

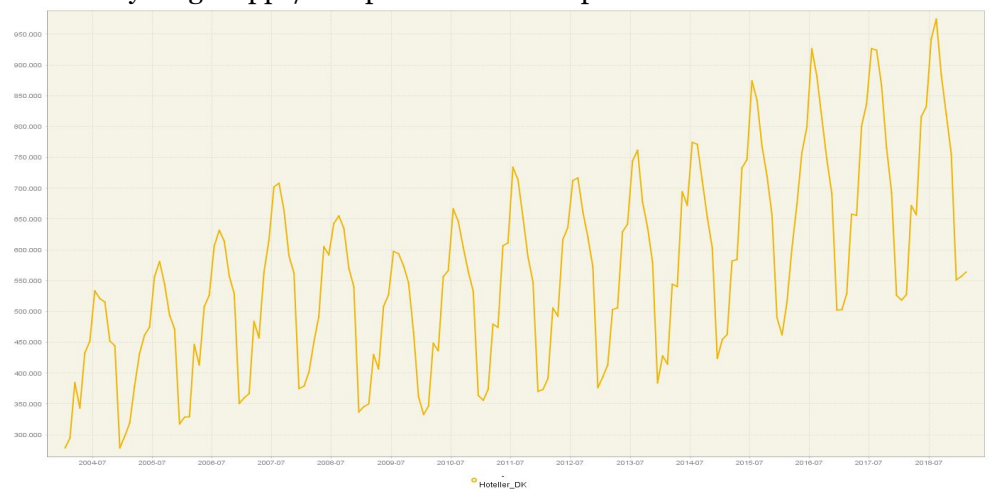
Introduktion til sæsonkorrektion

Hvorfor sæsonkorrigere?

En tidsrække med sæson består af forskellige mønstre og tilfældige udsving, som svarer til korte, irregulære og sæsonbestemte udsving (typisk under et år), mellem lange forretnings-cykluser (3 – 5 år) og lange bevægelser (over 5 år).

Man sæsonkorrigerer for at kunne studere den udvikling i tidsrækken, der ikke kan tilskrives sæsonbestemte udsving eller **sæsoneffekter**.

Figuren nedenfor viser en tidsserie med tydelige sæsonmønstre, dvs. der forekommer tydelige toppe/dale på de samme tidspunkter hvert år.



Hvad er sæsonkorrektion

- Sæsoneffekt** Økonomiske tidsrækker opgjort på måneds- eller kvartalsbasis er ofte påvirket af fænomener, der optræder på samme tid hvert år. Man siger, at tidsrækken 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 tidsrækken.
- Kalendereffekt** Endvidere er tidsrækker 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 tidsrækker, hvor det (i en månedstidsrække) 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.
- Forhåndskorrektion** For at kunne sammenligne på hinanden følgende måneder kan man standardisere tidsrækken for disse kalendereffekter. Denne såkaldte kalenderkorrektion er en del af **forhåndskorrektionen**, der forbereder tidsrækken til selve sæsonkorrektionen.

Sæsonkorrektionsprocessen Den samlede sæsonkorrektionsproces består altså af

- forhåndskorrektionen (inklusive eventuel kalenderkorrektion)
- sæsonkorrektion

og resulterer i **den sæsonkorrigerede tidsrække**, hvor tilstødende måneder/kvartaler er nemmere at sammenligne, da kalender- og sæsoneffekterne er fjernet.

Sæsonkorrektionen laves i Danmarks Statistik primært ved metoden X-13-ARIMA, hvorfor beskrivelsen baseres på denne.

Forhåndskorrektion

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

Arbejdsdagskorrektion Antallet af arbejdsdage kan variere fra måned til måned, og dette kan have stor betydning for økonomiske tidsrækker. 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 tidsrækker, 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 kvartalstidsrækker falde enten i første kvartal, i andet kvartal eller i begge. At korrigere for dette betegnes **påskekorrektion**.

Ekstreme observationer Under forhåndskorrektionen undersøges og korrigeres rådata endvidere for ekstreme observationer. De tre vigtigste typer af disse er:

- **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 (level shift)** 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 (transitory change)** 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.

Da beregningerne af de seneste sæsonkorrigerede værdier er baseret på værdier både bagud og fremad i tid ("forecast's") vil uforudsete hændelser (outliers) føre

til en større usikkerhed omkring bestemmelsen af den/de seneste sæsonkorrigerede værdier. Specielt når de seneste observerede værdier er afvigende i forhold til det ellers forventede forløb, så kan dette give anledning til, at de forecastede værdier ikke reelt afspejler vores forventning til fremtiden. Når de forecastede værdier ved næste produktionscyklus bliver erstattet af observerede værdier, så risikerer man derfor store revisioner af allerede offentliggjorte sæsonkorrigerede tal.

