

ISSN 2282-6599

# RIVISTA DI ECONOMIA E POLITICA DEI TRASPORTI

---

---

Anno 2018  
Numero 2

# R.E.PO.T



**SIET**

---

Rivista Scientifica della Società  
Italiana di Economia dei Trasporti e della  
Logistica

# ANALISI ECONOMETRICA DELL'INCIDENTALITÀ STRADALE: IL CASO DEL REGNO UNITO. I FATTORI SOCIO-ECONOMICI, TERRITORIALI, TECNOLOGICI.

**Mario Catalano<sup>1\*</sup>, Fabio Galatioto<sup>2</sup>**

<sup>1</sup> *Research fellow presso Transport Systems Catapult (United Kingdom)*

<sup>2</sup> *Senior Technologist presso Transport Systems Catapult (United Kingdom)*

Lo sviluppo di modelli per la previsione degli incidenti stradali può offrire a politici ed operatori della sicurezza stradale efficaci strumenti conoscitivi per la valutazione e selezione di politiche e misure volte alla prevenzione degli incidenti. La creazione di nuove applicazioni software, in grado di semplificare l'utilizzo di modelli di previsione degli incidenti complessi, accurati e derivanti da un'ampia esplorazione di potenziali variabili esplicative e forme matematiche, rappresenterebbe una significativa evoluzione dell'attuale stato dell'arte. Questo è proprio il fine della ricerca descritta in questa memoria, tuttora in corso di svolgimento e finanziata, nella sua prima fase, dal Dipartimento Ministeriale per i Trasporti del Regno Unito (DfT10 project, UK Department for Transport). Ad oggi, l'attività di ricerca ha prodotto un tool-kit web-based per la simulazione di scenari d'intervento in materia di sicurezza stradale nel Regno Unito. In questo articolo, s'illustra, in particolare, uno dei modelli econometrici del suddetto tool-kit, ovvero quello che stima la frequenza degli incidenti, a livello di micro-area, in funzione di variabili relative alle infrastrutture di trasporto, al profilo socio-demografico della popolazione locale e all'uso del suolo.

**Parole Chiave:** sicurezza stradale, supporto ai processi di pianificazione, modelli econometrici di previsione.

---

\* Autore a cui spedire la corrispondenza: Mario Catalano (mariocat73@gmail.com)

## 1 Introduzione e scopo della ricerca

Considerata la notevole crescita del trasporto motorizzato e dei relativi costi per la collettività negli ultimi decenni, l'analisi e la prevenzione degli incidenti stradali è una priorità per i decisori politici, le autorità locali e gli esperti di scienze sociali, che svolgono attività di ricerca sui temi dello sviluppo urbano e della mobilità. Infatti, nel 2010, l'Assemblea Generale dell'Organizzazione delle Nazioni Unite ha proclamato l'avvio di una decade di politiche per la sicurezza stradale (*Decade of Action for Road Safety*, Resolution 64/255), con l'obiettivo ambizioso di stabilizzare inizialmente e ridurre, in seguito, i livelli di mortalità stradale in tutto il mondo ([http://www.who.int/violence\\_injury\\_prevention/publications/road\\_traffic/UN\\_GA\\_resolution-54-255-en.pdf](http://www.who.int/violence_injury_prevention/publications/road_traffic/UN_GA_resolution-54-255-en.pdf)).

Lo sviluppo di modelli per la previsione degli incidenti stradali può offrire a politici ed operatori della sicurezza stradale efficaci strumenti conoscitivi per la valutazione e selezione di politiche e misure volte alla prevenzione degli incidenti. Uno studio recente, che ha coinvolto alcune decine di organizzazioni impegnate sul fronte della sicurezza stradale in Europa, in Australia e negli Stati Uniti d'America (autorità nazionali, società di gestione, istituti di ricerca, società di consulenza), rivela che l'impiego di modelli a supporto dei processi di pianificazione nel campo della sicurezza stradale è poco diffuso: circa il 60% delle organizzazioni intervistate si avvale di specifiche procedure per l'analisi delle alternative d'intervento, ma di rado o persino mai impiega modelli di previsione nel corso dei processi di valutazione (Yannis et al., 2016).

