

Orientering om säsongrensning

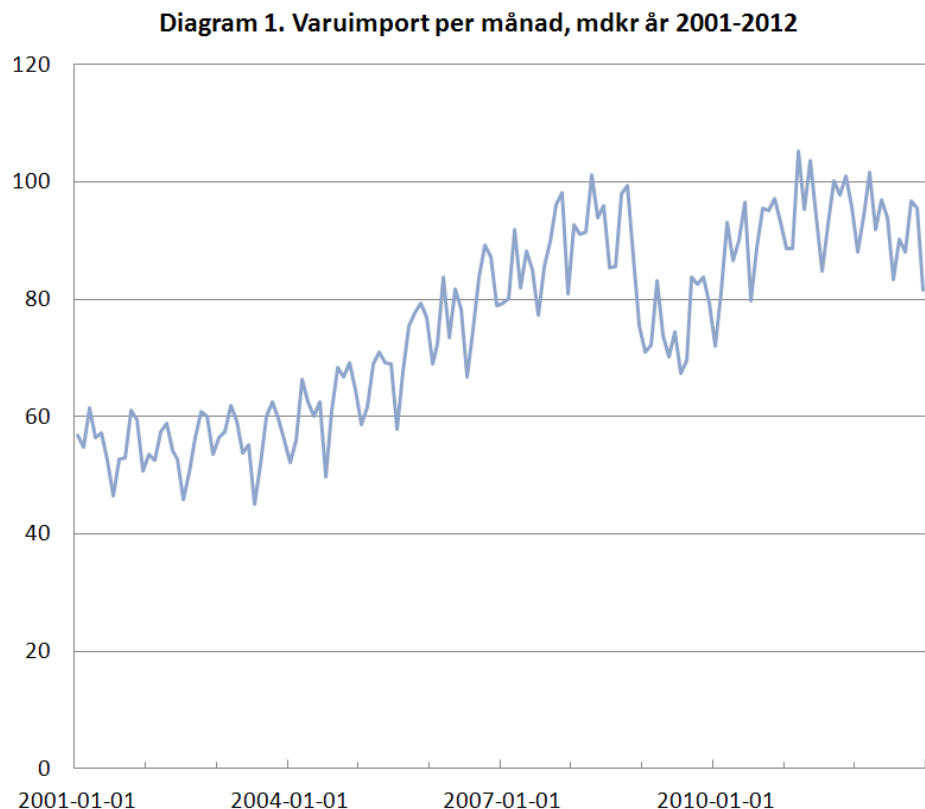
Säsongrensning (på engelska: seasonal adjustment) är ett sätt att förädla månadsstatistik och kvartalsstatistik så att den blir bättre jämförbar över tiden. I säsongrensad statistik har man räknat bort de svängningar som normalt följer av årtidernas och kalenderns växlingar. Detta gör statistiken mera talande för att tolkas och användas praktiskt.

Denna skrift syftar till att översiktligt orientera om säsongrensningen och dess roll i statistiken. Den riktar sig till alla som har att göra med månads- eller kvartalsstatistik i produktion eller användning. Beskrivningen utgår i första hand från arbetssättet vid Statistiska centralbyrån (SCB) men är till stor del generell för säsongrensningen som metod.

Först pekar vi i avsnitt 1 på problemet med säsongvariation, och i avsnitt 2 på sätt att möta problemet genom säsongrensning eller annat. Säsongrensningens idé med olika komponenter går igenom i avsnitt 3, och former för resultatets redovisning i avsnitt 4. Verktyg och praxis orienteras om i avsnitt 5, processen för säsongrensning i avsnitt 6, och i avsnitt 7 om sätt att samordna säsongrensningen mellan flera tidsserier. Slutligen ges i avsnitt 8 en sammanställning av ord och uttryck kring säsongrensning.

1 Säsongvariationen som problem

Diagram 1 ger ett exempel på månadsstatistik.



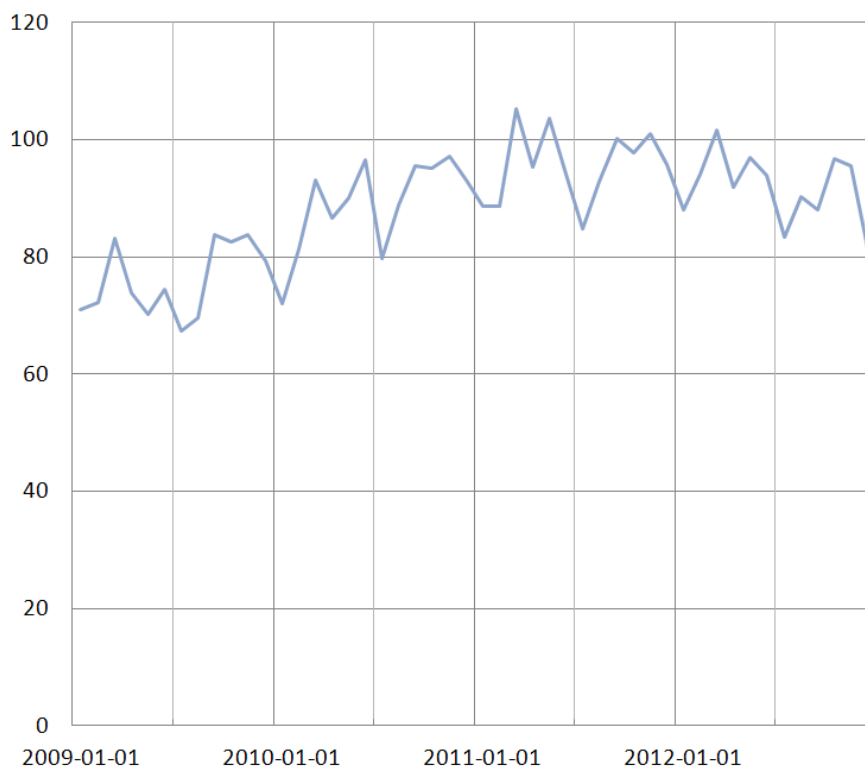
Kurvan i diagrammet följer varuimporten till Sverige i miljarder kronor månadsvis under åren 2001-2012. Kurvan visar en statistisk *tidsserie*, en serie av värden

på något vid tider efter varandra, i detta fall värden på varuimporten. Kurvan avser faktiskt observerade tal i löpande priser, utan några justeringar med hänsyn till säsong eller annat.

I diagrammet kan man se hur varuimporten har utvecklats över längre och kortare perioder. I de stora dragen kan man ana olika perioder när det har gått mest upp, gått mest ned eller mest legat still under några år i följd. Men kanske särskilt påfallande är de ständiga kortsiktiga svängningarna inom åren.

Tittar man noga kan man se att de korta svängningarna tenderar att upprepa sig likartat år efter år. Detta syns lite tydligare i diagram 1A, som visar fyra år på slutet i större tidsskala. Talen visar ett *säsongmönster*. Man kan till exempel se att talen tenderar att ligga särskilt lågt strax efter års- och halvårsskiften.

Diagram 1A. Varuimport per månad, mdkr år 2009-2012



I olika tidsserier är det vanligt med säsongmässig variation. Det är ofta snabba kast med något säsongmönster som gärna upprepar sig år efter år. Detta är naturligt, för mycket beror av årstiderna, såsom konsumtionsvanorna och olika branschens produktionsförutsättningar. Inte minst sommarsemestrarna inverkar.

Ett problem med säsongvariationen är att den lätt skymmer annat i statistiken. Den gör det svårt att jämföra talen över tiden och att läsa ut när det har kommit något nytt och oväntat, vilket ofta intresserar användare av statistiken.

2 Alternativa ansatser att hantera säsongvariationen

Det finns några olika ansatser att se bakom säsongvariationen i statistiken och blottlägga det mera intressanta som sker. Säsongrensning kan sägas vara det mest utvecklade alternativet för detta. Först ska nämnas lite om några enklare ansatser.

