

Introduktion til sæsonkorrektion

Hvad er sæsonkorrektion

- Sæsoneffekt** Økonomiske tidsserier opgjort på måneds- eller kvartalsbasis er ofte påvirket af fænomener, der optræder på samme tid hvert år. Man siger tidsserien er påvirket af **sæsoneffekter**. Et eksempel er husholdningernes større indkøb op til jul med resultatet – sæsoneffekten – at detailhandelen stiger fra oktober til november og fra november til december, mens den falder fra december til januar. Ved sæsonkorrektion forsøger man at fjerne sæsoneffekterne fra tidsserien.
- Kalendereffekt** Endvidere er tidsserier påvirket af kalenderens konkrete ugedagssammensætning. Fx vil produktionen som regel være højere i en måned med mange arbejdsdage end i måned med mange lørdage, søndage og helligdage. Denne effekt betegnes **kalendereffekt**. Den er aktuel i serier, hvor det (i en månedsserie) betyder noget hvilke dage der er fire af, og hvilke der er fem af i løbet af måneden. Påskens placering i enten marts eller april (eller i første eller andet kvartal) er et andet eksempel på en kalendereffekt. For at kunne sammenligne på hinanden følgende måneder kan man standardisere tidsserien for disse kalendereffekter. Denne såkaldte kalenderkorrektion er en del af forhåndskorrektionen, der forbereder tidsserien til selve sæsonkorrektionen.
- Sæsonkorrektionsprocessen** Den samlede sæsonkorrektionsproces består altså af
- forhåndskorrektionen (inklusiv eventuel kalenderkorrektion)
 - sæsonkorrektion
- og resulterer i **den sæsonkorrigerede serie**, hvor tilstødende måneder/kvartaler er nemmere at sammenligne, da kalender- og sæsoneffekterne er fjernet.
- Sæsonkorrektionen laves i Danmarks Statistik primært ved metoden X-12-ARIMA, hvorfor beskrivelsen baseres på denne.

Forhåndskorrektion

Formålet med forhåndskorrektionen er at søge at fjerne de variationer i rådata (den ukorrigerede serie), der skyldes kalendereffekter (arbejdsdags-, handelsdags- og påskeeffekter) og ekstreme observationer.

- Arbejdsdagskorrektion** Antallet af arbejdsdage kan variere fra måned til måned, og dette kan have stor betydning for økonomiske tidsserier. Antallet af arbejdsdage (mandag til fredag) i en måned kan således variere mellem 18 og 23 inden for et år. Dette kan – hvis vi ikke tager hensyn til det – føre til fejlagtige konklusioner vedrørende udviklingen i f.eks. produktiviteten. At korrigere for dette forhold benævnes arbejdsdagskorrektion.
- Handelsdagskorrektion** Vi ved også at intensiteten i f.eks. detailhandelen er forskellig på de forskellige ugedage. Lørdag er således erfaringsmæssigt en større handelsdag end mandag. Dette betyder, at vi i nogle serier, hvor handelsdage er relevante, vil se en

større handelsintensitet i måneder hvor der er fem lørdage end i måneder med fire lørdage. At korrigere for dette forhold betegnes handelsdagskorrektion.

