(z)}{\Gamma(z)\xi_{\theta}(z^{-1})} \right]_{+} \right|^2. \end{aligned}$$

På grunn av (7.8) er

$$\left| \frac{\phi(z)^2 K(z)}{\xi_{\theta}(z)^2} \right| + \frac{h^2}{\sigma^2} \left| \frac{\Delta_{\theta}(z) K(z)}{\Gamma(z)\xi_{\theta}(z)} \right|^2 = \frac{\tau^2}{\sigma^2} \left| \frac{K(z)}{\Gamma(z)} \right|^2.$$

Videre er

$$\text{Var}\{x_t\} = \mathcal{A}_{G_x}(z) = \sigma^2 \mathcal{A} \left| \frac{\phi(z)}{\Delta_{\theta}(z)} \right|^2$$

og

$$s^2 = \text{Var}\{v_t\} = h^2 \mathcal{A} \left| \frac{\psi(z)}{\Gamma(z)^2} \right|$$

slik at høyresiden av (7.16) reduserer seg til

$$\frac{\tau^2}{\sigma^4} \text{Var}\{x_t\} = \frac{s^2}{\sigma^4} + \frac{h^4}{\sigma^4} \mathcal{A} \left| \left[\frac{\Delta_\theta(z^{-1})K(z)}{\Gamma(z)\xi_\theta(z^{-1})} \right]_+ \right|^2.$$

Benytter vi nå setning (7.1) sammen med (7.16) får vi det ønskede resultatet.

Vi er nå istand til å vise hovedresultatet om forventet kvadratavvik.

Setning 7.5: La

$$v_j = -\lambda_j \xi(\lambda_j^{-1}) / \Delta(\lambda_j^{-1}) \Gamma'(\lambda_j^{-1}) \text{ og } v_0 = 0.$$

Forventet kvadratavvik for $\hat{x}_t(\ell)$ er gitt ved

$$\begin{aligned} E\{\hat{x}_t(\ell) - x_{t+\ell}\}^2 &= s^2 - \tau^2 \sum_{r,j} \frac{v_j v_r}{1-\lambda_j \lambda_r} - 2h^2 \sum_r v_r R_{q-p}(\lambda_r) \\ &- \frac{h^4}{\tau^2} \mathcal{A}\{R_{q-p}(z) R_{q-p}(z^{-1})\} + \tau^2 \sum_{r=0}^{\ell-1} \zeta^{\mathbf{x}}(r,0)^2 (1-\delta_{0\ell}) \end{aligned}$$

for $\ell \geq 0$.

Bevis: Lemma 7.3 gir

$$\begin{aligned} \mathcal{A} \left| \left[\frac{\Delta(z^{-1})K(z)}{\Gamma(z)\xi(z^{-1})} \right]_+ \right|^2 &= \mathcal{A} \left\{ \sum_{r,j} \frac{\tau^4 v_j v_r}{(1-\lambda_j z)(1-\lambda_r z^{-1})h^4} \right. \\ &\left. + 2 \sum_r \frac{\tau^2 v_r R_{q-p}(z)}{(1-\lambda_r z^{-1})\sigma^2} + |R_{q-p}(z)|^2 \right\} \end{aligned}$$

som sammen med (7.13) og lemma 7.5 gir resultatet ovenfor.

8. Spesielle modeller

La $K_2(z)$ være den genererende funksjon for autokovariansen til utvalgsfeilen innen utvalgsområdene. I følge rotasjonsplanen i AKU er autokovariansfunksjonen innen utvalgsområdene lik null for argumenter større enn 5, slik at $K_2(z)$ derfor blir et polynom i z og z^{-1} av femte grad. I samsvar med avsnitt 6 har vi at utvalgsfeilen mellom utvalgsområdene genereres av modellen

$$\Gamma v_{1t} = v_{1t} - \lambda v_{1t-1} - \mu v_{1t-4} + \lambda \mu v_{1t-5} = b_{1t}$$

hvor var $\{b_{1t}\} = h_1^2$. Den genererende funksjon for autokorrelasjonsfunksjonen til $\{v_{1t}\}$ blir følgende

$$1/\Gamma(z)\Gamma(z^{-1})$$

slik at

$$h^2 K(z) = h_1^2 + \Gamma(z) \Gamma(z^{-1}) K_2(z)$$

som i følge (7.8) medfører at $\xi(z)$ blir av grad 15.

La

$$\xi(z) = 1 + \sum_{j=1}^{15} v_j z^j,$$

og definer

$$\xi^{**}(z) = 1 + \sum_{j=1}^7 v_j z^{2j} (-1)^j,$$

$$\xi^{***}(z) = - \sum_{j=1}^8 v_j z^{2j-1} (-1)^{j+1},$$

som medfører at

$$\xi(iz) = \xi^{**}(z) - i \xi^{***}(z).$$

Ved bruk av setning 7.2 får vi etter litt regning

$$(8.1) \quad \zeta^{**}(0, z) = \frac{\Gamma(z)}{\xi(z)} \left[\frac{\{7\Gamma(1) + \Gamma'(1)\} \xi(1) - \Gamma(1) \xi'(1)}{4\Gamma(1)^2} (1-z^4) \right. \\ \left. + \frac{\xi(1)z(1+z)(1+z^2)}{4\Gamma(1)} + \frac{\xi(-1)(1-z)^2(1+z^2)}{8\Gamma(1)} \right. \\ \left. + \frac{[\xi^{**}(1) + z(\xi^{**}(1) + \xi^{***}(1))] (1-z)^2 (1+z)}{4(1+\lambda^2)(1-\mu)} \right].$$

Den beste lineære estimator er følgelig av typen

$$(8.2) \quad \xi \hat{x}_t(0)' = \Gamma P y_t$$

hvor P er en operator av fjerde grad.

For å bestemme forventet kvadratavvik må vi i følge setning 7.5 finne V_j og $R_{q-p}(z)$. Røttene til $\Gamma(z^{-1})$ er $\lambda, \alpha, -\alpha, i\alpha$ og $-i\alpha$ hvor $\alpha = \mu^{1/4}$ slik at

$$V_1 = \frac{\lambda^9 \xi(\lambda^{-1})}{(1-\lambda)(1-\lambda^4)(\lambda^4-\mu)}, \quad V_2 = \frac{\mu^2 \xi(\alpha^{-1})}{4(\alpha-\lambda)(1-\alpha)(1-\mu)}$$

$$V_3 = - \frac{\mu^2 \xi(-\alpha^{-1})}{4(\lambda+\alpha)(1+\alpha)(1-\mu)}$$

$$V_4 = - \frac{\mu^2 [\xi^{**}(\alpha^{-1}) - i \xi^{***}(\alpha^{-1})] (\lambda - \alpha^2 + i\alpha(1+\lambda))}{4(1-\mu)(1+\alpha^2)(\lambda^2 + \alpha^2)}$$

og

$$V_5 = - \frac{\mu^2 [\xi^{**}(\alpha^{-1}) + i \xi^{***}(\alpha^{-1})] (\lambda - \alpha^2 - i\alpha(1+\lambda))}{4(1-\mu)(1+\alpha^2)(\lambda^2 + \alpha^2)}.$$