Model til forhåndskorrektion Forhåndskorrekturen laves ved en regARIMA-model, som er en regressionsmodel, hvor residualen (den forhåndskorrigerede tidsrække) modelleres ved en ARIMA-model. En ARIMA-model er en tidsrække-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 tidsrække) til tid t

$\sum \beta_i x_{it}$ er forhåndskorrekturen til tid t

O_t er den forhåndskorrigerede tidsrække 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 tidsrækken. 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, da disse ekstremværdier beskriver virkelige hændelser (se næste afsnit). Korrektionen for kalendereffekterne er til gengæld **permanent**, og disse effekter er altså også fjernet i den endeligt sæsonkorrigerede tidsrække.

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 de forhåndskorrigerede tidsrække er lig rådata ($X_t = O_t$).

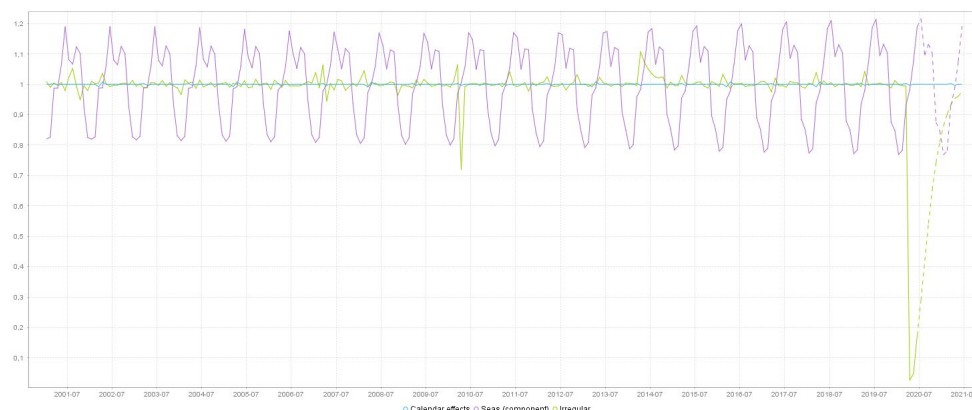
Sæsonkorrektion

Dekomponering Efter forhåndskorrekturen fortsættes med selve sæsonkorrektionen. Denne tager udgangspunkt i den forhåndskorrigerede tidsrække, som man ønsker at dele ind i tre komponenter (man dekomponerer tidsrækken):

- **Trenden** (T_t) repræsenterer den langsigtede tendens. Til sidst i processen tilbageføres eventuelle niveauskift til denne komponent (se afsnit om outliers ovenfor).
- **Sæsonkomponenten** (S_t) beskriver den del af variationen i en tidsrække, 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 forhåndskorrekturen). 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 tidsrække. 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 tidsrækken 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 tidsrække til tid t (evt. blot rådata)
Hvis størrelsen på sæsonkomponenten er niveauafhængigt, vælges ofte en multiplikativ model, dette testes automatisk af det valgte software.

Rent teknisk foregår dekomponeringen som følger:

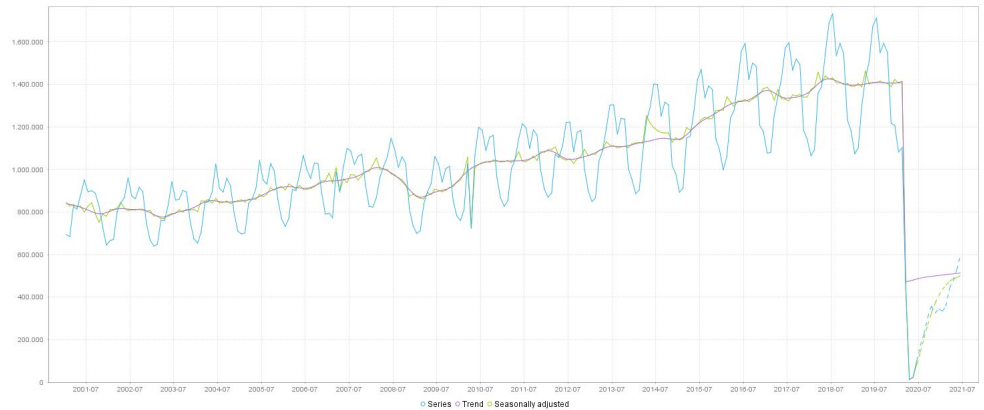
- i forbindelse med forhåndskorrekturen bliver den forhåndskorrigerede tidsrække modelleret ved en ARIMA-model
- den forhåndskorrigerede tidsrække fremskrives ved hjælp af den fundne ARIMA-model, dvs. der laves indirekte en prognose
- de tre komponenter bestemmes i en iterativ proces, der bygger på anvendelse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