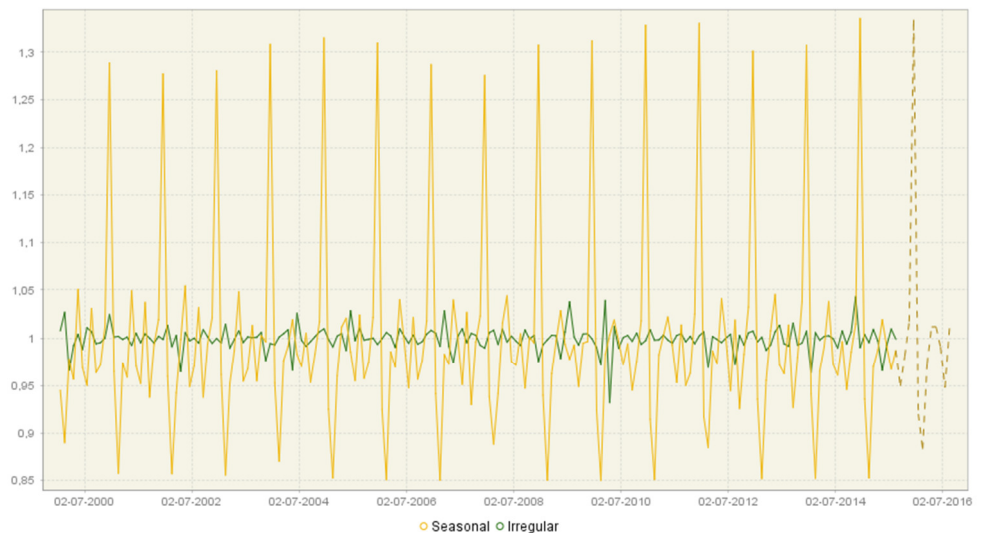
Påskekorrektion	Udover arbejds- og handelsdage kan de bevægelige helligdage – og her især påskens placering – have betydning for den økonomiske aktivitet. Påsken kan ligge i marts, i april eller delvist i marts og april. Tilsvarende kan den i kvartals-serier falde enten i første kvartal, i andet kvartal eller i begge. At korrigere for dette betegnes påskekorrektion.
Ekstreme observationer	Under forhåndskorrekturen undersøges og korrigeres rådata endvidere for ekstreme observationer. De tre vigtigste typer af disse er <ul style="list-style-type: none">- Additiv outlier er en ekstremværdi, som findes på et tidspunkt (måned eller kvartal), men ikke i de omkringliggende tidspunkter (før og efter). Et eksempel er korte strejker- Niveauskift skyldes fænomener, som påvirker rådata, sådan at de permanent kommer til at ligge på et højere/lavere niveau. Et eksempel kunne være salget af biler ved en nedsættelse af registreringsafgiften- Midlertidigt skift er et fænomen, som påvirker rådata på et tidspunkt, hvorefter denne påvirkning gradvist aftager over den følgende periode. Et eksempel er en prisstigning på kaffe, der vil medføre et øjeblikkeligt fald i den solgte mængde, men forventeligt vil stabilisere sig på det tidligere niveau igen efter et stykke tid.
Model til forhåndskorrektion	Forhåndskorrekturen laves ved en regARIMA model, som er en regressionsmodel, hvor residualen (den forhåndskorrigerede serie) modelleres ved en ARIMA model. En ARIMA model er en tidsserie model, som fanger tidsafhængigheden imellem perioderne. Matematisk kan det skrives som $X_t = \sum \beta_i x_{it} + O_t$ hvor X_t er rådata (den ukorrigerede serie) til tid t $\sum \beta_i x_{it}$ er forhåndskorrekturen til tid t O_t er den forhåndskorrigerede serie til tid t
Indhold af modellen	Med variabelen x_{it} angives et mål for den effekt, der korrigeres, f.eks. antallet af handelsdage i måned t , mens parameteren β_i angiver effektens indflydelse på serien. Regressionsdelen kan omfatte korrektion for flere effekter (angivet ved indekset i), og således bliver $\sum \beta_i x_{it}$ den samlede korrektion og O_t og den forhåndskorrigerede værdi til tid t .
Midlertidige og permanente korrektioner	Teknisk set foretages korrektionen for kalendereffekter og ekstreme observationer på samme måde, men der er alligevel en vigtig forskel. Korrektionen for de ekstreme observationer er nemlig kun midlertidig, idet den ophæves når sæsonkomponenten er identificeret (se næste afsnit). Korrektionen for kalendereffekterne er til gengæld permanent og disse effekter er altså også fjernet i den endeligt sæsonkorrigerede serie.
Hvilke effekter korrigeres?	Ved udarbejdelsen af de enkelte statistikker vurderes det, hvilke korrektioner der er relevante. Dette sker primært på baggrund af en række statistiske tests, men kalendereffekter bør derudover have en fornuftig økonomisk fortolkning. I tilfælde af, at der ikke konstateres kalendereffekter eller ekstreme observationer, falder forhåndskorrekturen bort og den forhåndskorrigerede serie er lig rådata ($X_t = O_t$).

