

Sammendrag

Utvikling av ulykkesmodeller for ulykker på riks- og fylkesvegnettet i Norge (2010-2015)

TØI rapport 1522/2016

Forfatter: Alena Høye

Oslo, 2016, 5 sider

Ulykkesmodeller er utviklet for riks- og fylkesvegnettet i Norge for å beregne normale antall personskadeulykker, lett skadde, hardt skadde, drepte og samlet antall drepte eller hardt skadde som en funksjon av trafikkmengde, strekningslengde, vegtype, fartsgrense, antall kjørefelt, antall kryss, ATK, vegbelysning og andre vegegenskaper. Den valgte modellformen (negative binomialmodeller med variabel overspredningsparameter) gjør det mulig å benytte resultatene bl.a. i evalueringer av trafikksikkerhetstiltak med den empiriske Bayes metoden som i dag anses som den beste for å redusere risikoen for å systematisk overestimere virkninger av trafikksikkerhetstiltak i før-etter studier. Resultatene kan også brukes i ulike verktøy som brukes av Statens vegvesen for å gjøre analyser av ulykkes situasjonen i vegnettet og for å beregne forventede virkninger av trafikksikkerhetstiltak.

Ulykkesmodellene som er utviklet på oppdrag av Vegdirektoratet er beregnet med ulykkesdata fra den nasjonale vegdatabanken (NVDB) fra 2010-2015. Rapporten bygger på tidligere rapporter om utvikling av ulykkesmodeller for det norske vegnettet (Ragnøy, Christensen & Elvik, 2002; Ragnøy & Elvik, 2003; Høye, 2014). Ragnøy et al. (2002) har utviklet ulykkesmodeller og en metode for å vekte sammen registrerte og normale ulykkestall til forventede ulykkestall. Basert på denne metoden har Ragnøy & Elvik (2003) gjort en trafikksikkerhetsanalyse av stamvegnettet. Høye (2014) har utviklet modeller som bygger på ulykkesdata fra 2006-2011 og en videreutvikling av modellene og metodene som er brukt av Ragnøy et al. (2002). Den aktuelle rapporten presenterer en oppdatering av modellene som er utviklet av Høye (2014).

Modellene kan brukes for å beregne normale antall personskadeulykker (Psu), lett skadde (LS), hardt skadde (HS), drepte og det samlede antall drepte eller hardt skadde (D/HS) på riks- og fylkesvegnettet. Normale ulykkestall beregnes som en funksjon av trafikkmengden og ulike vegegenskaper. Et normalt ulykkestall er antall ulykker som skjer i gjennomsnitt på en veg av en gitt lengde med en gitt trafikkmengde og gitte vegegenskaper i ett gitt tidsrom. Modellformen som er benyttet (negative binomialmodeller med variabel overspredningsparameter) gjør det også mulig å beregne forventede ulykkestall som en funksjon av normale ulykkestall, registrerte ulykkestall og overspredningsparameteren. Forventede ulykkestall er antall ulykker som man på lang sikt forventer at vil skje på en vegstrekning, ut fra generelle vegegenskaper og ulykkene som har skjedd på strekningen. Overspredningsparameteren brukes for å vekte sammen normale og registrerte ulykkestall. Overspredningsparameteren varierer som en funksjon av trafikkmengde og strekningslengde, slik at registrerte ulykkestall får en større vekt på vegstrekninger med store normale ulykkestall.

Når man beregner det forventede antall ulykker på en strekning med unormalt mange ulykker antar man at resultatet viser hvor mange ulykker man kan forvente på strekningen på lang sikt. Man tar dermed hensyn til at antall ulykker kan ha vært tilfeldig høyt og at antall ulykker på langt sikt mest sannsynlig vil ligge nærmere det normale ulykkestallet. Ved å benytte forventede ulykkestall, istedenfor registrerte ulykkestall, kan man i en før-etter studie unngå å overestimere virkningen av trafikksikkerhetstiltak. Tiltak settes ofte inn på vegstrekninger hvor det har vært spesielt mange ulykker og derfor vil antall ulykker i mange tilfeller gå ned (komme nærmere gjennomsnittet) uansett om man setter inn et tiltak eller ikke.