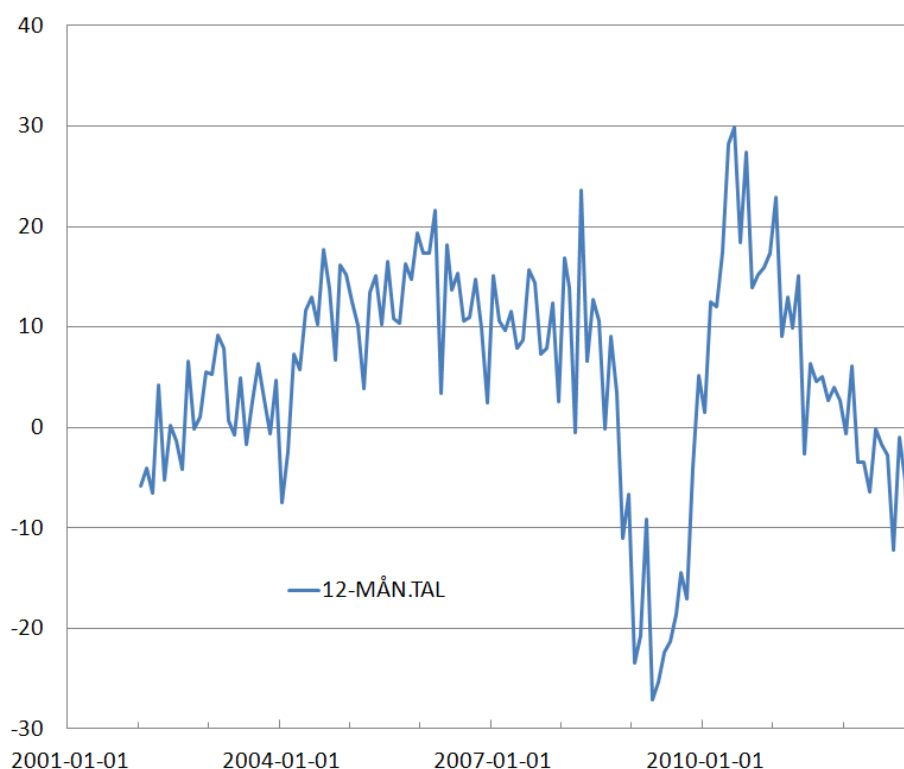
2.1 Ansatsen Årsförändringstal

En mycket vanlig och användbar ansats i månads- och kvartalsserier är att beräkna förändringen i procent under de senaste tolv månaderna, eller de senaste fyra kvartalen. Det är förändringen i procent till den aktuella månaden eller det aktuella kvartalet, från samma månad eller kvartal det närmast föregående året.

Till exempel, varuimporten som visas i diagram 1 låg på 96,5 mdr kr i oktober 2011. Det är 1,7 procent högre än nivån för samma månad ett år tidigare, oktober 2010, som var på 94,9 mdr kr. (Beräkning: $96,5 / 94,9 \times 100 - 100 = 1,7$ procent avrundat.)

I diagram 2 visas serien av sådana förändringstal i procent till aktuell månad, räknat från samma månad det närmast föregående året. Dessa procenttal kallas årsförändringstal eller tolvmånaderstal.

Diagram 2. Varuimport per månad - årsförändring i %



Årsförändringstalen uppnår i en enkel form syftet att visa utvecklingen ostört av säsongnormal variation. Detta genom att man tar samma månad året före och jämför med.

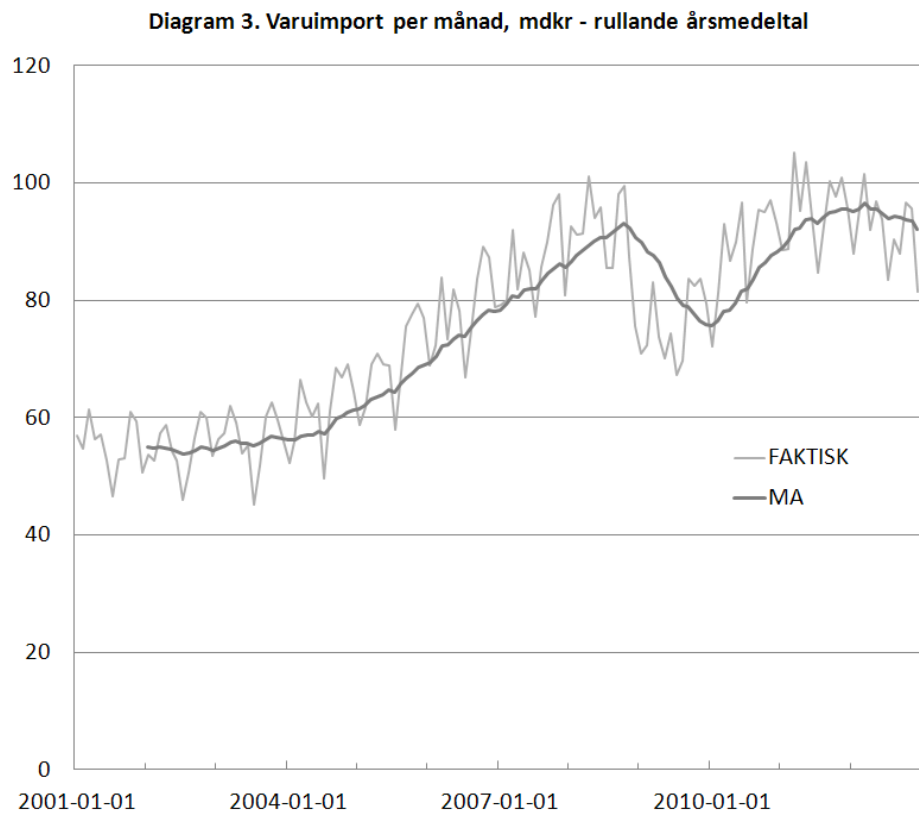
Fördelar med denna ansats är att det är lätt att ta till sig och fatta hur man har räknat. Förfarandet är väl etablerat på många håll och allmänt bekant för användare av ekonomisk statistik.

Nackdelar med denna ansats är att den inte renodlat avser det aktuella läget utan beror av läget ett år tidigare, som möjligen kan ha varit speciellt eller följt ett avvikande säsongmönster. Detta behöver användaren av resultaten ha i åtanke för att tolka dem rätt. Kurvan är också allmänt hoppig med starka oregelbundna variationer, vilket är typiskt för tal som avser förändring och inte nivå.

2.2 Ansatsen Retrospektivt enkelt glidande årsmedelvärde

En annan ganska enkel ansats är att beräkna ett retrospektivt enkelt (ovägt) glidande årsmedelvärde, även kallat rullande årsmedeltal. Detta beräknar man för varje månad som medelvärdet av seriens värden för de tolv senaste då gångna månaderna. Till exempel är det rullande årsmedeltalet för oktober 2011 lika med medelvärdet av seriens värden över tolv månadersperioden från och med november 2010 till och med oktober 2011. För kvartalsdata tar man på samma sätt medelvärdet över de fyra senaste kvartalen.

Diagram 3 visar det retrospektiva glidande årsmedelvärdet av varuimporten (kurvan MA), tillsammans med den faktiskt observerade serien från diagram 1.



Det glidande årsmedelvärdet störs inte av säsongnormal variation. Detta genom att årets alla månader kommer in på samma sätt i beräkningen, och säsongsvängningarna tar då ut varandra.

Fördelar är även med denna ansats att det är lätt att fatta hur man har räknat. Även detta förfarande är väl etablerat och bekant för många.