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

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

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

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



Betydning af fremskrivning

Fremskrivningen af tidsrækken er nødvendig, da der benyttes centrerede glidende gennemsnit til at dekomponere tidsrækken. Fremskrivningen af tidsrækken har en afgørende betydning for sæsonkorrektionen. I perioder, hvor tidsrækken ændrer sig meget, f.eks. når der indtræffer vendepunkter i en økonomi eller som konsekvens af COVID19 krisen, bliver fremskrivningen mere usikker. Den usikkerhed overføres direkte til sæsonkorrektionen, hvilket man bør tage hensyn til ved fortolkningen af sæsonkorrigerede data (jf. også afsnittet om ekstreme observationer).

Øvrige aspekter

Sæsonkorrektion af sumserier

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

Årsopregning

Sæsonkorrektion med X-13-ARIMA giver ikke automatisk konsistens mellem summen af den sæsonkorrigerede tidsrække og tilsvarende summen af den originale tidsrække. Normalt vil forskellen inden for et kalenderår være lille. **Årsopregning** dækker over, at man efter dekomponeringen korrigerer den sæsonkorrigerede tidsrække, således at den har samme sum som den oprindelige tidsrække inden for hvert kalenderår. Alternativt kan man lave årsopregning mellem den sæsonkorrigerede tidsrække og den kalenderkorrigerede tidsrække. Det er de færreste statistikker, der årsopregnes. Det bør bemærkes at det generelt ikke anbefales at foretage årsopregning.

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. I Danmarks Statistik er anbefalingen at der foretages revisioner for det indeværende kalenderår og tre år tilbage.

Kvaliteten af sæsonkorrektion

Det første spørgsmål, der skal afklares, når en tidsrække skal sæsonkorrigeres, er, om der er en sæsonstruktur i tidsrækken. Vurdering om der er sæsonstruktur gøres ud fra en række tests og grafer. For en del tidserier er sæsonmønstrene

ikke så tydelige og det kan i nogle tilfælde være vanskeligt at vurdere om en tids-
serie reelt indeholder sæson.

Hvis ikke der er vurderes at være sæsonstruktur, er det Danmarks Statistiks poli-
tik **ikke** at sæsonkorrigerer tidsrækken, men kun at bringe den originale tids-
række, eventuelt den forhåndskorrigerede tidsrække. Dette er i overensstem-
melse med Eurostats generelle ”guidelines”, som i øvrigt opstiller en række an-
befalinger i forbindelse med sæsonkorrektion (herunder forhåndskorrektion, in-
direkte/direkte sæsonkorrektion, årsopregning, revisioner etc.). Dokumentet
kan downloades fra Eurostats hjemmeside ([https://ec.europa.eu/euro-
stat/cros/content/ess-guidelines-sa-2015-edition_en](https://ec.europa.eu/eurostat/cros/content/ess-guidelines-sa-2015-edition_en)).

Hvis der er sæsonstruktur i tidsrækken, kan tidsrækken sæsonkorrigeres, og her
er X-13-ARIMA forsynet med en række diagnostiske mål, som belyser forskellige
aspekter af kvaliteten af den foretagne sæsonkorrektion. Disse mål bruges til at
vurdere sæsonkorrektionskvaliteten, eventuelt til at undersøge anvendelsen af
alternative modeller. Hvis kvaliteten af sæsonkorrektionen er meget dårlig bør
man ikke sæsonkorrigerer den pågældende tidsserie.

Software I Danmarks Statistik benyttes der næsten udelukkende sæsonkorrektionsprog-
rammet JDemetra+ (pt. version 2.2.2) til sæsonkorrektion, og som er det offici-
elt anbefalede program til sæsonkorrektion indenfor Eurostat. JDemetra er et
frit tilgængeligt software og kan downloades fra Eurostats hjemmeside
(https://ec.europa.eu/eurostat/cros/content/software-jdemetra_en).