Det gjenstår nå å finne $R_{q-p}(z)$. Siden $q-p = 5$ må vi ifølge lemma 7.3 finne den femte-deriverte i $u=0$ av uttrykket

$$\frac{\Delta(u) K(z) u^{10}}{(1-uz) \Gamma(u-1) u^5 \xi(u)}.$$

La

$$u^{10} K(u) = \sum_{j=0}^{10} T_j u^j,$$

$$A_1 = (\lambda v_1 - 1) \lambda^{-1}, A_2 = (\lambda v_2 - v_1) \lambda^{-1}, A_3 = (\lambda v_3 - v_2) \lambda^{-1}, A_4 = (\mu \lambda v_4 - \mu v_3 - \lambda) \lambda^{-1} \mu^{-1},$$

$$A_5 = (v_5 \lambda v - v_4 \mu - v_1 \lambda + 1) \lambda^{-1} \mu^{-1},$$

$$B_1 = -A_1, B_2 = A_1^2 - A_2, B_3 = 2A_1 A_2 - A_1^3 - A_3, B_4 = 2A_1 A_3 + A_1^4 + A_2^2 - 3A_1^2 A_2 - A_4$$

og

$$B_5 = 2A_1 A_4 + 2A_2 A_3 + 4A_1^3 A_2 - 3A_1^2 A_3 - 3A_1 A_2^2 - A_1^5 - A_5.$$

La videre

$$C_0 = T_5 - T_4 - T_1 + T_0 + B_1 (T_4 - T_3 - T_0) + B_2 (T_3 - T_2) + B_3 (T_2 - T_1)$$

$$+ B_4 (T_1 - T_0) + B_5 T_0,$$

$$C_1 = T_4 - T_3 - T_0 + B_1 (T_3 - T_2) + B_2 (T_2 - T_1) + B_3 (T_1 - T_0) + B_4 T_0,$$

$$C_2 = T_3 - T_2 + B_1 (T_2 - T_1) + B_2 (T_1 - T_0) + B_3 T_0,$$

$$C_3 = T_2 - T_1 + B_1 (T_1 - T_0) + B_2 T_0,$$

$$C_4 = T_1 - T_0 + B_1 T_0$$

og

$$C_5 = T_0.$$

Etter en del regning som vi ikke tar med her finner vi

$$(8.3) \quad R_5(z) = \lambda^{-1} \mu^{-1} \sum_{j=0}^5 C_j z^j$$

som gir

$$\mathcal{A}\{R_5(z) R_5(z^{-1})\} = \lambda^{-2} \mu^{-2} \sum_{j=0}^5 C_j^2.$$

Vi har dermed funnet eksplisitt alle de uttrykkene som inngår i formelen for forventet kvadratavvik for $\hat{x}_t(0)$.

Til slutt skal vi se på situasjonen når perioden $k=1$ og

$$\Gamma v_{1t} = v_{1t} - \lambda v_{1,t-1} = b_{1t}.$$

Setning 7.2 gir umiddelbart $\zeta^{\mathbf{x}}(\ell, z)$ og vi finner forventet kvadratavvik fra setning 7.5:

$$E\{\hat{x}_t(\ell) - x_{t+\ell}\}^2 = s^2 - \frac{\tau^2 V_1^2}{1-\lambda^2} - 2h^2 V_1 R_0(\lambda) - \frac{h^4}{\tau^2} \mathcal{A}\{R_0(z) R_0(z^{-1})\} + \tau^2 \sum_{r=0}^{\ell-1} \zeta^{\mathbf{x}}(r, 0)^2 (1-\delta_{o\ell})$$

der $V_1 = \xi(\lambda^{-1})/\Delta(\lambda^{-1})$.

Nå er

$$s^2 = h^2 \mathcal{A}\left\{\frac{K(z)}{\Gamma(z)\Gamma(z-1)}\right\} = \frac{K(\lambda)\lambda}{1-\lambda^2} h^2,$$

$$\tau^2 \xi(\lambda)\xi(\lambda^{-1}) = h^2 \Delta(\lambda)\Delta(\lambda^{-1}) K(\lambda)$$

og

$$R_0(z) = -k/\lambda$$

der $k = \{zK(z)\}_{z=0}$.

Herav

$$(8.4) \quad E\{\hat{x}_t(\ell) - x_{t+\ell}\}^2 = s^2 \left\{1 + \frac{2k\xi(\lambda^{-1})}{K(\lambda)} - \frac{\lambda\xi(\lambda^{-1})}{\xi(\lambda)} - \frac{\xi(\lambda)\xi(\lambda^{-1})(1+\lambda)}{(1-\lambda)^2 \lambda K(\lambda)^2}\right\} + \tau^2 \sum_{r=0}^{\ell-1} \zeta^{\mathbf{x}}(r, 0)^2 (1-\delta_{o\ell}).$$

Referanser

- Anderson, T. W. 1971: The Statistical Analysis of Time Series. Wiley, New York.
- Blight, B. J. N. and Scott, A. J., 1973: A stochastic model for repeated surveys. J. R. Statist. Soc. B 1, 61-66.
- Box, G. E. P. and Jenkins, G. M. 1970: Time Series Analysis, Forecasting and Control. Holden-Day, Inc. San Francisco.
- Dagsvik, J. 1974: Presisjonsgevinst ved bruk av sammensatt estimering i Byråets arbeidskraftundersøkelser. Statistisk Sentralbyrå, Arbeidsnotat (10 75/24).
- Doob, J. L. 1953: Stochastic Processes. Wiley, New York.
- Scott, A. J. and Smith, T. M. F. 1974: Analysis of repeated surveys using time series methods. JASA 69, 674-678.
- Whittle, P. 1963: Prediction and Regulation by Linear Least-Square Methods. English Universities Press Ltd, London.
- Wilson, G. J. 1969: Factorization of the generating function of a pure moving average process. SIAM J. Num. Analysis, 6,1.

Autokorrelasjonsestimater for sysselsatte pr. måned i perioden 1955-1970 for næringene: Industri, Bygg og anlegg, Offentlig, sosial og privat tjenesteyting, Varehandel, hotell og restaurantvirksomhet.