Nackdelar med denna ansats är att den inte följer det aktuella läget så bra utan beror mycket av tidigare lägen. Till exempel ser man i diagrammet att när serien går snabbt uppåt på det stora hela så hänger det glidande medelvärdet inte med riktigt. Då hamnar det glidande medelvärdet missvisande lågt i förhållande till de faktiska talen.

Det glidande årsmedelvärdet slätar ut lite mer och annorlunda än vad som behövs för att få bort säsongvariationen. Om detta är en för- eller nackdel i ett givet sammanhang beror på vad resultatet ska användas till. Enkla glidande årsmedelvärden är kanske inte så gängse i statistikpublicering, men användare som så

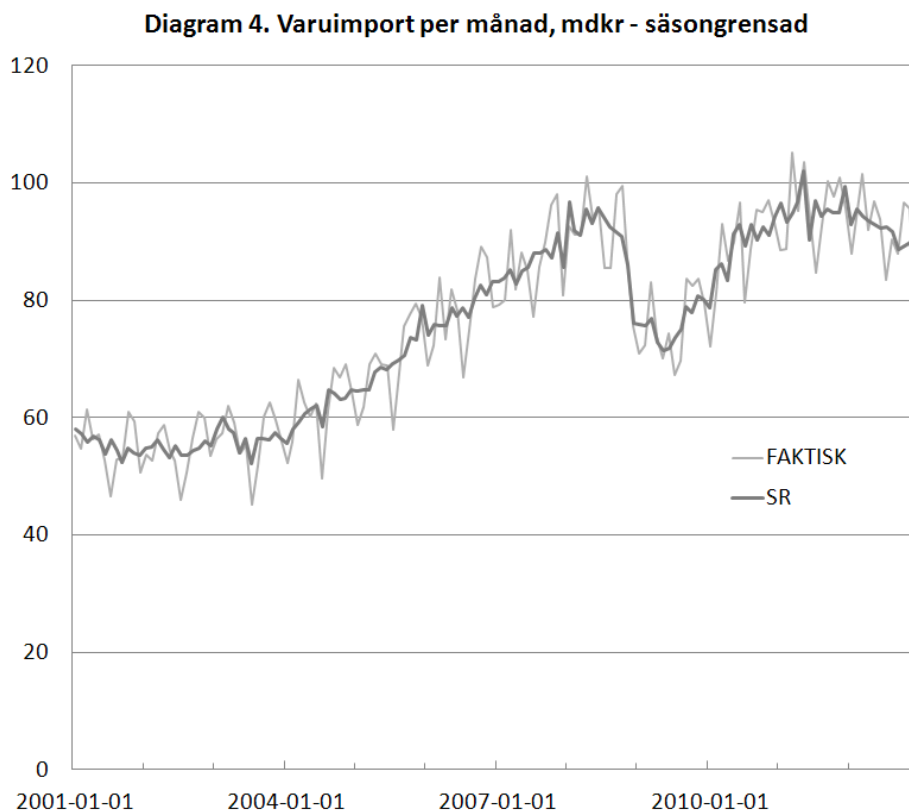
önskar kan lätt räkna ut dem själva i ett vanligt kalkylprogram, på nedladdade underlag från statistikdatabaser på webben.

De nämnda bristerna, med eftersläpning och vanligen för stark utjämning, gör att det retrospektiva enkla glidande årsmedelvärdet inte kan sägas leva upp till kraven på säsongrensning i modern mening.

2.3 Ansatsen Säsongrensning

Säsongrensning i modern mening är en ansats som nyttjar avancerade verktyg för statistisk analys men ger lättfattliga resultat. Säsongrensningen går ut på att räkna fram en serie som är väl jämförbar med den givna, faktiskt observerade serien men där den säsongnormala variationen har räknats bort. Det ger en serie som blir bättre jämförbar över tiden än den givna.

Diagram 4 visar en säsongrensad serie över varuimporten (kurvan SR), tillsammans med den faktiskt observerade serien från diagram 1.



Man ser att den säsongrensade kurvan väl följer den faktiskt observerade, utom att den inte svänger säsongmässigt. Och det var just meningen. Talen har blivit bättre jämförbara mellan månaderna när säsongvariationen inte är med längre.

Fördelar med säsongrensningen som ansats är att den är överlägset träffsäker på målet att ta hand om säsongvariationen. Den är inte så känslig för ovidkommande tillfälligheter som årsförändringstal är (avsnitt 2.1), och den siktar inte bredvid målet som retrospektivt glidande årsmedelvärde gör (avsnitt 2.2). Säsongrensningen ger resultat som är naturliga att tolka, vilket är till nytta inte minst när man tittar på många serier och har mycket att hålla reda på.

Nackdelar med säsongrensningen som ansats är att den för statistik användare i gemen kan verka lite som "svart låda" där gången inuti inte syns direkt. Resul-

taten i detalj beror också av olika val i upplägget. Vidare får man leva med att resultaten inte ligger fast utan kan *revideras bakåt* när serien aktualiseras med nya månader eller kvartal. Det är något annat än revideringar på grund av kompletteringar eller korrigeringar i underlaget till statistiken.

Den här skriften ska i fortsättningen orientera om hur säsongrensningen arbetar, vad som är bra att veta om detta i arbete med statistiska tidsseriedata.

2.4 Valet av ansats – säsongrensa (kanske) mera!

Säsongrensning är alltså en bland flera ansatser att hantera säsongvariation, och den ansatsen liksom de andra har både för- och nackdelar. Säsongrensning har på senare år fått ökad användning och skulle förmodligen kunna vara till fördel att nyttja ännu mera.

För att tjäna sitt syfte rätt och kunna inge förtroende hos statistikanvändarna behöver säsongrensningen planeras med eftertanke i ett tidigt skede, när formerna för analys och presentation av statistiken läggs upp. Man behöver utreda förutsättningarna i fråga om både materialens egenskaper och användarnas förväntningar. Detta behöver ske i samverkan mellan ämnesspecialister och metodstatistiker med specialistkompetens inom säsongrensning.

Ett skäl att säsongrensa i produktion av officiell statistik är att ge jämförbarhet mellan olika statistikanvändares redovisningar. När statistikproducenten (som SCB) säsongrensar med metoder och kvalitet som passar huvudanvändarna kan dessa slippa att göra och redovisa egna säsongrensningar som kunde riskera att inte bli fullt jämförbara.

Vissa slags förutsättningar kan vara svåra att hantera oavsett ansats. Det gäller kanske särskilt om säsongmönstret inte är stabilt utan kan ändras nyckfullt från år till år. Då riskerar säsongrensningen att inte fungera rätt och bli förledande, kanske mer än enklare ansatser som då möjligen kan vara att föredra.

Säsongrensade serier eller andra förädlade resultat ska lämpligen normalt komplettera men inte ersätta de faktiskt observerade talen i statistikredovisningen. Somliga användare kan med goda skäl behöva utgå direkt från faktiskt observerade oförädlade ”rådata”, till exempel för ekonometriska analyser med inbyggt säsongbeaktande.

3 Idén med modern säsongrensning

3.1 En tidsserie delas upp i komponenter

Statistiska analyser allmänt sett bygger ofta på *modeller* där observerade data antas bero av andra storheter på vissa sätt. Säsongrensning utgår ofta från en modell där talen i den observerade serien ses som resultat av att vissa tal har multiplicerats ihop med varandra. I formel kan modellens idé uttryckas så här:

$$(1) \quad Y = TC \times S \times K \times O \times I$$

Här står Y för ett tal i den givna serien, alltså det faktiskt observerade värdet för en given månad eller ett givet kvartal, i ett givet år. Beteckningarna TC , S , K , O och I står för storheter som också avser den månaden eller det kvartalet, i det året. De är storheter som inte kan observeras direkt utan skattas genom beräkningarna i säsongrensningen. De har olika roller som strax ska gås igenom, och de benämns enligt följande:

TC = Trendcykel (eller trend)
 S = Säsongeffekt
 K = Kalendereffekt
 O = Outliereffekt
 I = Irreguljär effekt

Det intressanta med modellen (1) är att den ger en uppdelning av den givna serien Y i delarna eller *komponenterna* TC , S , K , O och I .

