

Att dö eller att inte dö? Analys av trafikolyckors dödsutfall i Sverige

Pavel Lukashin*

Juni 2017

Sammanfattning

Den här uppsatsen har till syfte att förklara dödligt utfall av trafikolyckor. I denna studie används datamaterial för trafikolyckor i Sverige mellan åren 2005-2016. Genom logistisk regression har vi skattat sannolikheten att en trafikolycka leder till ett dödligt utfall givet olika förutsättningar som väg- och väderförhållandena, geografiska koordinater samt årstiden och inblandade typer av fordon. Då en del observationer saknar värden för vissa av de förklarande variabler används imputering för att estimeras och ersätta dessa värden. För att testa hur väl denna modell kan prediktera utfallet av framtida trafikolyckor används metoden korsvalidering, där man anpassar modellen på en del av datan och predikterar på en annan.

Analysen har visat att dödligheten i trafikolyckorna beror på inblandningen av tunga fordon, fotgängare, cyklister och övrigt trafikelement (se Figur 8), men att det även finns geografisk och temporal inverkan. Av de väder- och vägförhållanden som fanns med i den givna datamängden var det sikten som influerade responsutfallet mest.

Som prediktionsmodell visade det sig att modellen var överanpassad på vår datamängd. För att hantera överanpassning hade man kunnat använda regularisering, det vill säga olika sätt att hantera skattningar av de parametrar som modellen använder.

*Postadress: Matematisk statistik, Stockholms universitet, 106 91, Sverige.
E-post: p.lukashin@gmail.com. Handledare: Tom Britton, Benjamin Allévius.