Tabell 1. Sysselsatte i industri x_t

Autokorrelasjon ρ_r	r	ρ_r for x_t	ρ_r for $x_t - x_{t-1}$	ρ_r for $x_t - x_{t-12}$	ρ_r for $x_t - x_{t-1} - x_{t-12} + x_{t-13}$
	1	0.965	0.079	0.953	-0.116
	2	0.925	-0.451	0.914	0.082
	3	0.904	0.043	0.869	0.117
	4	0.882	0.412	0.812	0.098
	5	0.844	-0.099	0.747	0.245
	6	0.809	-0.653	0.661	0.044
	7	0.802	-0.163	0.571	-0.102
	8	0.805	0.392	0.490	0.103
	9	0.792	0.075	0.398	0.115
	10	0.776	-0.458	0.298	-0.116
	11	0.780	0.059	0.209	0.148
	12	0.781	0.822	0.108	-0.404
	13	0.748	0.126	0.042	0.007
	14	0.709	-0.448	-0.023	0.028
	15	0.689	-0.021	-0.090	-0.153
	16	0.670	0.366	-0.143	-0.134
	17	0.638	-0.091	-0.184	-0.078
	18	0.608	-0.626	-0.217	-0.026
	19	0.602	-0.201	-0.247	-0.043
	20	0.609	0.340	-0.271	-0.101
	21	0.604	0.082	-0.284	-0.129
	22	0.595	-0.405	-0.285	0.136
	23	0.604	-0.008	-0.296	-0.079
	24	0.612	0.753	-0.299	-0.057
	25	0.588	0.168	-0.299	0.071
	26	0.556	-0.426	-0.304	-0.139
	27	0.541	-0.049	-0.297	0.097
	28	0.527	0.344	-0.298	-0.005
	29	0.503	-0.057	-0.298	-0.031
	30	0.479	-0.580	-0.296	-0.036
	31	0.478	-0.213	-0.293	-0.003
	32	0.489	0.317	-0.290	0.013
	33	0.486	0.108	-0.288	0.062
	34	0.478	-0.381	-0.293	-0.215
	35	0.485	-0.040	-0.279	-0.085
	36	0.491	0.696	-0.260	0.098
	37	0.467	0.167	-0.249	-0.189
	38	0.434	-0.388	-0.221	0.029
	39	0.416	-0.092	-0.196	-0.108
	40	0.400	0.332	-0.164	0.004
	41	0.373	-0.038	-0.133	0.027
	42	0.344	-0.544	-0.105	-0.053
	43	0.338	-0.214	-0.070	-0.075
	44	0.343	0.293	-0.030	0.039
	45	0.335	0.124	0.004	-0.076
	46	0.322	-0.321	0.048	0.116
	47	0.321	-0.046	0.080	0.054
	48	0.321	0.607	0.106	-0.098

Tabell 2. Sysselsatte i Bygg og anlegg x_t

Autokorrelasjon ρ_r	r	ρ_r for x_t	ρ_r for $x_t - x_{t-1}$	ρ_r for $x_t - x_{t-12}$	ρ_r for $x_t - x_{t-1} - x_{t-12} + x_{t-13}$
	1	0.846	0.404	0.783	-0.104
	2	0.567	0.020	0.599	-0.168
	3	0.282	-0.070	0.487	-0.094
	4	0.015	-0.123	0.429	0.029
	5	-0.212	-0.378	0.364	0.093
	6	-0.321	-0.654	0.260	-0.104
	7	-0.230	-0.368	0.147	-0.023
	8	-0.025	-0.106	0.066	0.021
	9	0.216	-0.051	0.001	0.084
	10	0.471	0.040	-0.091	0.034
	11	0.710	0.408	-0.201	0.141
	12	0.818	0.812	-0.346	-0.455
	13	0.678	0.358	-0.329	0.019
	14	0.426	0.018	-0.300	-0.006
	15	0.168	-0.051	-0.272	0.011
	16	-0.073	-0.108	-0.253	0.051
	17	-0.279	-0.364	-0.244	-0.154
	18	-0.370	-0.593	-0.179	0.163
	19	-0.289	-0.371	-0.129	0.061
	20	-0.088	-0.083	-0.122	-0.015
	21	0.146	-0.035	-0.132	-0.113
	22	0.391	0.064	-0.101	0.012
	23	0.612	0.375	-0.070	0.005
	24	0.714	0.755	-0.045	-0.028
	25	0.586	0.334	-0.017	-0.019
	26	0.354	0.022	0.014	0.124
	27	0.114	-0.065	0.001	0.066
	28	-0.107	-0.114	-0.026	-0.099
	29	-0.289	-0.300	-0.034	0.062
	30	-0.376	-0.562	-0.059	-0.062
	31	-0.291	-0.348	-0.059	-0.016
	32	-0.097	-0.070	-0.056	-0.071
	33	0.121	-0.011	-0.026	0.017
	34	0.342	0.070	-0.003	0.108
	35	0.539	0.333	-0.009	-0.024
	36	0.628	0.694	-0.006	0.092
	37	0.503	0.285	-0.027	-0.033
	38	0.288	-0.009	-0.020	-0.058
	39	0.075	-0.063	0.010	-0.014
	40	-0.117	-0.082	0.032	-0.012
	41	-0.282	-0.267	0.056	-0.013
	42	-0.362	-0.508	0.073	0.015
	43	-0.287	-0.319	0.078	-0.035
	44	-0.110	-0.039	0.095	0.069
	45	0.080	0.001	0.083	0.051
	46	0.270	0.044	0.053	-0.068
	47	0.445	0.315	0.042	-0.005
	48	0.519	0.621	0.035	-0.127

Tabell 3. Sysselsatte i Varehandel, hotell og restaurant x_t

Autokorrelasjon ρ_r	r	ρ_r for x_t	ρ_r for $x_t - x_{t-1}$	ρ_r for $x_t - x_{t-12}$	ρ_r for $x_t - x_{t-1} - x_{t-12} + x_{t-13}$
	1	0.973	0.082	0.901	-0.182
	2	0.944	-0.479	0.835	-0.000
	3	0.926	-0.224	0.770	0.065
	4	0.912	0.177	0.693	0.114
	5	0.894	0.102	0.593	0.027
	6	0.871	-0.283	0.489	-0.101
	7	0.859	0.102	0.404	0.054
	8	0.847	0.161	0.307	0.006
	9	0.830	-0.206	0.210	-0.069
	10	0.817	-0.464	0.126	-0.055
	11	0.814	0.085	0.054	0.048
	12	0.810	0.913	-0.029	-0.457
	13	0.784	0.081	-0.021	0.102
	14	0.756	-0.449	-0.033	-0.042
	15	0.739	-0.214	-0.036	-0.005
	16	0.727	0.170	-0.037	-0.099
	17	0.709	0.097	-0.018	-0.051
	18	0.686	-0.265	0.010	0.246
	19	0.675	0.089	-0.007	-0.194
	20	0.664	0.147	0.014	-0.028
	21	0.649	-0.187	0.039	0.081
	22	0.637	-0.430	0.049	0.027
	23	0.634	0.076	0.052	-0.041
	24	0.631	0.848	0.064	0.172
	25	0.606	0.076	0.037	-0.007
	26	0.580	-0.412	0.009	0.008
	27	0.565	-0.202	-0.020	-0.001
	28	0.553	0.165	-0.051	-0.018
	29	0.536	0.095	-0.079	0.102
	30	0.515	-0.255	-0.127	-0.243
	31	0.505	0.087	-0.130	0.126
	32	0.495	0.132	-0.158	0.007
	33	0.480	-0.172	-0.188	-0.046
	34	0.468	-0.401	-0.211	-0.105
	35	0.466	0.073	-0.212	0.056
	36	0.462	0.776	-0.226	-0.018
	37	0.438	0.069	-0.230	-0.084
	38	0.413	-0.377	-0.216	0.019
	39	0.397	-0.196	-0.204	0.025
	40	0.384	0.162	-0.198	-0.003
	41	0.368	0.076	-0.189	-0.056
	42	0.347	-0.232	-0.168	0.155
	43	0.337	0.088	-0.175	-0.074
	44	0.327	0.117	-0.167	0.024
	45	0.312	-0.156	-0.164	-0.042
	46	0.300	-0.369	-0.149	0.007
	47	0.297	0.068	-0.138	0.024
	48	0.292	0.699	-0.130	-0.034