Säsongrensningen utförs med en programvara som tar den givna serien som indata. Programmet räknar fram ett mellanresultat i form av skattade serier över komponenterna, och från dem sedan slutresultatet i form av den säsongrensade serien som vi kommer till (avsnitt 4.1).

Kommentar om sorter. I modellen (1) räknas komponenten TC i samma sort eller måttenhet, till exempel kronor, som den givna serien Y . De övriga komponenterna, S , K , O och I , är däremot "sortlösa" proportionsfaktorer och varierar kring värdet 1. I modellen (1) förutsätts att seriens värden Y alltid är större än 0, och alla komponenterna är alltid större än 0.

3.2 En alternativ uppdelning

Modellen (1) kallas *multiplikativ*, eftersom komponenterna multipliceras ihop med varandra. Ibland använder man i stället en *additiv* modell, där komponenterna adderas ihop. Detta ger en alternativ form för uppdelningen i komponenter. Komponenterna har här motsvarande mening men blir matematiskt lite annorlunda. I formel ser den additiva modellens idé ut så här:

$$(2) \quad Y = TC + S + K + O + I$$

Multiplikativ modell som (1) passar ofta bra för ekonomiska serier som uttrycks i kronor och alltid är större än noll. Additiv modell som (2) behövs dock för nettoposter där värdena kan växla mellan över och under noll, och den kan även passa bra för serier som räknas i antal personer, såsom antalet sysselsatta.

Kommentar om sorter. I den additiva modellen (2) räknas alla komponenterna, TC , S , K , O och I , i samma sort eller måttenhet, till exempel kronor, som den givna serien Y . Komponenterna S , K , O och I varierar här kring värdet 0 och tar värden både över och under 0.

Kommentar om ekonomiska serier. Att multiplikativ modell som (1) ofta passar för ekonomiska serier kan ha att göra med sådant som inflationen och tillväxten i ekonomin. Inflationen och den ekonomiska tillväxten ändrar efter hand skalan för belopp i kronor. Då kan det passa bättre att behandla effekterna av säsong med mera som vissa proportioner (enligt multiplikativ modell) än som vissa belopp i kronor (enligt additiv modell).

De följande beskrivningarna av komponenterna gäller i princip för både multiplikativ och additiv modell.

3.3 Komponenten Trendcykel (TC)

Komponenten Trendcykel (TC), alternativt kallad *trend*, mäter nivån för serien och rör sig trögt. Den visar bara långsiktiga rörelser som i princip pågår över längre tid än ett år i taget. I säsongrensningsprogrammet nyttjar man att den trög-rörliga egenskapen går att karaktärisera matematiskt.

Kommentar om alternativ för trendcykeln. Det kan förekomma i ekonometriska studier att man delar upp trendcykeln ytterligare i delkomponenter. Det kan vara dels en delkomponent "trend", som då får avse enbart längre rörelser över kanske fem år eller mer, dels en delkomponent "cykel" som avser konjunkturcykler över några få år. För tillämpning i officiell statistik räcker det dock normalt med trendcykeln utan denna ytterligare uppdelning.

3.4 Komponenten Säsongeffekt (S)

Komponenten Säsongeffekt (S) följer den regelbundna normala säsongvariationen som upprepar sig likartat år efter år i en jämn rytm. Denna variation beror på årstidernas inverkan i sådant som konsumtionsvanor och produktionsförutsättningar. I säsongrensningsprogrammet nyttjar man att den årsrytmiska egenskapen går att karaktärisera matematiskt.

Lite överkurs. En tidsserie som rör sig kring en konstant nivå kan behandlas matematiskt enligt teorin för spektralanalys. Den delas då upp som en sammanvägning av ”rena” pendelsliknande svängningsförlopp med olika svängningstakt eller frekvens. Frekvenserna räknas i antal svängningar eller cykler per år och tar sina värden på en kontinuerlig skala upp till en övre gräns, som är 6 cykler per år för månadsserier och 2 cykler per år för kvartalsserier (kallad Nyquist-frekvensen). *Säsongkomponenten* är den del av sammanvägningen som svarar mot frekvenser på heltal cykler per år. De frekvenserna är 1, 2, 3, 4, 5 eller 6 cykler per år för månadsserier, och 1 eller 2 cykler per år för kvartalsserier.

3.5 Komponenten Kalendereffekt (K)

Komponenten Kalendereffekt (K) mäter inverkan av sådant som att de olika veckodagarnas antal under en månad eller ett kvartal varierar lite mellan åren, genom att veckorna faller lite olika i kalendern. En serie över till exempel produktionsvolym under månaden beror naturligt av antalet arbetsdagar, och motsvarande kan även gälla andra serier. Man kan även beakta ”påskeffekten”, verkan av hur påsken infaller. Kalenderns verkan på serien skattas numera helst utifrån seriens observerade värden med en regressionsanalys. (Mer i avsnitt 6.5.)

3.6 Komponenten Outliereffekt (O)

Komponenten Outliereffekt (O) tar hand om exceptionellt utstickande värden i serien. Det är värden som avviker så starkt från vad som kan förväntas enligt modellen att det vore osannolikt vid annars observerad variabilitet. Denna komponent behöver skiljas ut för att skattningmetoderna för trendcykeln och säsongeffekten brister i robusthet mot utstickande värden. (Mer i avsnitt 6.4.)

Kommentar om terminologin. *Extremvärden* säger man ibland men den termen är lite mindre tydlig. För den står annars ofta för ett värde som är högst eller lägst någonstans vilket inte är samma sak som att ett värde sticker ut. Dock är problemet lite annorlunda här än när det gäller outliers, utstickande värden bland uppgifter för enskilda företag eller så.

3.7 Komponenten Irreguljär effekt (I)

Komponenten Irreguljär effekt (I) fångar upp de oregelbundna variationer som finns i den givna serien och inte förklaras av de andra komponenterna. ”Brus” i en mening, eller restkomponent om man så vill. Till denna komponent bidrar källor av olika slag:

- Oregelbundna variationer i de verkliga förhållandena som serien beskriver.
- Slumpmässig statistisk osäkerhet i den observerade serien genom olika osäkerhetskällor såsom urvalsosäkerhet och slumpmässig mätosäkerhet.
- Modellosäkerhet i skattningen av de övriga komponenterna.

Kommentar om slumpmodelleringen. När säsongrensningsprogrammet skattar komponenterna i modellen och gör statistiska test så antas att den irreguljära komponenten är slumpmässig med (logaritmisk) normalfördelning, oberoende mellan månader eller kvartal.

4 Resultatet

4.1 Resultat i form av Säsongrensad serie

Den primära formen för redovisning av säsongrensningens resultat är den säsongrensade serien. Ett exempel på en säsongrensad serie visades i diagram 4 (avsnitt 2.3) och kommenterades där.

Idén med den säsongrensade serien är att man där har räknat bort säsongeffekten och även kalendereffekten såsom beskrevs i avsnitt 3. Det var nämligen primärt dessa två effekter som störde jämförbarheten över tiden i den givna serien.

Uttryckt i formel kan man uppfatta den säsongrensade serien som att man i modellen (1) stryker säsongkomponenten och kalenderkomponenten. För den säsongrensade serien SR blir då kvar:

$$(3) \quad SR = TC \times O \times I$$

I praktiken når man samma beräkningsresultat genom att i stället ta fasta på säsongeffekten och kalendereffekten, och räkna bort dem (dividera bort dem) ur den givna serien, enligt:

$$(4) \quad SR = Y / S / K$$

För den additiva modellen (2) blir den säsongrensade serien på motsvarande sätt:

$$(5) \quad SR = TC + O + I$$

Den kan då beräknas enligt:

$$(6) \quad SR = Y - S - K$$

4.2 Resultat i form av Kalenderkorrigerad serie

Det finns efterfrågan även på serier som är enbart kalenderkorrigerade och inte säsongrensade. Då ska kalendereffekten men inte säsongeffekten vara borträknad.