Sæsonkorrektion

Dekomponering Efter forhåndskorrektionen fortsættes med selve sæsonkorrektionen. Denne tager udgangspunkt i den forhåndskorrigerede serie, som man ønsker at delte ind i tre komponenter (man dekomponerer serien):

- **Trenden** (T_t) repræsenterer den langsigtede tendens. Til sidst i processen lægges eventuelle niveauskift til denne komponent.
- **Sæsonkomponenten** (S_t) beskriver den del af variationen i en tidsserie, som sker inden for et år. Disse årlige bevægelser er mere eller mindre stabile over tid, hvad angår retning og størrelse. Altså er det (mere eller mindre) faste måneds/kvartals-effekter.
- **Den irregulære komponent** (I_t) er den del af de observerede rådata, der ikke indgår i trenden eller sæsonkomponenten (eller er blevet taget ud i forhåndskorrektionen). Den omtales ofte som tilfældig variation eller uforklaret variation. Til sidst i processen lægges eventuelle ekstremværdier (dog ikke niveauskift) til denne komponent.

Figuren nedenfor viser sæsonkomponenten og den irregulære komponent for en økonomisk tidsserie. Vi ser, at der ikke er nogen tydelig systematik tilbage i den irregulære komponent, men at sæsonkomponenten har et meget tydeligt gentaget mønster.



Modelvalg ved dekomponering

Som hovedregel antages det at tidsserien kan skrives enten som sum eller som produkt af de tre komponenter:

$$O_t = T_t + S_t + I_t \quad \text{additiv model}$$

$$O_t = T_t \cdot S_t \cdot I_t \quad \text{multiplikativ model}$$

hvor O_t er den forhåndskorrigerede serie til tid t (evt. blot rådata)

Hvis størrelsen på sæsonkomponenten er niveuafhængigt, da vælges ofte en multiplikativ model, dette testes automatisk det valgte software.

Rent teknisk foregår dekomponeringen som følger:

- i forbindelse med forhåndskorrektionen bliver den forhåndskorrigerede serie modelleret ved en ARIMA model

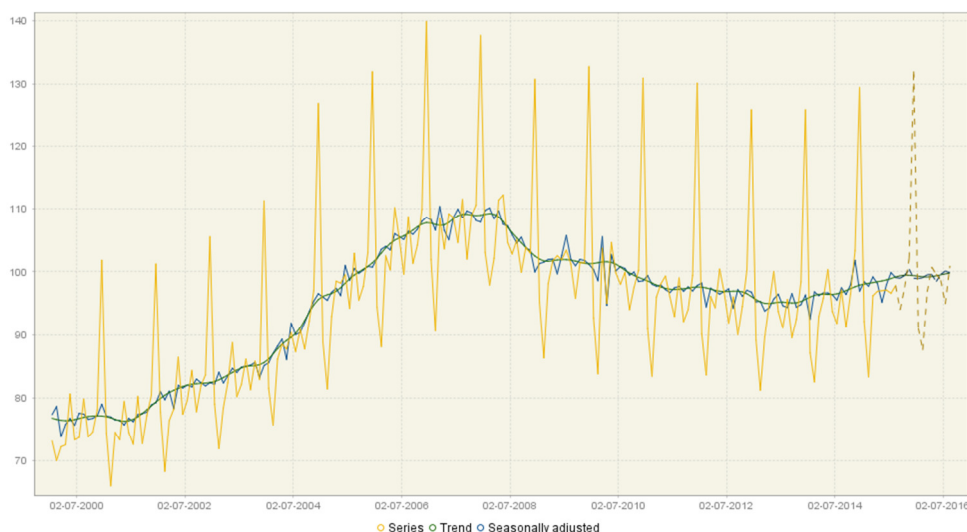
- den forhåndskorrigerede serie fremskrives ved hjælp af den fundne ARIMA model, dvs. der laves indirekte en prognose
- de tre komponenter bestemmes i en iterativ proces, der bygger på anvendelse af forskellige typer glidende gennemsnit af den således udvidede forhåndskorrigerede serie