Tabell 4. Sysselsatte i Off. sosial og privat tjenesteyting x_t

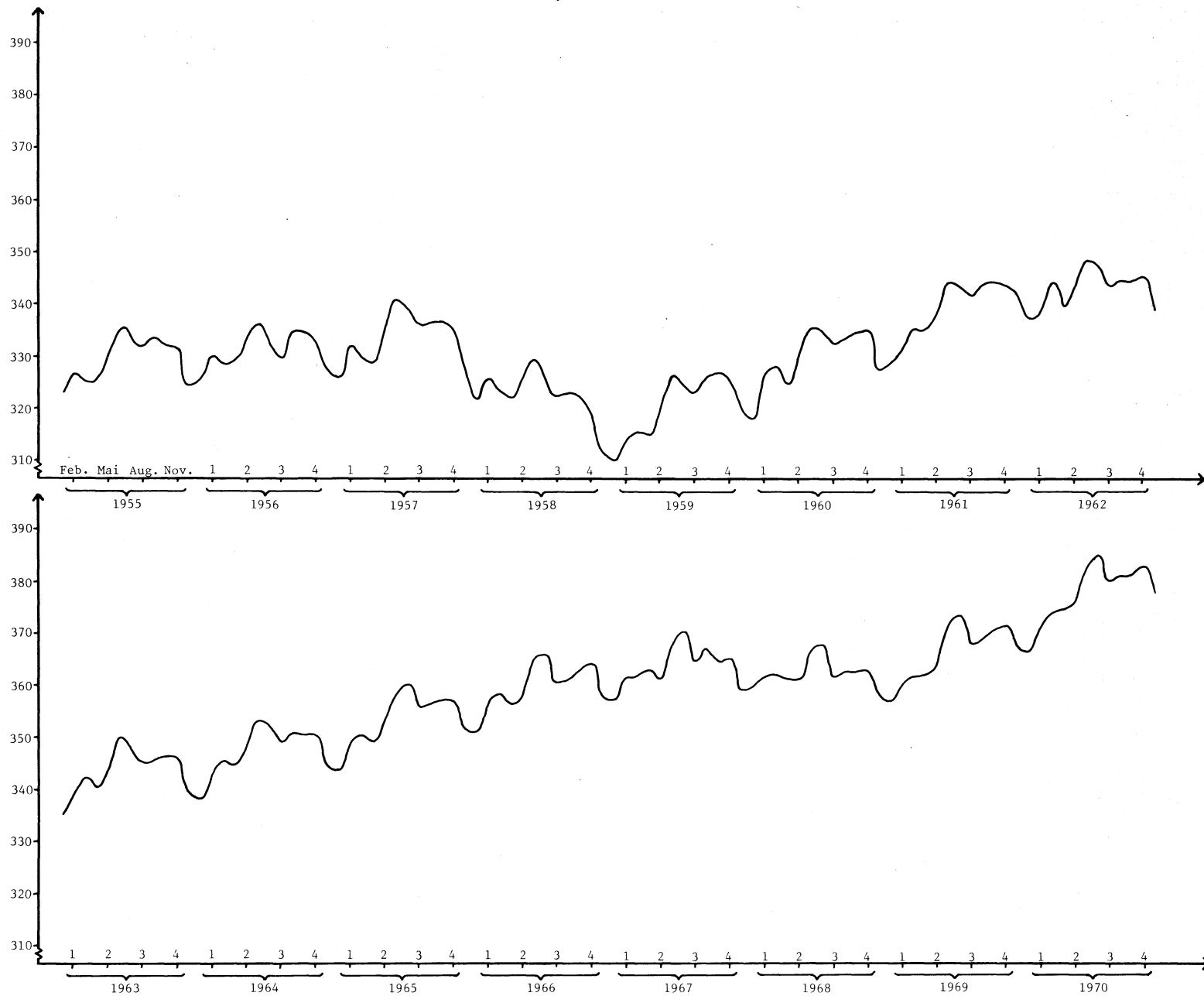
Autokorrelasjon ρ_r	r	ρ_r for x_t	ρ_r for $x_t - x_{t-1}$	ρ_r for $x_t - x_{t-12}$	ρ_r for $x_t - x_{t-1} - x_{t-12} + x_{t-13}$
	1	0.980	0.124	0.951	-0.282
	2	0.959	-0.395	0.916	0.071
	3	0.939	-0.357	0.877	-0.090
	4	0.923	0.227	0.850	-0.081
	5	0.908	0.084	0.828	-0.005
	6	0.891	-0.100	0.807	-0.082
	7	0.874	0.107	0.785	0.077
	8	0.856	0.240	0.758	0.067
	9	0.837	-0.316	0.727	0.097
	10	0.820	-0.400	0.690	-0.112
	11	0.805	0.127	0.663	0.202
	12	0.790	0.832	0.623	-0.528
	13	0.772	0.143	0.616	0.094
	14	0.753	-0.357	0.602	0.029
	15	0.736	-0.347	0.586	0.018
	16	0.721	0.215	0.566	0.165
	17	0.707	0.068	0.535	-0.018
	18	0.692	-0.078	0.497	0.160
	19	0.675	0.078	0.462	-0.117
	20	0.659	0.209	0.435	-0.068
	21	0.641	-0.297	0.416	-0.079
	22	0.625	-0.369	0.404	0.073
	23	0.611	0.099	0.388	-0.050
	24	0.597	0.784	0.378	0.113
	25	0.580	0.152	0.356	0.074
	26	0.561	-0.345	0.328	-0.031
	27	0.545	-0.350	0.302	-0.053
	28	0.531	0.161	0.282	-0.102
	29	0.519	0.059	0.267	-0.008
	30	0.506	-0.084	0.261	-0.129
	31	0.491	0.084	0.250	0.162
	32	0.476	0.202	0.232	-0.042
	33	0.460	-0.262	0.214	0.136
	34	0.445	-0.351	0.185	-0.037
	35	0.432	0.095	0.162	-0.027
	36	0.420	0.717	0.141	0.000
	37	0.404	0.107	0.128	-0.043
	38	0.388	-0.321	0.120	-0.028
	39	0.373	-0.328	0.116	0.089
	40	0.360	0.164	0.105	-0.041
	41	0.349	0.046	0.096	0.064
	42	0.337	-0.067	0.081	0.101
	43	0.324	0.052	0.063	-0.162
	44	0.310	0.187	0.060	0.079
	45	0.296	-0.261	0.051	-0.111
	46	0.283	-0.333	0.060	-0.002
	47	0.272	0.103	0.060	0.103
	48	0.261	0.654	0.053	-0.059

