

Sesongjustering – Avvikling av korona-håndtering

Bakgrunn

Mandag 30. mars 2020 ble notatet «Sesongjustering under Koronakrisen» distribuert fra Metodeseksjonen til fagavdelingene.

Notatet ga anbefalinger om spesiell sesongjustering og bakgrunnen for dette ble gjennomgått.

Seinere, 17. desember 2020, kom et nytt notat som påpekte at krisehåndteringen med outliere skulle fortsette i 2021.

I september forsvant de fleste nasjonale koronatiltakene og vi planla å normalisere sesongjusteringen, men det ble en ny runde med tiltak.

Lørdag 12. februar forsvant igjen de fleste nasjonale koronatiltakene. Nå er det endelig snart på tide å avslutte spesiell håndtering av sesongjustering. Nedenfor er informasjon om avviklingen. Etterpå diskuteres noen problemstillinger litt nærmere. Legg spesielt merke til at vi, for outlier-deteksjon, anbefaler at faste kritiske t-verdi-grenser tas i bruk.

Mars eller 1. kvartal 2022 blir siste korona-outlier

Vi tar utgangspunkt i at metoden som har vært i bruk i SSB er løpende korrigering med outliere. Siste måned man skal ha korona-outlier er den første måneden som regnes som normal. Dette vil sikre at trenden etter kriseperioden har riktig utgangspunkt. Når april-tallet er på plass er det viktig at mars-outlieren omgjøres til nivåskift (LS) dersom punktekstremverdier var spesifisert (AO).

For kvartalsvise statistikker er anbefalingen at 1. kvartal 2022 blir siste kvartal med korona-outlier. Når 2. kvartal 2022 er på plass skal 1. kvartal 2022 omgjøres til nivåskift (LS).

Begrunnelsen for valget av mars er at vi regner med at situasjonen er normalisert og at mars er en helt normal måned. Strengt tatt ville en tilsvarende vurdering ført til at 2. kvartal 2022 også skulle bli outlier (første helt normale kvartal). Men avveiningen av fordeler og ulemper gjør at valget lander på 1. kvartal 2022 som siste outlier. Dette vil si at kvartalsstatistikker har hatt mer enn to hele år med outliere.

Dette er en felles anbefaling. Avviklingen kan settes dersom statistikkansvarlig har kunnskap som tilsier dette. Spesielt kan det være aktuelt å legge til enda en outlier for kvartalsstatistikker. Dette

har også sammenheng med hvordan tallene for kvartalsstatistikker er fremkommet. Denne problemstillingen kan vi diskutere nærmere på et sesongjusteringsforum.

Flere detaljer

Trend gjennom koronaperioden

Beregning av trend har opphørt under koronaperioden. Dersom trend var ønskelig, var løsningen å bruke nivåskift for alle outliere slik at trend og sesongjusterte tall ble sammenfallende. Bedre trendestimer kan i ettertid oppnås med en re-modellering av hele koronaperioden, inkludert sesongjusteringen. Dette vil kreve manuelt arbeid. Hvis dette skal gjøres er det bedre å vente med slik modellering til flere observasjoner er på plass.

En enklere tilnærming er å beholde sesongjusterte tall og heller finne en trend ved å glatte de sesongjusterte tallene i etterhånd. Det kan f.eks. velges glidende gjennomsnitt (3 eller 5 mnd.) eller et annet glattingsfilter.

Annet sesongmønster etter koronaperioden?

Sesongjustering handler om å korrigere for et kjent sesongmønster. Dersom sesongmønsteret ikke er kjent, så kan vi heller ikke sesongjustere. Det kan innebære at vi må ta pause, minst tre år, fra sesongjusteringen.

På den annen side er gradvise endringer av sesongmønsteret noe som er innebygget i sesongjusteringsmetoden. Det velges automatisk et sesongfilter ut fra seriens lengde og egenskaper. Jo kortere filteret er, desto mer brå endringer i sesongmønsteret håndteres.

Dersom en serie som har hatt et stabilt sesongmønster får en relativt brå endring etter koronaperioden, kan det medføre at automatikken etter hvert vil endre filterlengden. Men det vil kanskje ta to år før informasjonen er tydelig nok til automatikken gjør endring.

Dersom man er sikker på at sesongmønsteret endres kraftigere enn før, går det an å ligge i forkant. Automatikken kan overstyres i den forstand at sesongfilteret kan spesifiseres manuelt, f.eks. ved `x11{seasonalma = S3X3}`, slik at et kort filter velges. Samlet sett kan en manuell styring av sesongfilteret gi mindre revisjon enn ellers, forutsatt at antagelsen om at en relativt stor endring i sesongmønsteret stemmer.