Kalendereffekten är visserligen ofta mindre betydande än säsongeffekten men kan å andra sidan vara svårare att genomskåda intuitivt för den som tolkar statistiken. Detta genom att den är av mera teknisk karaktär och beror av detaljer som skiftar mellan åren på ett komplicerat sätt. För somliga användningar kan det därför passa att statistikproducenten hjälper till med kalenderkorrigeringen men annars lämnar säsongbeaktandet till användarens tolkning och analys.

I formel kan man uppfatta den kalenderkorrigerade serien som att man i modellen (1) stryker kalenderkomponenten. För den kalenderkorrigerade serien KK blir då kvar:

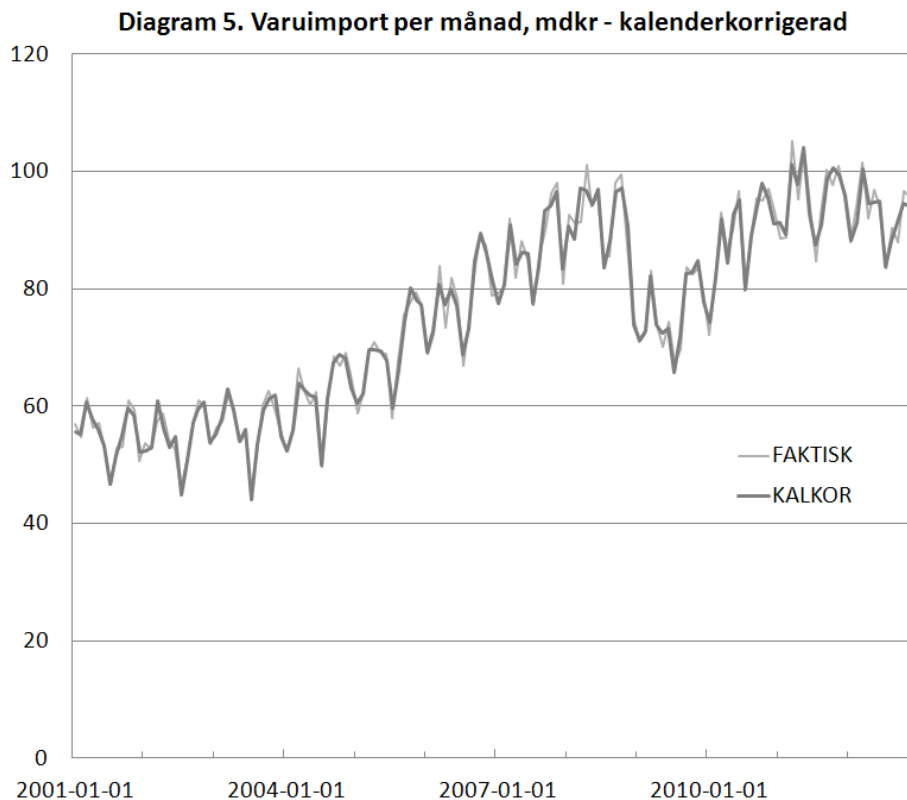
$$(7) \quad KK = TC \times S \times O \times I$$

Den serien kan då beräknas enligt:

$$(8) \quad KK = Y / K$$

På motsvarande sätt kan man kalenderkorrigera utifrån den additiva modellen (2) när den har valts.

Diagram 5 visar en kalenderkorrigerad serie över varuimporten (kurvan KALKOR), tillsammans med den faktiskt observerade serien från diagram 1.



Man ser i diagrammet att kalenderkorrigeringen har en ganska marginell inverkan. Men även små skillnader kan vara viktiga i en del sammanhang, såsom små konjunkturtendenser i stabila branscher, så kalendereffekten kan vara annat än försumbar i situationen.

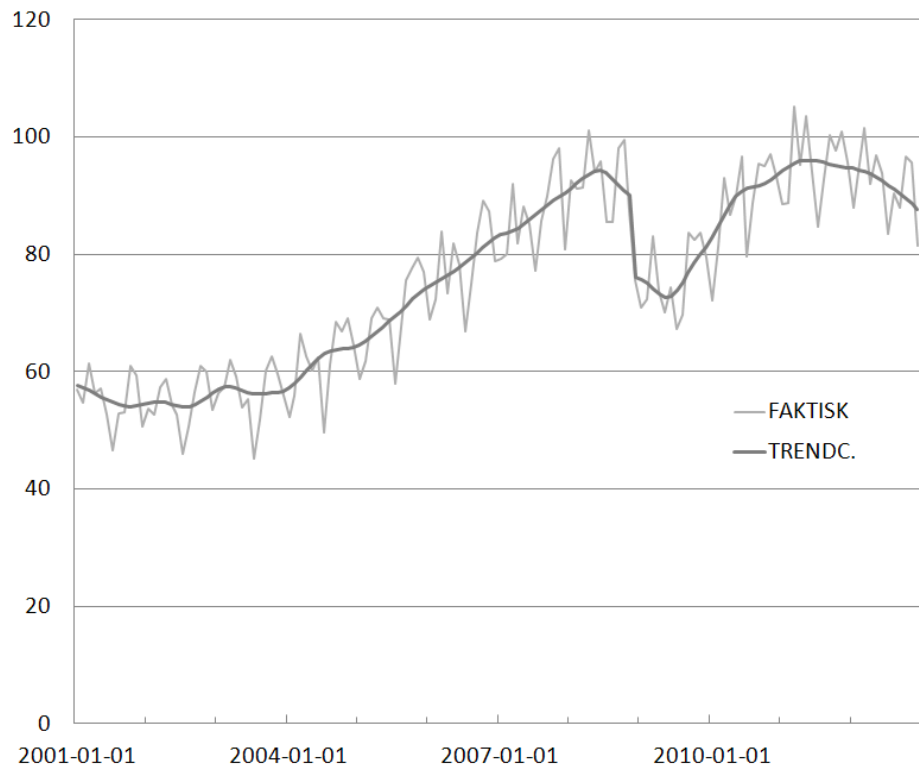
4.3 Resultat i form av Trendcykel

Det finns också efterfrågan på redovisning av komponenten trendcykel separat. Då handlar det om att gå längre än att eliminera den säsongnormala variationen och kalendereffekten. I trendcykeln har man i princip rensat bort alla snabba kortsiktiga fluktuationer, oavsett om de är säsongnormala eller inte. Det kan göra trendcykeln intressant tolkbar i en del användningar, även om den beror av den valda beräkningsmetoden och i en del fall av ingrepp i outlierhanteringen.

Diagram 6 visar en serie som avser trendcykel för varuimporten (kurvan TRENDC.), tillsammans med den faktiskt observerade serien från diagram 1.

Med sin jämna utslätade kurva i diagrammet påminner trendcykeln lite om det retrospektiva glidande medelvärdet i diagram 3 (avsnitt 2.2). Trendcykeln som redovisningsform har dock fördelar framför det senare, främst i att den ligger mera rätt i nivå och hänger med bra i nivå även när serien i stort drar åt ett håll under några år.

Diagram 6. Varuimport per månad, mdkr - trendcykel



Det kan kanske vara lockande att se trendcykelkurvan som en bild av något slags bakomliggande tendens i ekonomin, men den beror ändå av beräkningsmetoden och man får akta sig lite för att övertolka.