Når modeltypen er fastlagt og komponenterne bestemt, findes den **sæsonkorrigerede series** (A_t) ved at fjerne sæsonkomponenten fra den forhåndskorrigerede serie:

$$A_t = O_t - S_t = T_t + I_t \quad \text{i tilfælde af additiv model}$$

$$A_t = O_t/S_t = T_t \cdot I_t \quad \text{i tilfælde af multiplikativ model}$$

Figuren nedenfor viser den faktiske serie (ikke-sæsonkorrigeret), trenden og den sæsonkorrigerede serie for den samme økonomiske tidsserie, som blev vist tidligere. Da der er meget tydelige og stærke sæsoneffekter bliver der justeret meget i dette tilfælde.



Betydning af fremskrivning

Fremskrivningen af serien er nødvendig da der benyttes centrerede glidende gennemsnit til at dekomponere serien. Fremskrivningen af serien har en afgørende betydning for sæsonkorrektionen. I perioder hvor serien ændrer sig meget, f.eks. når der indtræffer vendepunkter i en økonomi, bliver fremskrivningen mere usikker. Den usikkerhed overføres direkte til sæsonkorrektionen, hvilket man bør tage hensyn til ved fortolkningen af sæsonkorrigerede data.

Øvrige aspekter

Sæsonkorrektion af sumserier

Hvis en serie der er sum af flere underserier skal sæsonkorrigeres, kan dette gøres på to forskellige måder. **Direkte sæsonkorrektion** er anvendt, hvis sumserien og underserierne sæsonkorrigeres hver for sig. I dette tilfælde vil den sæsonkorrigerede sumserie normalt ikke svare til summen af de sæsonkorrigerede underserier. **Indirekte sæsonkorrektion** er anvendt, hvis den sæsonkorrigerede sumserie beregnes som summen af de sæsonkorrigerede underserier.

Årsopregning

Sæsonkorrektion med X-12-ARIMA giver ikke automatisk konsistens mellem summer af den sæsonkorrigerede serie og tilsvarende summer af den originale

serie. Normalt vil forskellen inden for et kalenderår være lille. **Årsopregning** dækker over, at man efter dekomponeringen korrigerer den sæsonkorrigerede serie, således at den har samme sum som den oprindelige serie inden for hvert kalenderår. Alternativt kan man lave årsopregning mellem den sæsonkorrigerede serie og den kalenderkorrigerede serie. Det er de færreste statistikker, der årsopregnes.

Revisioner De sæsonkorrigerede tal bliver som regel revideret ved ny publicering. Når rådata suppleres med nye måneds- eller kvartalstal, foretages sæsonkorrektionen i princippet på et nyt datamateriale, hvilket kan føre til, at tidligere offentliggjorte tal ændres. I enkelte tilfælde kan resultatet være, at vækstrater og tendenser observeret inden for de seneste måneder bliver ændret.

Kvaliteten af sæsonkorrektion Det første spørgsmål, der skal afklares, når en tidsserie skal sæsonkorrigeres, er, om der er en sæsonstruktur i serien. Hvis ikke er det Danmarks Statistiks politik ikke at sæsonkorrigerer serien, men kun at bringe den originale serie, eventuelt den forhåndskorrigerede serie.

Hvis der er sæsonstruktur i serien kan serien sæsonkorrigeres, og her er X-12-ARIMA forsynet med en række diagnostiske mål, som belyser forskellige aspekter af kvaliteten af den foretagne sæsonkorrektion. Disse mål bruges til at vurdere sæsonkorrektionsens kvalitet eventuelt til at undersøge anvendelsen af alternative modeller.

Software I Danmarks Statistik benyttes der Demetra 2.0 og X-12 (dos version 0.3). Fra starten 2016 begynder produktionen delvist at overgå til det nye software JDemetra+. I Dette software hedder metoden X-13-ARIMA men afviger ikke fra beskrivelsen i dette dokument.