Det enkleste er å fortsette uten slik overstyring og i mange tilfeller vil det også være det beste.

Hva om det blir flere koronatiltak?

Perioden vi har vært igjennom har vært ekstraordinær. Vi kan ikke vite hva fremtiden innebærer angående koronaviruset eller andre virus og hvor ofte det vil være større eller mindre tiltak. En stor nasjonal alarm må til dersom metodeseksjonen igjen skal komme med en fellesanbefaling for

sesongjusteringen. Det normale er at de statistikkansvarlige har ansvaret for egne serier og følger med på utviklingen som påvirker disse.

Fremskrevne sesongfaktorer

Dersom fremskrevne sesongfaktorer har vært brukt, må sesongjusteringsmodellene etter hvert oppdateres. Hele koronaperioden kan håndteres ved å legge inn outliere på samme måte som for andre serier i SSB. En mer manuell modellering av hele koronaperioden kan også tenkes, slik som nevnt under avsnittet om trend. Da kan man vurdere å gjøre bruk av «temporary change» (TC). Det bør ventes med evt. slik modellering til flere observasjoner er på plass.

Outlier-deteksjon og t-verdi-feil

Når det gjelder andre outliere enn korona-outliere, så er beste praksis å fryse fast gamle outliere og bare se etter nye på slutten av serien (det siste året). Denne praksisen er bare i noen grad implementert i SSB. Etter korona-perioden, er det en god anledning til å gjøre en endring. Eventuelle outliere før korona bør fryses fast. I X-13ARIMA-SEATS betyr dette at gamle outliere kodes inn regression-bolken sammen med korona-outlierne og at outlier-span defineres til å starte første observasjon etter korona-perioden. I store systemer med svært mange serier kan det være praktisk vanskelig å gjennomføre slik fastfrysing. Videre drøfting av denne problemstillingen ligger utenfor rammen av dette dokumentet.

Vi anbefaler samtidig å sette en fast kritisk deteksjonsgrense. Helt konkret anbefaler vi at outlier-bolken kodes slik:

Månedsserier:

```
Regression{
  variables=(LSS2020.3-2022.3)
}
outlier{
  critical = 4.25
  span = (2022.4, )
}
```

Kvartalsserier:

```
Regression{
  variables=(LSS2020.1-2022.1)
}
outlier{
  critical = 4.00
  span = (2022.2, )
}
```

Her er det også vist, i grå skrift, hvordan det er mulig, i siste versjon av X-13ARIMA-SEATS, å kode serien av korona-outliere på en kompakt måte. I eksemplene over forutsettes det at alle ekstremverdier før Korona også inkluderes/hardkodes i regression-bolken. Det er avhengig av de enkelte tidsseriene, og er derfor ikke tatt med her.

Årsaken til at disse kritiske grensene anbefales er to-delt:

1. Den automatiske kritiske grensa avhenger av antall observasjoner som undersøkes. Dette betyr svært lave grenser for korte span. Selv om hypotesetestingsteorien bak dette er riktig, så er den faktiske konsekvensen at outlier-terskelen senkes enormt ved frysing. Dette er noe vi ikke anbefaler, og det er spesielt uheldig med tanke på at automatisk endring av outlier-valg kan medføre revisjoner. For lange serier er de vanlige grenseverdiene ca. 4,00 eller 3,75 for henholdsvis måneds- og kvartalsserier.

2. Vi har tidligere (14. april 2021) påpekt at en feil i programmet X-13ARIMA-SEATS medfører en systematisk feil i t-verdier når vi har korona-outliere. Dette medfører lavere terskel for automatiske outliere, spesielt for korte tidsserier. Det går an å gi spesifikke beregninger for dette. En forenkling er å innføre en tommelfingerregel om at grensa bør økes med 0,25 slik at anbefalingen blir 4,25 eller 4,00.

Programmet JDemetra+ har ikke den systematiske feilen i t-verdier. Ved bruk av dette programmet anbefaler vi faste kritiske grenser på 4,00 eller 3,75 for henholdsvis måneds- og kvartalsserier. Hvor stor konsekvens denne anbefalingen har i praksis er faktisk litt uklart. Det er nemlig et misforhold mellom hva dokumentasjonen sier (at grense avhenger av span-lengde) og hva programmet faktisk gjør. Vi regner med at misforholdet rettes i fremtidige versjoner, men hva rettingen innebærer vet vi ikke.

Om man ikke tar i bruk frysing av outliere og outlier-span anbefaler vi allikevel at faste kritiske grenser benyttes. Dette gjelder spesielt ved bruk av X-13ARIMA-SEAT pga. den systematiske feilen.

Kontakt Øyvind Langsrud (oyl@ssb.no)