5 Verktyg och praxis

5.1 Metod och programvara

Standardverktyget för säsongrensning vid SCB bygger på en metod och programvara som heter X-12-ARIMA och har tagits fram vid U.S. Census Bureau. Detta verktyg används i många länder. Den implementering av verktyget som är fastställd standard vid SCB ingår i SAS[®] programsystem och heter där procedur X12.

Ett alternativt verktyg som tidigare var standard vid SCB och ännu kan förekomma övergångsvis är TRAMO-SEATS från Spaniens centralbank. Metodmässigt är detta verktyg i stort sett likvärdigt med X-12-ARIMA som dock har praktiska fördelar, bland annat genom tillgången i SAS[®] programvarumiljö. Resultaten kan skilja sig mellan de två verktygen men beror mera på vilka inställningar som väljs. Metodmässigt bygger TRAMO-SEATS något mera uttryckligt på modeller i vissa led såsom trendcykelskattningen, men det är en nyansskillnad, för båda verktygen beror av modeller. En efterföljare till X-12-ARIMA, kallad X-13ARIMA-SEATS, har utvecklats vid U.S. Census Bureau och går att ställa in för motsvarande funktionalitet som TRAMO-SEATS. Den har inte införts i produktionen vid SCB. En ytterligare alternativ programvara kallad JDemetra+ har tagits fram på initiativ av Eurostat och innefattar X-12-ARIMA och TRAMO-SEATS, men har inte heller införts vid SCB.

5.2 Konsten att säsongrensa

Programvaran för säsongrensning går att ställa in enligt alternativa val på flera punkter i utförandet (vilket ska exemplifieras i avsnitt 6.4-6.5). Konsten att säsongrensa handlar mycket om att kunna välja dessa inställningar så att det säkrar resultatens tillförlitlighet och användbarhet. Genom att säsongrensning inte är något helt entydigt definierat begrepp kommer det in element av bedömning, och där behövs insikt om statistikens innehåll och statistik användarnas perspektiv.

Att låta programmet välja inställningar helt automatiskt efter datamaterialet undviks i regel. Helt automatiska val kunde nämligen låta slumpvariationer slå igenom för mycket och även bortse från vad man kan känna till om outliers eller annat. Flera möjliga inställningar kan tekniskt passa datamaterialet nästan lika bra men ge delvis ganska olika resultat, och det kunde slå något godtyckligt. Det är därför normalt bättre att undvika ändra inställningarna mellan produktionsomgångarna så länge de inte börjar passa mycket för illa.

Programmet producerar mått på osäkerheten i skattningen av komponenterna och en rikhaltig mängd *diagnostika*, processdata som indikerar hur säsongrensningen fungerar på datamaterialet. Detta tjänar till stöd för säsongrensningsspecialistens bedömningar.

5.3 Tekniska vägledande dokument

Valmöjligheterna i säsongrensningssprocessen medför att det i stort behövs en samsyn på hur dessa möjligheter ska användas, för att en pålitlig kvalitet ska kunna hållas. Principer för hur man säsongrensar vid SCB finns därför dokumenterade i vägledande tekniska dokument i Verksamhetsstödet. Det är dels ett dokument om *Praxis* för säsongrensning, dels en *Vägledning* (i skrivande stund under arbete). Dessa tekniska dokument vänder sig i första hand till metodstatistiker och ansvariga inom säsongrensning, men att de finns för kvalitetssäkringen är bra att veta även för andra. Där ges även referenser till litteratur för närmare teknisk information och underbyggnad. Ett dokument på EU-nivå med motsvarande målgruppsprofil är:

ESS Guidelines on Seasonal Adjustment,

http://epp.eurostat.ec.europa.eu/portal/page/portal/research_methodology/methodology/seasonal_adjustment

6 Processen för säsongrensning

6.1 Hur programmet räknar

Programvaran för säsongrensning utför beräkningarna i flera olika steg. Syftet är att räkna fram skattade serier över komponenterna som beskrevs i avsnitt 3, och från dem sedan den säsongrensade serien och andra resultat som beskrevs i avsnitt 4.

Ett viktigt internt hjälpmedel i programmets beräkningar är att skriva fram och förlänga tidsserien med i regel minst drygt ett år framåt och bakåt. Detta görs i programmet genom att serien analyseras enligt en tidsseriemodell av en typ kallad *ARIMA* (AutoRegressive Integrated Moving Average). *ARIMA*-modellen bygger framskrivningen på vissa specificerbara antaganden om mätbara statistiska samband mellan värden kort efter varandra i serien. Framskrivningen presenterar man i regel inte utåt som resultat utan den tjänar bara som ett internt mellansteg i programmet.

I ett tidigt skede skattas kalendereffekten och outliereffekten. Då används en så kallad RegARIMA-modell, ett slags regressionsmodell som känner av hur stort genomslaget är på seriens värden. Den modellen fungerar som om den resterande serien utan de nämnda effekterna skulle följa en ARIMA-modell.

När kalendereffekten och outliereffekten är borträknade kan trendcykeln och säsongkomponenten skattas på serien som är kvar. I standardverktyget X-12-ARIMA sker detta med så kallade *filter*, som är ett slags vägda glidande medelvärden. För att ge rätt nivå och precision är de vägda glidande medelvärdena i första hand symmetriska på så sätt att de räknas på månader eller kvartal som ligger symmetriskt före och efter den månad eller det kvartal som avses. Där kommer det till pass att serien skrevs fram.

Filtreringarna sker i några steg på varandra. Först görs en preliminär trendfiltrering, och därpå en preliminär säsongfiltrering på skillnaden mellan serierna före och efter trendfiltreringen. Sedan följer ännu en trendfiltrering och en säsongfiltrering, på skillnaden mellan serierna före och efter föregående filtrering. Då kan trendcykeln och säsongeffekten falla ut tillräckligt rent i sina komponenter.

Det alternativa verktyget TRAMO-SEATS arbetar delvis på liknande sätt med filter för att skatta trendcykeln och säsongkomponenten, men filtren är där inte så fasta som i X-12-ARIMA utan delvis anpassade efter en skattad ARIMA-modell.

6.2 Design, årsöversyn, produktion – roller och ansvar

För att säkra kvaliteten och integriteten i säsongrensningen på en statistikprodukt behöver man planera för den på förberedande stadier, i designprocessen. Förutsättningarna i fråga om statistikanvändarnas behov och datamaterialens egenskaper ska då utredas. Test av körning med olika inställningar och val kan göras på de serier det gäller, så långt fram som de har kunnat observeras då.

Säsongrensningsspecialister, dvs. metodstatistiker med särskild kompetens på säsongrensningsspecialiteter, behöver här samverka konsultativt med produkt- och metodansvariga, ämnesspecialister och statistikanvändare.

En *årlig översyn* av upplägget i säsongrensningen på statistikprodukterna behöver göras av säsongrensningsspecialister. Där företar man nya tester och följer upp diagnostika (processdata) med tanke på om upplägget behöver modifieras till följd av ändrade förutsättningar.

Den löpande produktionen varje månad eller kvartal kan normalt utföras med rutiner och inställningar som lades fast i årsöversynen. Säsongrensningsspecialister behöver då ofta inte medverka men kan lämpligen finnas till hands att kunna rådfrågas om något oväntat dyker upp. Processdata sparas.

Kommentar om en detalj. Varje gång serien har aktualiserats med en ny månad eller ett nytt kvartal så skattar programmet ånyo de utnyttjade ARIMA- och RegARIMA-modellerna. De skattningarna kan alltså grundas på hela den aktualiserade seriens data. Modellernas specifikationer ligger däremot fast mellan årsöversynerna. Dessa principer kallas *partial concurrent adjustment*.

6.3 Datamaterialets förutsättningar

En grundfråga redan i den förberedande planeringen är om serierna lämpar sig att säsongrensas. Ett första krav är att serierna är långa nog, normalt minst fem år, för att kunna vittna tillräckligt väl om säsongmönstret.