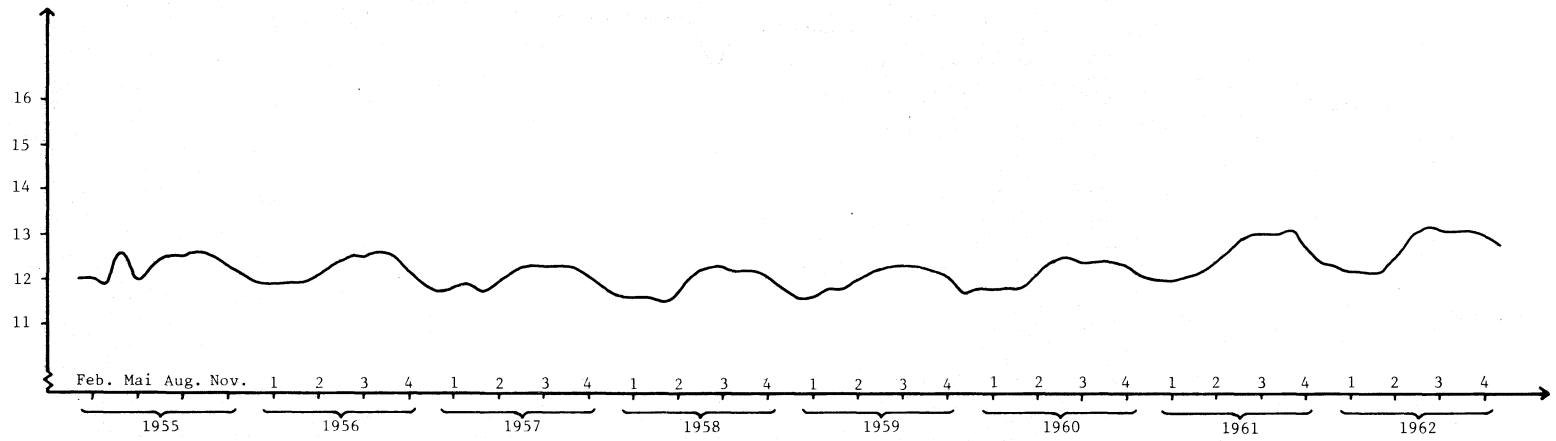
Tabell 5. Variansestimater for det tilfeldige sjokk a_t

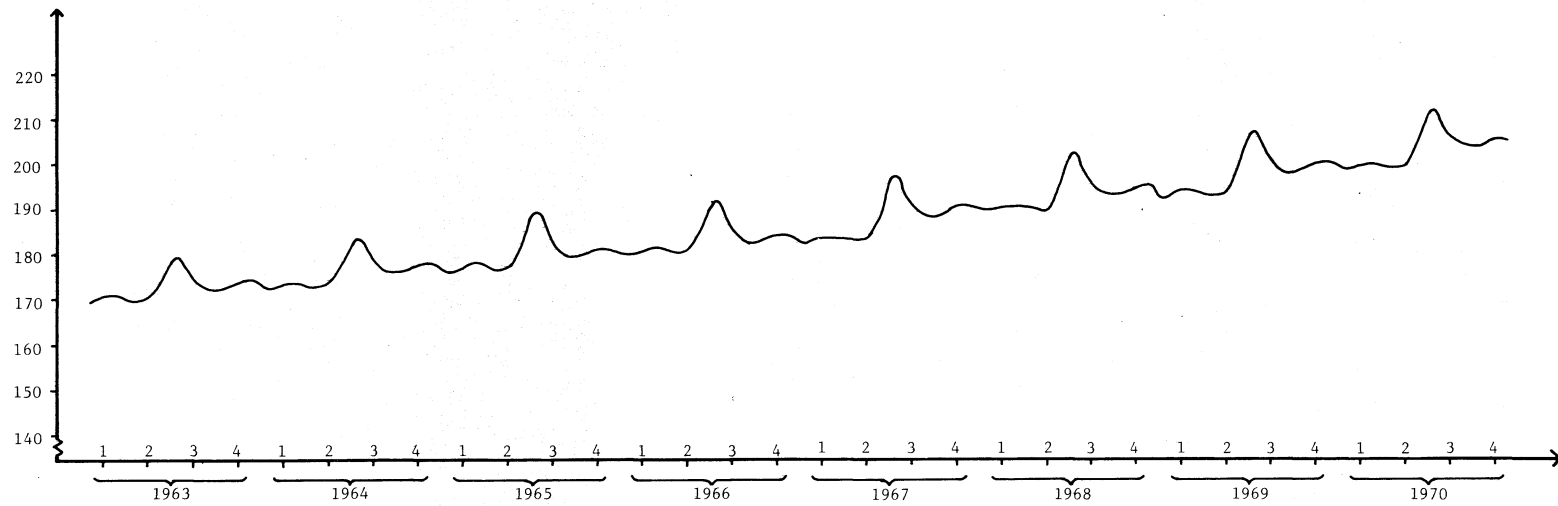
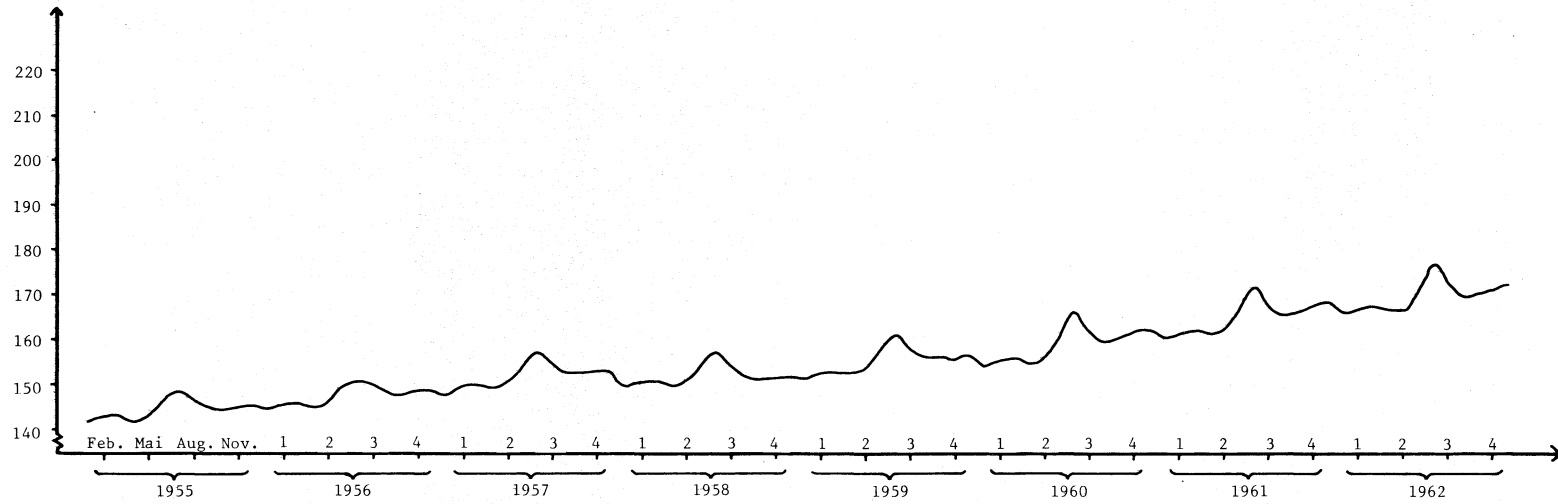
Industri	$2,5 \cdot 10^6$
Bygg og anlegg	$3,6 \cdot 10^6$
Varehandel, hotell og restaurant	$0,4 \cdot 10^6$
Off. sosial og privat tjenesteyting	$0,5 \cdot 10^6$