Modell og prediktorvariabler

Modellformen som er brukt for alle skadegrader er en negativ binomialmodell (NB modeller) med variabel overspredningsparameter. Modellformen er følgende:

$$E(n) = e^{\sum_i \text{Prediktor}_i * \text{Koeff}_i}$$

$E(n)$ er det predikerte (normale) ulykkestall (dvs. antall Psu, LS, HS, drepte eller D/HS), prediktorene er trafikkmengden og ulike vegegenskaper og i er subskript for prediktorene. I tillegg inneholder modellen prediktorer og koeffisienter for overspredningsparameteren.

Fordelen med denne modellformen er at den tar hensyn til overspredningen i ulykkesdata og at overspredningen varierer som en funksjon av trafikkmengde, segmentlengde og antall år. Dermed er det mulig å benytte resultatene bl.a. i evalueringer av trafikksikkerhetstiltak med den empiriske Bayes (EB) metoden. Denne anses i dag som den beste for å redusere risikoen for å systematisk overestimere virkninger av trafikksikkerhetstiltak i før-etter studier. I en EB studie sammenligner man det registrerte ulykkestallet etter at et tiltak er satt inn på en veg med det forventede ulykkestallet. Det forventede ulykkestallet beregnes ved å vekte sammen det registrerte ulykkestall og det normale ulykkestall. Det normale ulykkestall er antall ulykker som er predikert av modellen, dvs. det gjennomsnittlige antall ulykker på en vegstrekning med de gitte vegegenskapene. Vekten er en funksjon av overspredningsparameteren og det normale antall ulykker. Det forventede antall ulykker ligger alltid mellom det normale og det registrerte antall.

Prediktorvariablene i modellene for alle skadegrader er følgende:

- **Segmentlengde og antall år:** Segmentlengde og antall år inngår i alle modellene som eksponeringsvariabler, slik at det normale antall ulykker alltid øker proporsjonalt med segmentlengde og antall år. Antall år er prediktor fordi det ikke foreligger data fra alle år for alle segmentene; segmenter med vesentlige endringer (som for eksempel endringer av fartsgrensen) inngår kun med ulykkesdata fra etter endringen i datafilen som ligger til grunn for modellberegningene.
- **Trafikkmengde (ÅDT):** Trafikkmengden inngår i modellene som $\ln(\hat{ADT})$. Dermed er den prosentvise økningen av normale ulykkestall uavhengig av hvor høy trafikkmengden er i utgangspunktet. Hver økning av trafikkmengden med X prosent medfører alltid den samme prosentvise økningen for hver skadegrad.
- **Fartsgrense:** For hver fartsgrense er det definert en dummyvariabel som er én ved den respektive fartsgrensen og null ellers. Dummyvariabler er valgt framfor en tallvariabel for å ta hensyn til generelle forskjeller mellom veger med ulike fartsgrenser som gjør at antall ulykker kan gå opp og ned mellom ulike fartsgrenser.