Utöver att serierna är långa nog krävs för säsongrensningen också att saknade värden och nollvärden inte är för vanligt förekommande och att säsongmönstret är något så när stabilt. Bristande stabilitet i säsongmönstret kan uppstå i följande situationer:

- Om den säsongnormala årsrytmiska variationen tenderar att överflyglas av stark oregelbunden variation i den observerade serien, ”kraftigt brus”.
- Om säsongerna för produkter eller annat lätt förskjuts inom året, så att de ibland kommer tidigare eller senare än annars.

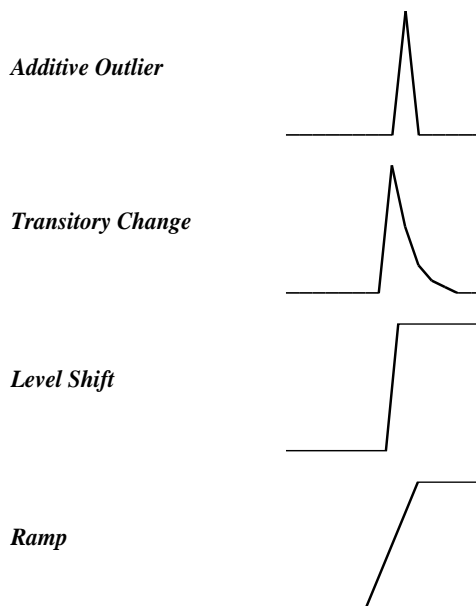
Redan i planeringen får man också utreda andra aspekter på kvaliteten i de observerade serierna, såsom bristande jämförbarhet över tiden genom metodomläggningar eller annat. För jämförbarhetsstörningar får bedömas om de gör säsongrensningen olämplig eller till exempel kan hanteras som outliers (jfr nästa avsnitt).

I en del tillämpningar kan man ha att göra med ett stort antal serier samtidigt. Då får man avväga mellan hänsyn till dels seriernas förutsättningar var för sig, dels jämförbarheten mellan serierna, dels begränsningarna i utredningsresurserna.

6.4 Outlierhantering

För hanteringen av outliers, utstickande värden, behövs bedömningar och val i upplägget och utförandet av säsongrensningen, ibland med stor inverkan på resultatet. Olika typer av ”utstickande” kan urskiljas.

En typ är den som kallas ”Additive Outlier”. Där sticker serien ut endast för en månad eller ett kvartal, och hoppar sedan direkt tillbaka till en mera ”normal” nivå. Andra typer är ”Transitory Change”, där återgången dröjer något, och ”Level Shift”, där serien inte går tillbaka snart utan fortsätter i anslutning till den nya nivån. En ytterligare typ är ”Ramp”, där serien successivt skiftar nivå. Se följande diagram.



I outlierhanteringen ingår ofrånkomligen bedömningar. Detta gäller inte minst det intressanta slutet på serien, vad som händer nu. Typen av outlier kan där inte bestämmas på en fullgod empirisk grund, för man kan inte veta hur serien ska röra sig i fortsättningen, utan man får ta stöd i bedömningar och konventioner.

För att indikera var i serien outliers förekommer nyttjar säsongrensningsprogrammet statistiska test. Dessa testar om det observerade värdet avviker signifikant starkt från ett förväntat värde enligt modellen. Ett val som behöver göras är att sätta gränsen för hur osannolikt starkt avvikande det observerade värdet ska behöva vara för att klassas som outlier.

Risk kan finnas att en månad eller ett kvartal genom slump kan klassas olika mellan produktionsomgångarna, och gå fram och tillbaka mellan att klassas som outlier och inte. Denna metodrelaterade osäkerhetskälla motverkas genom regler om en viss stabilisering i outlieridentifieringen.

Outlierhanteringen kan ha stark inverkan på statistikresultaten. Det är därför angeläget att grunda outlierhanteringen på klara principer utifrån kunskap om innehållet i statistiken och användarnas perspektiv.

6.5 Kritiska val: Kalendervariabel, framskrivningsmodell

Ett val som behöver göras i upplägget är valet av "kalendervariablerna", de förklarande variablerna i regressionsanalysen som skattar kalendereffekten. En förenkling är att som enda kalendervariabel genomgående ta antalet helgfria veckodagar, alltså normala arbetsdagar, relaterat till dess långsiktiga medelvärde över kalendrar för olika år. Detta val är knappast idealt för exempelvis detaljhandeln men kan ha en fördel som enhetligt val i ett större sammanhang. Oavsett valet av kalendervariabel är det en fördel att regressionsanalysen med RegARIMA-modell avpassar kalenderkorrigeringen till att motsvara den empiriska kalendereffekten på den serie det gäller.

Val som behöver göras gäller också specifikation av ARIMA-modellen för framskrivningen av serien. Valen där är särskilt viktiga med hänsyn till tillförlitligheten i det för tunga användningar intressanta slutet på serien som ger bilden av nuläget.

6.6 Säsongrensningens kvalitet

Om säsongrensningen fungerar som avsett så förbättrar den statistikens kvalitet i fråga om jämförbarheten över tiden, utan att för mycket påverka dess kvalitet i fråga om tillförlitlighet, aktualitet eller annat. Risker för kvaliteten har att göra med risken att säsongrensningen inte fungerar helt som avsett, och den risken i sin tur beror i praktiken kanske mest på om serien kan variera på ett oregelbundet sätt som säsongrensningsmodellen har svårt att klara. Finns sådana oregelbundenheter så tillför de osäkerhet i skattningarna av komponenterna och därmed osäkerhet i den säsongrensade serien, enligt principen "skräp in – skräp ut".

Det finns inte något vedertaget sammanfattande mått på säsongrensningens kvalitet. Det får främst ankomma på säsongrensningsspecialisten att bedöma kvalitetsaspekterna utifrån processdata från säsongrensningsprogrammet. Bland annat kan det då vara belysande att se efter om värdena på den irreguljära effekten fördelar sig enligt det slumpmönster som teoretiskt kan förväntas.

I fråga om relevansen har säsongrensningen principiella begränsningar när det gäller de intressanta *vändpunkterna*, tidpunkterna när det just ska till att vända i samhällsekonomin. Vändpunkterna kan gå att utläsa tydligt något år efteråt. Men säsongrensningen kan knappast fås att helt säkert signalera om vändpunkterna just när de händer och ligger i slutet av serien man har – och kanske är som intressantast.

Kommentar om teoretiska kvalitetskriterier. Ett exempel på ett möjligt teoretiskt kriterium som anförs ibland men kanske kan diskuteras är så kallad idempotens: att serien borde återges opåverkad om programmet körs en gång till på en serie som redan är säsongrensad. Exakt kan detta inte uppfyllas av filtren i sig utan bara konstruerat genom att programmet anpassar bearbetningen och till exempel hoppar över att räkna bort säsongeffekten om den inte är statistiskt signifikant. Ett annat kriterium är att revideringarna bakåt när serien aktualiseras inte ska vara större än nödvändigt, och det gynnas om man är försiktig med att ändra inställningarna. Ett basalt kriterium är om säsongeffekten är statistiskt signifikant; är den inte det så har säsongrensningen kanske inte bidragit med annat än ökad osäkerhet.

7 Samordnad säsongrensning av flera serier

7.1 Problemet om konsistens i aggregering

I statistiska och ekonomiska tabeller med sammanräknade summor väntar man sig att talen ska vara *summakonsistenta* med varandra, så att totalposter stämmer överens exakt med summorna av de delposter som ingår. Till exempel exporten totalt väntas stämma överens med summan av alla de olika branschernas export.