Tale mancanza di una cultura della pianificazione che si fondi sulle evidenze scientifiche è probabilmente ascrivibile anche alla scarsa disponibilità, accuratezza e facilità d'uso di strumenti di supporto ad hoc. Infatti, attualmente esistono alcune applicazioni web rivolte al settore, che, tuttavia, possono offrire descrizioni d'interventi già sperimentati, semplici dati statistici sugli impatti rilevati in certi contesti territoriali o, al più, coefficienti per modificare le previsioni di modelli di stima dei tassi d'incidentalità, sviluppati per configurazioni standard dei siti d'interesse (Yannis et al., 2016). Pertanto, la creazione di nuove applicazioni software in grado di semplificare l'utilizzo di modelli di previsione degli incidenti complessi, accurati e derivanti da un'ampia esplorazione di potenziali variabili esplicative e forme matematiche, rappresenterebbe una significativa evoluzione dell'attuale stato dell'arte.

Questo è proprio il fine della ricerca descritta in questa memoria, tuttora in corso di svolgimento e finanziata, nella sua prima fase, dal Dipartimento Governativo per i Trasporti del Regno Unito (DfT10 project, UK Department for Transport), con lo scopo di studiare i fattori esplicativi e gli effetti (anche in termini di congestione ed emissioni inquinanti) degli incidenti stradali nel Regno Unito, nonché sviluppare metodi e modelli per contribuire alla definizione di politiche e misure a favore della sicurezza stradale. Ad oggi, l'attività di ricerca ha prodotto un tool-kit web-based (MAIA, Models and methods for collision prediction and Impact Assessment) per la simulazione di scenari d'intervento in materia di sicurezza stradale nel Regno Unito, a diverse scale territoriali (per una descrizione del tool-kit, cfr. Galatioto et al., 2018). In particolare, l'applicazione MAIA incorpora modelli microeconomici per la previsione delle principali dimensioni del fenomeno: frequenza degli incidenti, numero di feriti per incidente e gravità delle conseguenze per la salute umana. Tali modelli sono il frutto di un'ampia sperimentazione di metodi econometrici e di machine learning. Inoltre, la piattaforma web si caratterizza per una significativa facilità d'utilizzo, richiedendo al fruitore soltanto la selezione del modello d'interesse e l'inserimento dei valori delle variabili esplicative, sì da definire lo scenario di simulazione. È in fase di sviluppo la funzione di stima dei modelli disponibili, in modo da consentire all'utente l'aggiornamento dei loro coefficienti nel tempo e l'elaborazione di nuovi modelli relativi a specifici contesti locali.

In questo articolo, s'illustra, in particolare, uno dei suddetti modelli, ovvero quello che stima la frequenza degli incidenti, a livello di micro-area, in funzione di variabili relative alle infrastrutture di trasporto, al profilo socio-demografico della popolazione locale e all'uso del suolo. Come si vedrà nel prosieguo, tale analisi ha consentito di ottenere strumenti ed



evidenze empiriche utili per definire, a livello politico, obiettivi strategici di promozione della sicurezza stradale.

## 2 Il modello di stima della frequenza d'incidenti stradali a livello di micro-area

Questo modello è stato sviluppato per l'analisi e la previsione dei tassi d'incidentalità stradale nell'ambito di un'area geografica del tipo LSOA (Lower Layer Super Output Area), in relazione ad un intervallo temporale annuale. L'acronimo LSOA si riferisce ad un livello della gerarchia territoriale ideata per la raccolta di dati statistici in Inghilterra e nel Galles. Un'area del tipo LSOA presenta in media una popolazione residente di 1.500 unità. In particolare, per la stima dei coefficienti del modello sono stati utilizzati i dati sugli incidenti stradali con almeno un ferito rilevati dagli organi di polizia secondo il protocollo denominato STATS19; pertanto, il fenomeno degli incidenti da cui scaturiscono solo danni per i veicoli non è stato ancora oggetto di studio, a causa della difficoltà a reperire dati affidabili. Infine, il modello è stato validato valutando il grado di accuratezza delle sue previsioni rispetto ad osservazioni non incluse nel campione impiegato per la stima.