- **Antall kjørefelt:** For antall kjørefelt er det definert dummyvariabler på samme måte og av de samme grunnene som beskrevet for fartsgrense. Segmenter med kun ett kjørefelt er utelatt fra datasettet fordi dette er en liten og svært heterogen gruppe av helt ulike typer veger og dermed ikke noen meningsfylt kategori.
- **Plankryss, rundkjøringer og ramper (planskilte kryss):** For X-kryss, T-kryss, rundkjøringer og ramper er det beregnet variabler som er den naturlige logaritmen av antallet per kilometer pluss én (pluss én for å unngå å ta logaritmen av null).
- **Type veg:** Det er definert fem dummyvariabler for ulike typer veg: 1 = motorveg; 2 = tofeltsveg med planskilte kryss (tidligere betegnet som motortrafikkveg); 3 = TEN-T-veg (ikke motorveg / tofeltsveg med planskilte kryss); 4 = øvrig europa-/riksveg (ikke motorveg / tofeltsveg med planskilte kryss / TEN-T-veg); 5 = fylkesveg. Disse vegklassene er valgt fordi det finnes generelle forskjeller i vegstandarden (bl.a. kjørefelt- og skulderbredde) mellom klassene; f.eks. har alle motorveger midtrekkverk og TEN-T-veger har generelt en noe høyere standard enn øvrige europa- og riksveger).
- **ATK:** Det er definert tre dummyvariabler for automatisk trafikkontroll (ATK), punkt-ATK (PATK), streknings-ATK (SATK) i én retning og SATK i begge retninger. For PATK er det ikke skilt mellom strekninger med PATK i en vs. begge retninger. Strekninger som ligger inntil 100 meter oppstrøms og 3 km nedstrøms for et ATK-punkt regnes som PATK-strekninger. For SATK regnes alle strekninger mellom de to fotoboksene som SATK-strekninger.
- **Vegbelysning:** Det er definert én dummyvariabel for vegbelysning.
- **Midtdeler / -rekkverk:** Alle segmentene er delt inn i segmenter med både midtdeler og midtrekkverk, segmenter med midtdeler og uten midtrekkverk, segmenter med midtrekkverk og uten midtdeler, og segmenter uten verken midtdeler eller midtrekkverk. For hver av de fire gruppene er det definert én dummyvariabel. Det er ikke definert to separate dummyvariabler (én for med/uten midtdeler og én for med/uten midtrekkverk) fordi det kan være interaksjonseffekter mellom midtdeler og midtrekkverk, dvs. at virkningen av midtrekkverk kan være forskjellig på veger med og uten midtdeler og at virkningen av midtdeler kan være forskjellig på veger med og uten midtrekkverk.
- **Forsterket midtoppmerking:** Forsterket midtoppmerking er kombinasjonen av vanlig oppmerking og rumleriller. Denne variabelen omfatter kun forsterket midtoppmerking på veger som ikke er motorveger og som ikke har midtdeler eller midtrekkverk. Det skilles ikke mellom ulike typer forsterket midtoppmerking.
- **Fylke:** For hvert fylke er det definert én dummyvariabel. Disse skal fange opp generelle forskjeller mellom fylkene (f.eks. forskjeller i topografi, vær og befolkningstetthet).
- **Konstantterm:** Alle modellene inneholder en konstantterm.

For hver av modellene er det i tillegg estimert koeffisienter for å beregne **overspredningsparameteren** som en funksjon av segmentlengde, antall år og trafikkmengde. Overspredningen synker med økende segmentlengde, antall år og trafikkmengde.

Hvor gode er modellene?

Det er beregnet flere ulike goodness-of-fit (GOF) indikatorer: Pseudo-R², mean square prediction error (MSPE) og Elvik-index. Når man sammenligner ulike modeller (med ulike prediktorer) viser GOF indikatorene kun svært små forskjeller mellom ulike modellvarianter for hver skadegrad. Sammenlignet med modellene av Høye (2014) er de aktuelle modellene omtrent like «gode» eller bedre. I motsetning til modellene til Høye (2014) er avvikene mellom de faktiske og predikerte totale ulykkestallene forholdsvis små for alle skadegrader. Sorterer man avvikene etter trafikkmengden er det likevel en del avvik mellom predikerte og registrerte ulykkestall i enkelte ÅDT-grupper.

Regneark Ulykkesmodeller

I vedlagte regneark Ulykkesmodeller 2016.xlsx kan man beregne:

- Predikerte (normale) antall Psu, LS, HS, drepte og D/HS som en funksjon av de prediktorvariablene som er beskrevet ovenfor
- En overspredningsparameter for hver skadegrad
- En vekt for hver skadegrad som kan brukes til å beregne forventede ulykkes/-skadetall
- Hvis man i tillegg oppgir registrerte ulykkes- og skadetall, beregnes også forventede antall Psu, LS, HS, drepte og D/HS.

Det er også mulig å omregne resultater til år mellom 1997 og 2030.