Säsongrensar man utan vidare de olika posterna var för sig kommer dock summorna normalt inte att stämma exakt som de borde. De skattningar som ingår i säsongrensningen kan nämligen slå lite olika. Summakonsistensen kan dock vara ett tungt användarkrav, och det går att uppnå den genom olika ansatser med samordning mellan serierna som strax beskrivs.

Kommentar om ett relaterat problem. Ett något annorlunda problem med aggregeringskonsistens förekommer för serier med relativa tal såsom indextal. En indexserie över produktionsvolymerna i en bransch kan någon gång naturligt röra sig åt motsatt håll mot alla serierna för de delbranscher som ingår, men efter säsongrensning kan denna paradox uppträda på ett onaturligare och mera störande sätt. Vi ska här inte gå närmare in på detta och i det följande behandlas bara summakonsistens i serier av sammanräknade summor, inte problematiken för indexserier.

7.2 Ansatsen med indirekt säsongrensning

Direkt säsongrensning kallas det när totalserier och delserier säsongrensas var för sig, och det ger som sagt problem med bristande summakonsistens. *Indirekt* säsongrensning går ut på att först säsongrensa enbart delserierna och sedan räkna ihop dessa till totalnivå. Då kan man få ett summakonsistent system av säsongrensade serier på både delnivå och totalnivå.

Denna indirekta princip tillämpas när man räknar samman från länder till EU-nivå, och även till exempel inom USA. Inom Sverige däremot används den bara begränsat. För i en relativt liten ekonomi som den svenska kan delserierna vara mera influerade av oregelbundenheter, och metoden kunde då göra totalserien onödigt osäker. Det senare har bekräftats i empiriska studier.

7.3 Ansatsen med ensning av metodspecificeringarna

En annan ansats att uppnå summakonsistens är att eliminera orsakerna till störningar i summakonsistensen genom att utföra säsongrensningen med tillräckligt enhetliga specificeringar mellan serierna. Åtgärderna gäller särskilt:

- Additiv (inte multiplikativ) modell för alla serierna.
- Samma modellspecifikation för framskrivningen i alla serierna.
- Samma eller sammanjämkad outlierhantering i alla serierna.
- Samma eller samordnad skattning av kalendereffekten i alla serierna.
- Samma val av trendfilter och säsongfilter för alla serierna.

Ansatsen med ensning av metodspecificeringarna tillämpas för säsongrensningen i Arbetskraftsundersökningarna (AKU), vars förutsättningar gör att den passar.

7.4 Ansatsen med rekonciliering av resultaten

En ytterligare ansats att uppnå summakonsistens är att säsongrensa serierna var för sig och sedan i efterhand jämka eller rekonciliera delserierna så att summan fås att stämma med totalserien. Detta utförs med en automatiserad beräkning enligt en regel som minimerar ingreppets påverkan enligt ett valt kriterium.

Ansatsen med rekonciliering av resultaten är huvudmetod vid SCB för att uppnå summakonsistens, enligt dokumentet om Praxis för säsongrensning. Metoden tillämpas för säsongrensningen i Nationalräkenskaperna (NR). Planer finns på ett mera generellt verktyg för metoden.

Metoder för konsistens med årsstatistik. Liknande metoder som den nu beskrivna finns för att efter årets slut justera månads- eller kvartalsstatistik så att summorna över året fås att stämma med motsvarande årsstatistik, så kallad *benchmarking* (i en speciell mening).

7.5 Ansatsen med multivariat strukturmodell

Teoretiskt tilltalande vore att ställa upp en multivariat strukturmodell som förklarar de olika observerade serierna samtidigt och ger skattade summakonsistenta säsongrensade serier. Teorin för sådana modeller finns och är etablerad i ekonomisk forskning, en form av så kallade *state-space-modeller*. Men för att tiden ska bli mogen för tillämpning på säsongrensning inom officiell statistik kan idéer och konsekvenserna för resultatens tolkbarhet behöva belysas mera och förankras hos användare och andra, även internationellt.

8 Ord och uttryck kring säsongrensning

Ordförklaringar med hänvisningar till avsnitt.

ARIMA-modell är en metod för att analysera, skriva fram och förlänga *tidsserier* under vissa antaganden, vilken används internt i säsongrensningsverktygen. (Avsnitt 6.1.)

Direkt säsongrensning innebär att man säsongrensar summaposter och delposter var för sig. (Avsnitt 7.2.)

Extremvärden är här ett annat ord för *outliers* (se *Outliereffekt*). Förväxla inte med extremvärden i betydelsen största och minsta värdena. (Avsnitt 3.6.)

Filter är ett slags vägda glidande medelvärden som används i säsongrensningsverktyg för att skatta *trendcykeln* och *säsongkomponenten*. (Avsnitt 6.1.)

Indirekt säsongrensning innebär att man först säsongrensar delposter och sedan räknar ihop dem till summaposter. (Avsnitt 7.2.)

Irreguljär effekt är en komponent i *säsongrensningsmodellen*. Denna komponent uttrycker oregelbundna variationer som inte fångas upp i de andra komponenterna. (Avsnitt 3.1, 3.7.)

Kalendereffekt är en komponent i *säsongrensningsmodellen*. Denna komponent mäter inverkan av kalenderns egenskaper såsom antal arbetsdagar under månaden eller kvartalet. (Avsnitt 3.1, 3.5, 6.5.)

Kalenderkorrigerig innebär att räkna bort *kalendereffekten* från en *tidsserie*. (Avsnitt 4.3, 6.5.)

Outliereffekt är en komponent i *säsongrensningsmodellen*. Denna komponent mäter inverkan av *outliers*, värden i tidsserien som avviker starkt från förväntat värde och särbehandlas för att inte störa säsongrensningen. Förväxla inte med outliers i en annan betydelse för data om enskilda uppgiftsobjekt. (Avsnitt 3.1, 3.6, 6.4.)

RegARIMA-modell är ett slags regressionsanalysmodell som används i säsongrensningsverktyg för att skatta *kalendereffekten* och *outliereffekten*. (Avsnitt 6.1.)

Serie är här ofta ett kortare ord för *tidsserie*.

Strukturmodeller, state-space-modeller är modeller som ligger till grund för vissa alternativa metoder för säsongrensning, hittills inte mycket använda för officiell statistik. (Avsnitt 7.5.)

Summakonsistent sägs statistikredovisningen vara om summaposterna stämmer korrekt med summorna av de delposter som ingår. För säsongrensad statistik behövs vissa grepp för att detta ska uppfyllas. (Avsnitt 7.1.)

Säsongeffekt är en komponent i *säsongrensningsmodellen*. Denna komponent mäter inverkan av årstiderna, dvs. *säsongvariationen*. (Avsnitt 3.1, 3.4.)

Säsongrensning är en analysansats som räknar bort *säsongeffekten* och *kalendereffekten* från en *tidsserie*. (Avsnitt 2.3, 3.1-3.7, 4.1.)

Säsongrensningsmodellen beskriver hur en tidsserie kan delas upp i olika komponenter och ger en grund för *säsongrensning* av serien. (Avsnitt 3.1.)

Säsongvariation, eller säsongnormal variation, är sådan variation som normalt följer av årstidernas skiftningar. (Avsnitt 1.)

Tidsserie är en serie värden på en statistisk storhet vid tider eller perioder (såsom månader, kvartal eller år) efter varandra. (Avsnitt 1.)

TRAMO-SEATS är ett verktyg med metod och programvara för säsongrensning. (Avsnitt 5.1.)

Trend är här ofta ett kortare ord för *trendcykel*. Ordet kan annars förekomma i andra betydelser. (Avsnitt 3.3.)

Trendcykel är en komponent i *säsongrensningsmodellen*. Denna komponent mäter tidsseriens nivå och långsiktiga rörelser över längre tid än ett år. (Avsnitt 3.1, 3.3, 4.3.)

X-12-ARIMA är ett verktyg med metod och programvara för säsongrensning. (Avsnitt 5.1.)