### 2.1 Specificazione del modello

La specificazione del modello per la stima della frequenza attesa d'incidenti stradali in un'area del tipo LSOA comprende variabili relative alle infrastrutture di trasporto, al profilo socio-demografico della popolazione locale e all'uso del suolo. La selezione di queste variabili si basa sull'analisi dei rapporti di correlazione lineare nell'ambito di un più ampio insieme iniziale di potenziali fattori esplicativi, nonché sull'esame della significatività statistica dei relativi coefficienti stimati, che hanno consentito l'esclusione di variabili ridondanti o irrilevanti. La specificazione finale consta degli elementi di seguito illustrati:

- *Estensione della rete stradale nell'area (km)*. Questo regressore è stato utilizzato per rappresentare l'esposizione al rischio d'incidente stradale in luogo dei flussi di traffico, i quali non sono disponibili per tutte le aree LSOA considerate. In particolare, come proxy del traffico veicolare, è stata scelta la somma pesata delle estensioni delle varie tipologie di strada presenti nell'area, i cui pesi sono stati quantificati seguendo le raccomandazioni del Dipartimento Governativo per i Trasporti (Department for Transport, 2017), in modo da riflettere le proporzioni tra le diverse capacità stradali. L'idea di avvalersi di questa proxy s'ispira al lavoro di ricerca di Rose et al. (2009), che dimostra come tale variabile sia un efficace determinante dei livelli d'inquinamento generato dal traffico, non distante dai livelli di accuratezza raggiunti dalle misure dirette dei flussi veicolari.
- *Fattori d'uso del suolo*. È diffusamente dimostrato che il numero delle potenziali destinazioni all'interno di una certa area cresca al crescere della densità di residenti e lavoratori, così riducendo la distanza media percorsa. In generale, tutto ciò aumenta la congestione locale e la frequenza delle collisioni tra veicoli (cfr. Litman, 2007, per un esame della letteratura scientifica di riferimento). Per tale ragione, la densità della popolazione residente per LSOA (n. residenti per km<sup>2</sup>) è stata inclusa nella specificazione del modello. Inoltre, anche il numero di scuole per LSOA è stato inserito, alla luce di alcuni dati statistici ufficiali (Department for Transport, 2018a) che mostrano come nel 2017, in Gran Bretagna, la più elevata percentuale di decessi giovanili (23%) si concentri nelle ore in cui gli studenti lasciano le loro sedi scolastiche (3:00-5:00 p.m.). Infine, anche il numero di pub per LSOA è stato integrato, per tener conto del fenomeno della guida in stato di ebbrezza.
- *Livelli di deprivazione nell'area*. In particolare, due indicatori del set degli indici di deprivazione sono stati selezionati (Department for Communities and Local Government, 2015):

1. *Crime*, che misura il rischio di danni materiali e personali dovuti a violenze e furti e può essere considerato un indicatore dell'attitudine generale verso il rispetto delle regole (comprese quelle stradali) a livello locale;
  2. *Not staying on in education post 16*, che misura la percentuale di giovani che hanno deciso di non proseguire gli studi scolastici dopo i 16 anni d'età. Tale indice, ai fini dello studio in esame, può essere considerato un indicatore del ruolo della formazione scolastica nella vita dei più giovani, proprio quando questi, in generale, conseguono la licenza di guida di un'automobile.
- *Struttura demografica dell'area*. Più in dettaglio, la percentuale delle donne d'età compresa tra 16 e 64 anni e quella dei giovani dai 16 ai 29 anni sono state incluse nella specificazione del modello. La scelta della seconda variabile, in particolare, scaturisce dalla considerazione che i guidatori più giovani sono più esposti al rischio d'incidente stradale, a causa della loro mancanza d'esperienza e di un comportamento di guida solitamente più impetuoso e meno prudente. Infatti, le statistiche ufficiali mostrano come nel 2017, in Gran Bretagna, i giovani dai 16 ai 29 anni rappresentino circa il 14% dei titolari di licenza di guida di un'auto (Department for Transport, 2018b), ma costituiscano il 34% dei guidatori deceduti o gravemente feriti negli incidenti automobilistici (Department for Transport, 2018a).