Sysselsatte i industri

Fig. 1

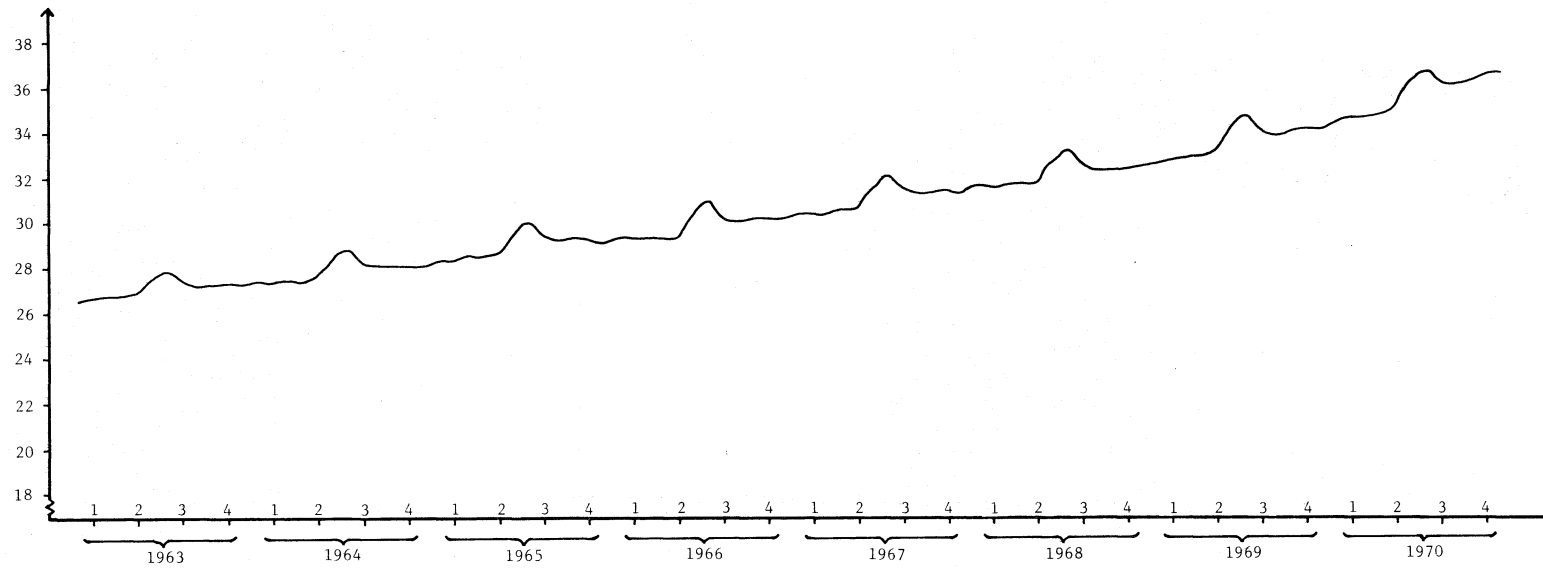
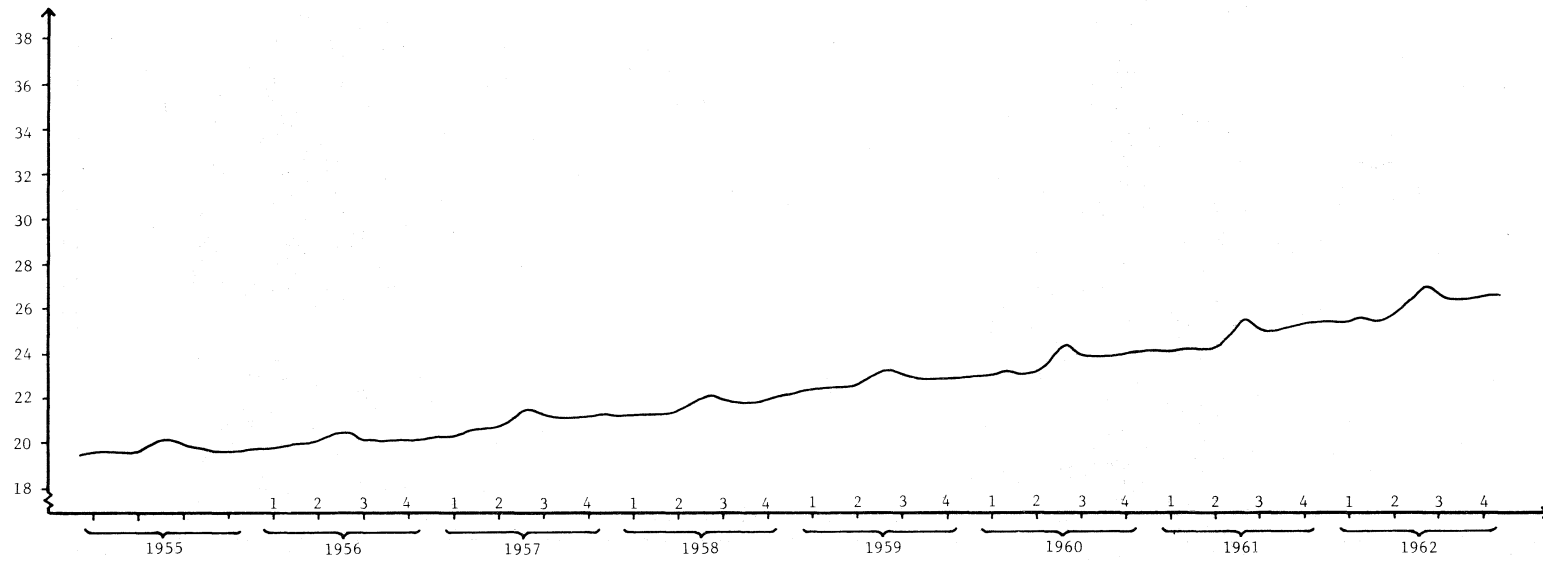