## 2.2 Struttura e stima del modello

Per la stima del modello in esame, sono stati utilizzati dati relativi all'anno 2015 e al solo contesto Inglese, in quanto gli indicatori di deprivazione sono determinati in modo differente nei Paesi che costituiscono il Regno Unito e l'Inghilterra è l'area più popolata.

La variabile d'interesse è il numero annuale d'incidenti stradali per LSOA, con almeno un ferito e rilevati secondo il protocollo STATS19 dagli organi della Polizia. Questa ha la natura di variabile di conteggio, che richiede l'impiego di metodi parametrici non lineari per l'analisi delle sue relazioni con i fattori esplicativi prescelti. Ci si riferisce a quella classe di modelli che presentano come variabile dipendente una variabile casuale discreta (Cameron and Trivedi, 2005), che può assumere una distribuzione del tipo Poisson o Binomiale Negativa (es.: probabilità di osservare 1, 2, ..., incidenti nell'anno per LSOA).

Dalla figura 1, emerge che il numero d'incidenti stradali per LSOA in Inghilterra nel 2015 non può essere rappresentato con una distribuzione di Poisson, in quanto un notevole grado di sovradisersione è presente nei dati (la varianza è maggiore della media); pertanto, il modello più appropriato per la grandezza in esame è quello, più generale, che assume la forma della distribuzione binomiale negativa (Cameron and Trivedi, 2005). Tale modello ipotizza che la variabile dipendente, ossia la probabilità che si verifichi un certo numero di eventi relativi al fenomeno considerato, abbia una distribuzione di Poisson, ma condizionata rispetto ad una componente casuale:

$$Pr(y|\lambda) = \exp(-\lambda) \cdot \lambda^y / y! \quad (1)$$

Essendo il parametro  $\lambda$  una variabile casuale:

$$\lambda = \mu \cdot v \quad (2)$$

Dove  $\mu$  è funzione di un vettore di regressori  $x$ , generalmente  $\mu = \exp(x' \cdot \beta)$ , e  $v (> 0)$  è una variabile casuale con densità di probabilità  $g(v|\alpha)$ . Tutto ciò implica una forma latente di eterogeneità, in quanto diverse osservazioni possono presentare differenti valori del parametro  $\lambda$ , ma parte di questa variabilità è dovuta ad una componente casuale ( $v$ ).

Da quanto precede, deriva che la probabilità che la variabile dipendente assuma un certo valore  $y$ , incondizionata rispetto alla componente casuale  $v$ , ma condizionata in relazione ai parametri  $\alpha$  e  $\mu$ , può essere così formulata:

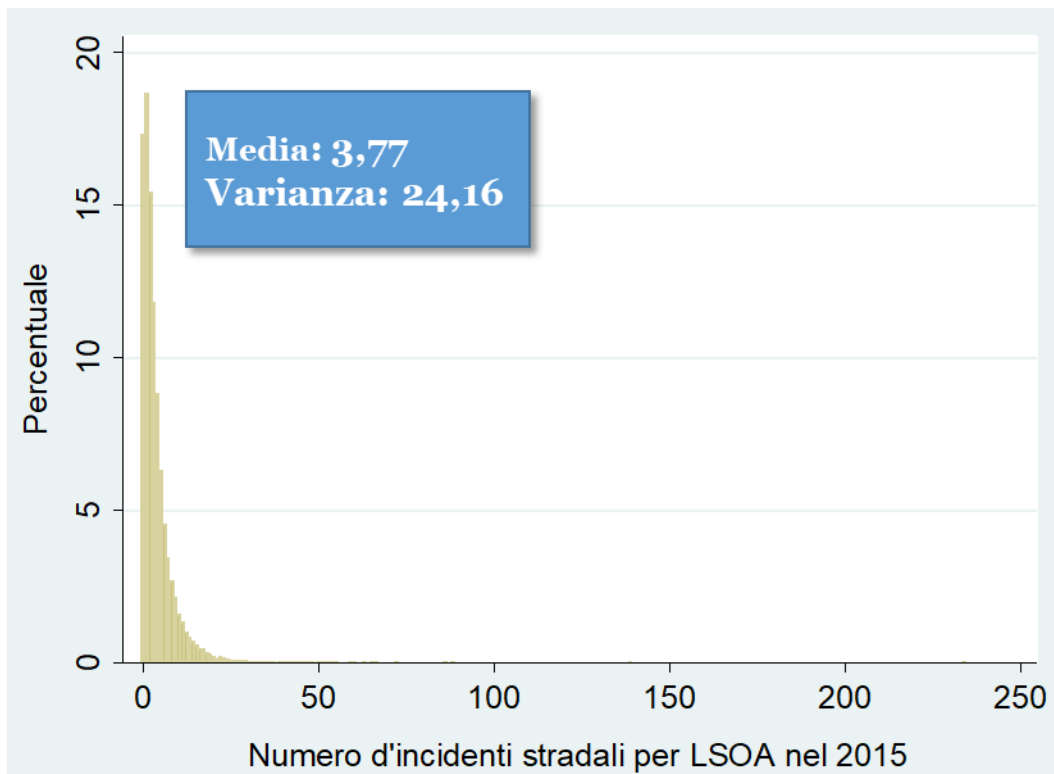
$$Pr(y|\mu, \alpha) = \int Pr(y|\mu, v) \cdot g(v|\alpha) \cdot dv \quad (3)$$

Si tratta, dunque, di un modello misto in cui la probabilità incondizionata di un certo numero di eventi si determina tramite un integrale lungo la densità della componente casuale che genera l'eterogeneità latente. Nel caso in cui la densità  $g(v|\alpha)$  sia del tipo gamma con media 1 e varianza  $\alpha$ , il suddetto integrale può essere risolto analiticamente.

Infine, la media (numero atteso di eventi) e la varianza della variabile dipendente possono essere così formulate:

$$\text{Media} = E(y|\mu, \alpha) = \mu \tag{4}$$

$$\text{Varianza} = V(y|\mu, \alpha) = \mu \cdot (1 + \alpha \cdot \mu)$$



**Figura 1. Distribuzione del numero d'incidenti in Inghilterra per LSOA nel 2015 (dati STATS19)**

Il modello in esame è stato stimato con il metodo della massima verosimiglianza, tenendo conto delle correlazioni tra le LSOA appartenenti allo stesso ente locale (district) nella determinazione della matrice di varianza-covarianza dei coefficienti incogniti. Il coefficiente stimato di una certa variabile esplicativa può essere interpretato come una semielasticità, ossia la variazione relativa del numero atteso d'incidenti all'anno indotta da una variazione unitaria della variabile considerata. Se, invece, il regressore in esame è misurato in scala logaritmica o in termini relativi, il suo coefficiente stimato è un'elasticità. Così, ad esempio, come si può evincere dalla tabella 1 che presenta i risultati della stima, l'aumento unitario del numero di scuole secondarie nell'area è associato ad una crescita di circa il 20% del numero atteso annuale d'incidenti stradali; nel caso della percentuale di giovani dai 16 ai 29 anni, poi, un aumento di 10 unità percentuali è associato ad una crescita di circa il 23% del numero atteso d'incidenti all'anno.

**Tabella 1. Stima del modello per la frequenza d'incidenti stradali a livello di micro-area.**

Variabili esplicative	Coefficienti	Errori Standard	z	P> z	[Intervallo Confidenza]	
Log (estensione della rete stradale in km)	0,5200	0,0183	28,36	0,0000	0,4840	0,5559
Densità della popolazione (10 <sup>3</sup> res./km <sup>2</sup> )	0,0592	0,0037	15,92	0,0000	0,0519	0,0665
Numero di pub	0,0648	0,0041	15,89	0,0000	0,0568	0,0727
Numero di scuole secondarie	0,1991	0,0113	17,64	0,0000	0,1770	0,2212
<i>Crime*</i>	0,0611	0,0138	4,44	0,0000	0,0341	0,0880
<i>Not staying on in education post 16 (%)</i>	0,3438	0,1384	2,48	0,0130	0,0726	0,6151
Percentuale di donne 16-64 anni (%)	-2,9795	0,4123	-7,23	0,0000	-3,7877	-2,1713
Percentuale di giovani 16-29 anni (%)	2,3528	0,3181	7,40	0,0000	1,7294	2,9762
Costante	1,3037	0,2313	5,64	0,0000	0,8503	1,7571
Alpha (varianza della componente casuale)	0,5816	0,0168			0,5495	0,6154

Numero di osservazioni = 32844

Log pseudoverosimiglianza = -75577,522

Wald  $\chi^2(8) = 2129,85$

Prob >  $\chi^2 = 0,0000$

Pseudo  $R^2 = 0,0652$

\* Punteggio frutto di una standardizzazione di indicatori elementari (relativi a diverse fattispecie di crimine) e di analisi fattoriale per la determinazione dei pesi con i quali integrare gli indicatori elementari in un unico indice dal significato più generale (Department for Communities and Local Government, 2015).

Dall'esame dei risultati presentati nella tabella 1, si possono trarre le seguenti considerazioni sul ruolo dei fattori esplicativi prescelti:

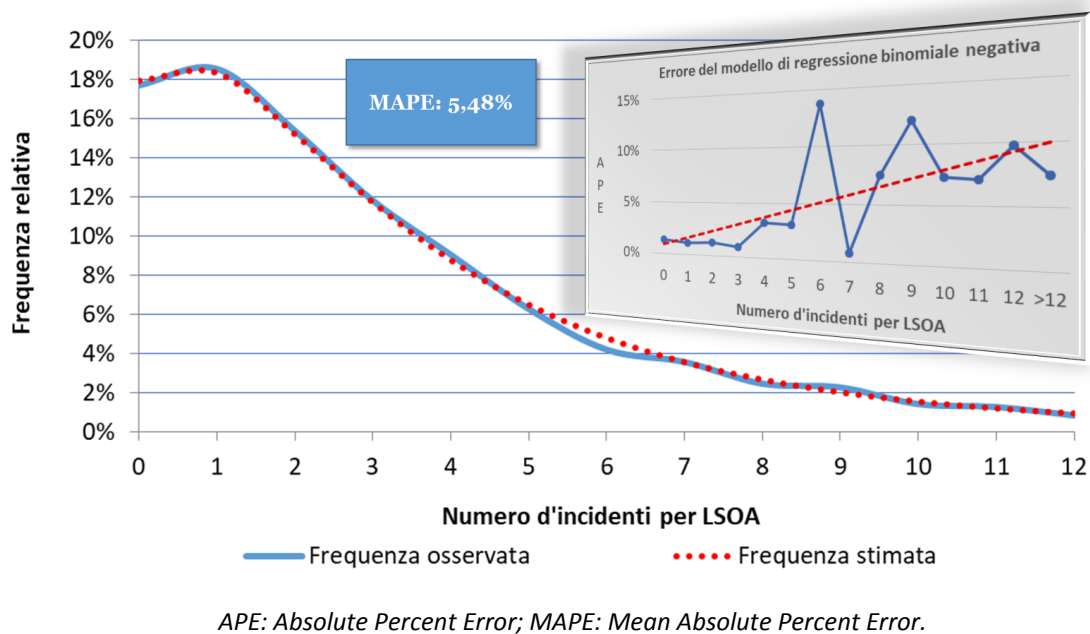
- il segno positivo del coefficiente della variabile di esposizione (estensione della rete stradale in km) conferma che più elevati livelli di mobilità su strada nell'area (misurati indirettamente, in questo caso) sono associati ad una crescita del numero atteso d'incidenti all'anno.
- Una più elevata densità della popolazione residente comporta, verosimilmente a causa di fenomeni di congestione, un più alto rischio di collisione tra veicoli e danni alle persone.
- Il numero delle scuole secondarie nella LSOA è una variabile che ha una significativa relazione diretta con il tasso d'incidentalità stradale, mentre non significativo è risultato il numero di scuole primarie e dell'infanzia. Inoltre, anche il numero di pub nell'area ha un significativo impatto negativo sui livelli di sicurezza stradale. Tutto ciò suggerisce che i poli attrattori e generatori di spostamenti compiuti dai più giovani e quelli in cui si consumano alcolici meritano una grande attenzione come potenziali fonti di comportamenti rischiosi su strada.
- Il coefficiente stimato della variabile *crime* dà un'interessante conferma: più elevati tassi di criminalità, essendo questi un indicatore dell'attitudine generale verso il rispetto delle regole, corrispondono a più diffusi comportamenti di guida illegali e pericolosi.
- Anche il coefficiente della variabile *Not staying on in education post 16* ha un segno positivo. Questo interessante risultato sembra rivelare un ruolo favorevole della formazione per la sicurezza stradale e merita certamente un approfondimento in futuro.
- In relazione alla proporzione di donne residenti dai 16 ai 64 anni, si può dire che, laddove è probabile che la proporzione di guidatori di sesso femminile sia più alta, la frequenza degli incidenti stradali tende ad essere minore.