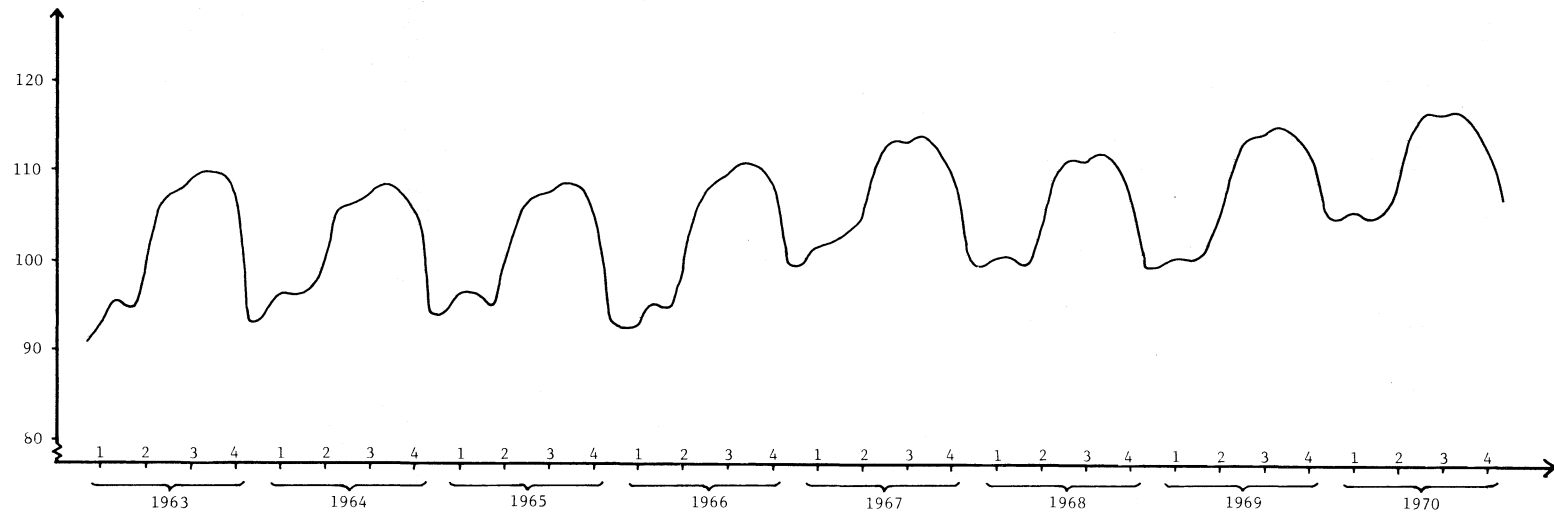
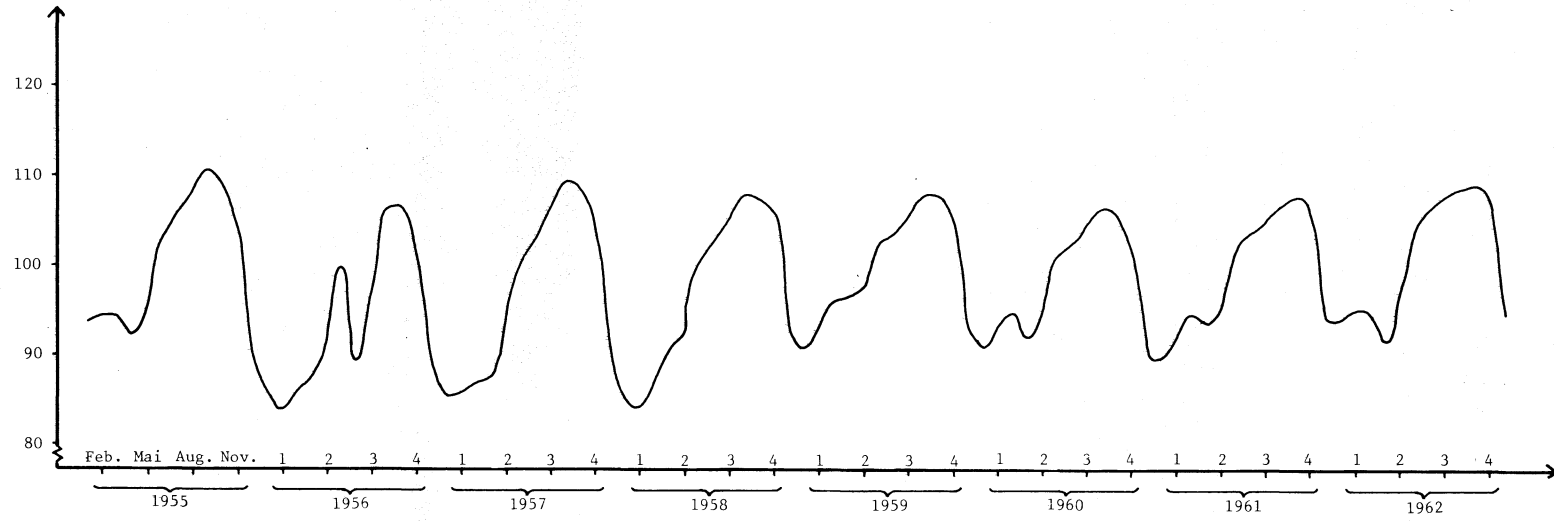




Sysselsatte i bank- og finansieringsvirksomhet,
forsikringsvirksomhet, eiendomsdrift og
forretningsmessig tjenesteyting

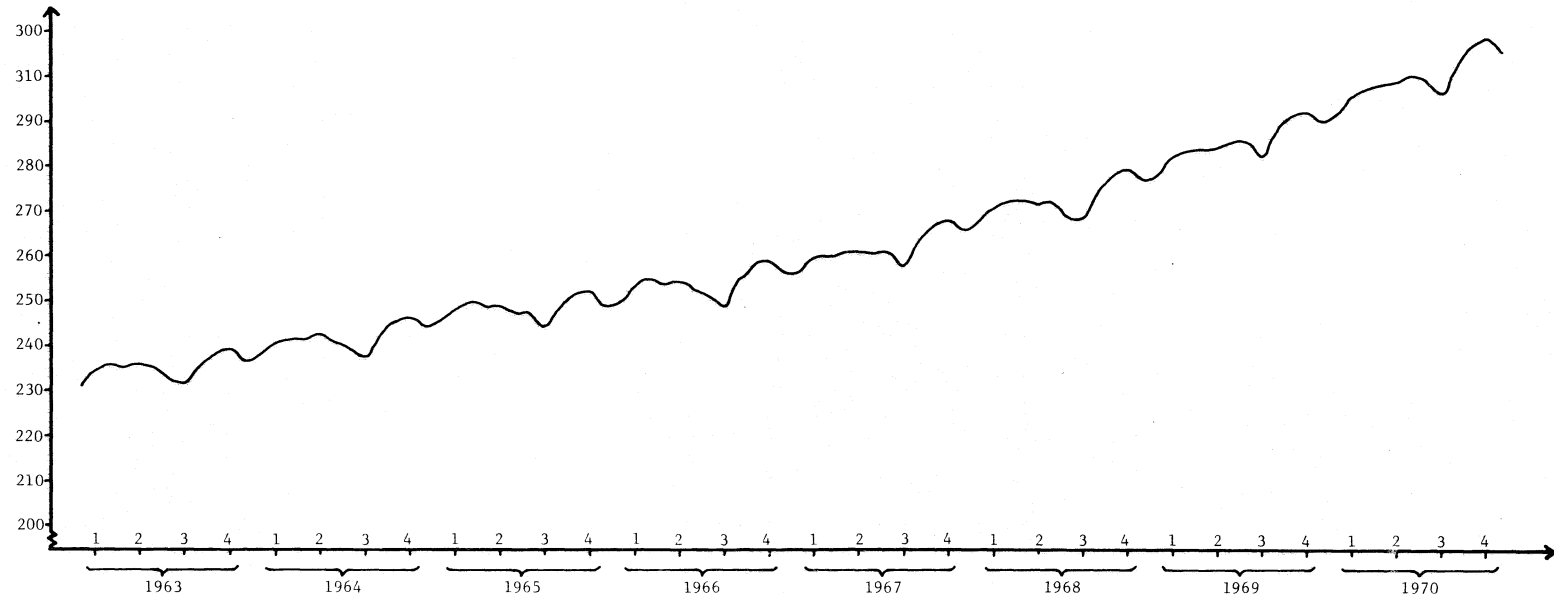
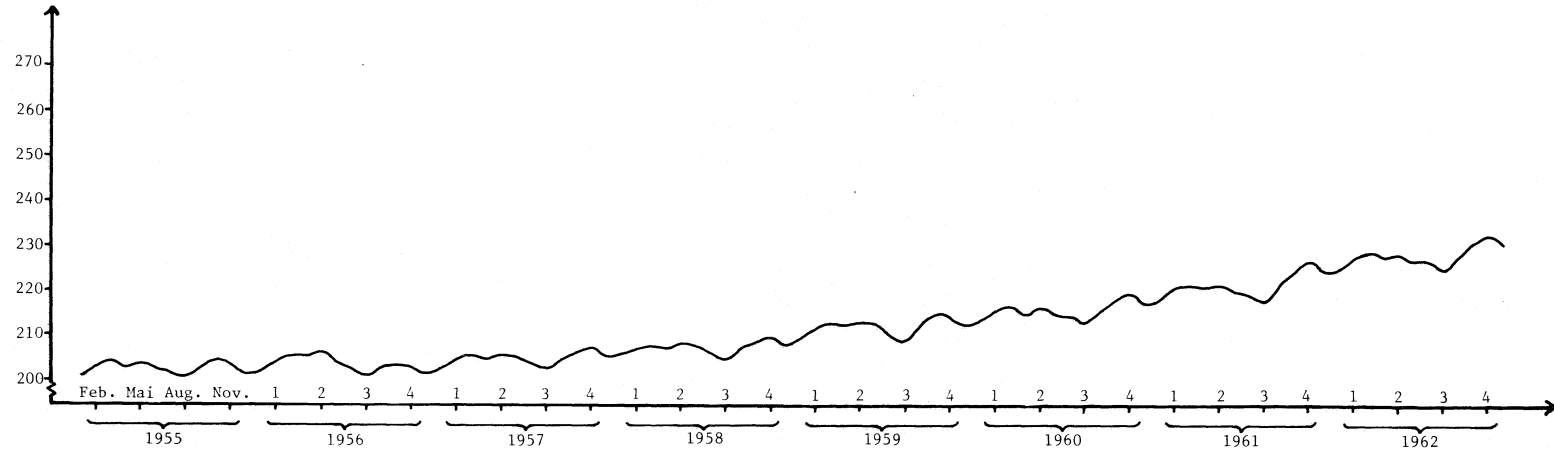
Fig.4