- Infine, laddove la percentuale di giovani residenti dai 16 ai 29 anni è maggiore, cresce la frequenza attesa degli incidenti, il che avvalorava l'ipotesi di comportamenti di guida più rischiosi nel caso dei più giovani, in linea con altri studi in materia (in particolare, cfr. Butcher, 2016).

### 2.3 Validazione del modello

Il modello di regressione binomiale negativa è stato validato valutando la sua performance in relazione ad osservazioni non incluse nel campione di stima. Più in dettaglio, il set di dati sul numero d'incidenti stradali con feriti per LSOA e sulle variabili esplicative descritte in precedenza, relativo al contesto inglese nel 2015, è stato suddiviso in modo casuale in due campioni: uno per la stima (70% delle osservazioni) ed uno per la validazione (il restante 30%, ossia circa 10.000 osservazioni).

La figura 2 mostra, per ogni valore del numero d'incidenti per LSOA nell'intervallo 0-12 (95% delle osservazioni), il confronto tra le frequenze relative osservate e quelle stimate con il modello di regressione binomiale negativa<sup>1</sup> (Cameron and Trivedi, 2005).



**Figura 2. Distribuzione del numero d'incidenti per LSOA in Inghilterra nel campione di validazione: frequenze osservate versus frequenze stimate (2015).**

La figura 2 dimostra che il modello di regressione binomiale negativa è in grado di riprodurre con accuratezza, nel campione di validazione, le frequenze osservate per i diversi valori del numero d'incidenti, commettendo un errore percentuale in valore assoluto (APE) che fluttua attorno ad un trend crescente e presenta una media (MAPE) molto contenuta, ossia 5,48%. È importante, altresì, sottolineare che, come il suddetto trend

<sup>1</sup> Ciascuna calcolata come media, nel campione di validazione, della probabilità stimata con il modello per il relativo numero d'incidenti.

indica, l'errore di previsione del modello tende ad aumentare per i casi più problematici e meno frequenti (numero d'incidenti per LSOA più elevato).

### 3 Conclusioni

Questo articolo ha presentato parte dei risultati di un'attività di ricerca, in corso di svolgimento, sul tema del supporto alla pianificazione in materia di sicurezza stradale tramite modelli di previsione degli incidenti stradali. La prima fase della ricerca è stata finanziata dal Dipartimento Governativo per i Trasporti del Regno Unito (DfT10 project, UK *Department for Transport*), con lo scopo di studiare i fattori esplicativi e gli effetti (anche in termini di congestione ed emissioni inquinanti) degli incidenti stradali nel Regno Unito, nonché sviluppare metodi e modelli a supporto della formulazione di politiche e misure a favore della sicurezza stradale. La prima fase dello studio ha prodotto una piattaforma web-based per la simulazione di scenari d'intervento attraverso modelli microeconomici per la stima della frequenza, del numero di feriti e della gravità degli incidenti stradali.