Sysselsætning i offentlig, sosial og
privat tjenesteyting

Fig.6



Sysselsatte i transport, lagring, post og telekommunikasjon

Fig. 7

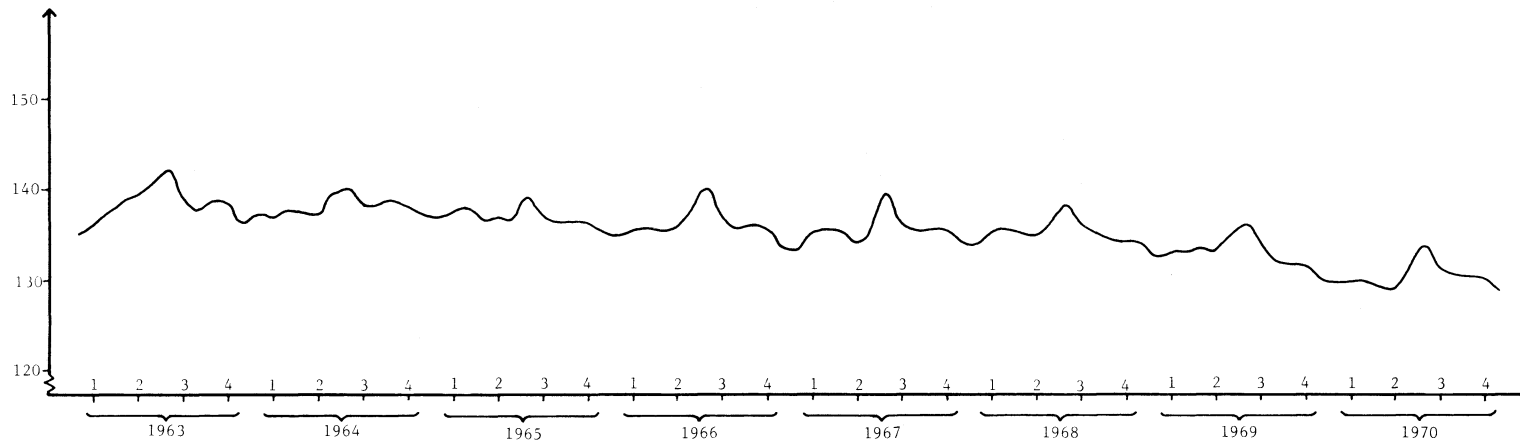
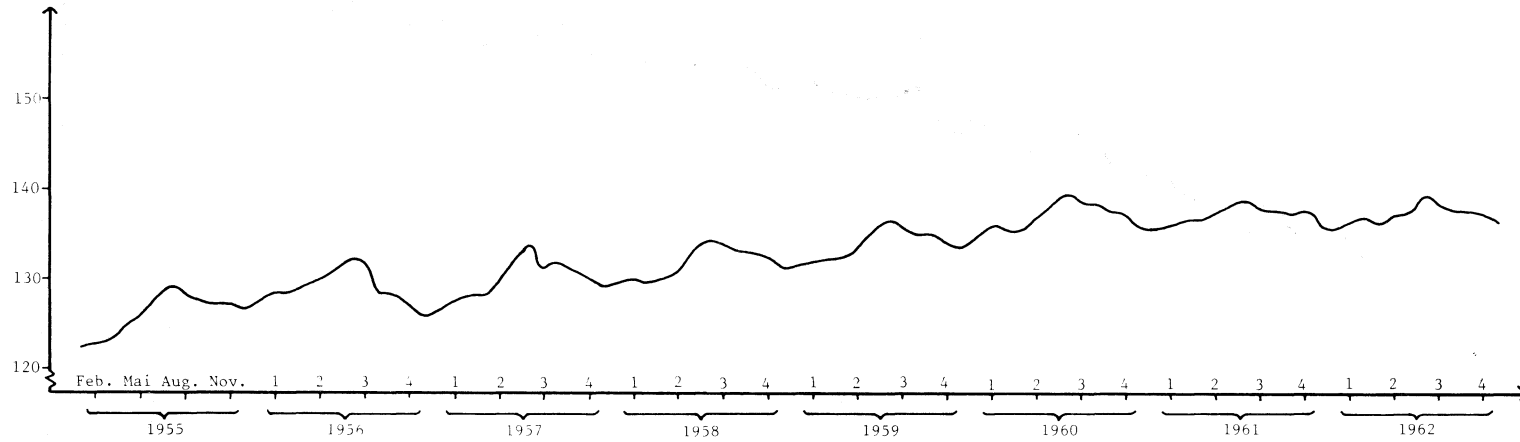


Fig. 8

$$x_t = x_{t-1} + \lambda_{t-1} - \lambda_{t-2} + \epsilon_t + \theta \epsilon_{t-1}$$