Questa memoria ha illustrato un modello relativo al primo dei tre aspetti, mettendo in luce come il tasso d'incidentalità a livello di micro-area (LSOA) dipenda da variabili relative alle infrastrutture di trasporto, al profilo socio-demografico della popolazione locale e all'uso del suolo. Inoltre, il modello descritto ha mostrato un significativo peggioramento della performance di previsione per le manifestazioni più gravi e meno frequenti del fenomeno oggetto di studio (tassi elevati d'incidentalità). Per tale ragione, in futuro, s'intende esplorare logiche di modellazione del problema di tipo diverso, ovvero quella nota nella letteratura scientifica di riferimento come *ensemble modelling* o *learning* (Alikhani et al., 2013; Catalano and Galatioto, 2017; Sohn and Lee, 2003). In particolare, si vorrebbe sperimentare un approccio che preveda la suddivisione del campione di osservazioni in cluster di casi simili, la stima di modelli di previsione specifici per ciascun cluster e lo sviluppo di un meta-modello in grado di selezionare, per una certa combinazione dei valori delle variabili esplicative, il modello di previsione relativo al cluster di minore distanza (ad esempio, in termini di metrica euclidea) dallo scenario dato.

#### *Riferimenti bibliografici*

Alikhani, M., Nedaie, A. and Ahmadvand, A. (2013). Presentation of clustering-classification heuristic method for improvement accuracy in classification of severity of road accidents in Iran. *Safety Science* 60, 142-150.

Butcher, L. (2016). *Young and novice drivers*. House of Commons Library Briefing Paper, Number SN517.

Cameron, A.C. and Trivedi, P.K. (2005). *Microeconometrics: Methods and Applications*. Cambridge University Press, New York.

Catalano, M. and Galatioto, F. (2017). Enhanced transport-related air pollution prediction through a novel metamodel approach. *Transportation Research Part D* 55, 262-276.

Department for Transport (2017). TAG UNIT A5.4 - Marginal External Costs (this TAG Unit is guidance for the appraisal practitioner and is part of the family "A5 - UNI-MODAL APPRAISAL"). Disponibile all'indirizzo web <https://www.gov.uk/transport-analysis-guidance-webtag>.

Department for Transport (2018a). Reported Road Casualties Great Britain: 2017 Annual Report. Disponibile all'indirizzo web [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/755698/rrcgb-2017.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/755698/rrcgb-2017.pdf).

Department for Transport (2018b). GB Driving Licence Data. Disponibile all'indirizzo web <https://data.gov.uk/dataset/d0be1ed2-9907-4ec4-b552-c048f6aec16a/gb-driving-licence-data>.

Department for Communities and Local Government (2015). The English Indices of Deprivation 2015 - Technical Report. Disponibile all'indirizzo web



[https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/464485/English\\_Indices\\_of\\_Deprivation\\_2015\\_-\\_Technical-Report.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/464485/English_Indices_of_Deprivation_2015_-_Technical-Report.pdf).

Galatioto, F., Catalano, M., Shaikh, N., McCormick, E. and Johnston, R. (2018). Advanced accident prediction models and impacts assessment. *IET Intelligent Transport Systems* 12 (9), 1131-1141.

Rose, N., Cowie, C., Gillett, R. and Marks, G. B. (2009). Weighted road density: A simple way of assigning traffic-related air pollution exposure. *Atmospheric Environment* 43, 5009-5014.

Sohn, S.Y. and Lee, H.S. (2003). Data fusion ensemble and clustering to improve the classification accuracy for the severity of road traffic accidents in Korea. *Safety Science* 44, 1-14.

Yannis, G., Dragomanovits, A., Laiou, A., Richter, T., Ruhl, S., La Torre, F., Domenichini, L., Graham, D., Karathodorou, N. and Li, H. (2016). Use of accident prediction models in road safety management - an international inquiry. *Transportation Research Procedia* 14, 4257-4266.