

## **Regresní analýza nehodovosti na světelně řízených křižovatkách**

RNDr. Martin Hála, CSc.

*Abstract: Various statistical methods and tools were used and tested when analyzing traffic accidents. The main stress was put on the use of regression and correlation tools. First we derived regression models for the total count of accidents on the given intersection. In the second part of the research we tried to apply regression methods when looking for the relationship between the number of accidents on a given intersection related to specified maneuvers of the cars involved and the traffic volumes related to the approaches to the intersection used by these cars. We derived very interesting and helpful results.*

### **1. Úvod**

Významnou součástí výzkumného záměru „Udržitelná výstavba“ je tvorba a modifikace matematických a počítačových modelů rizikové analýzy procesů a systémů. Cílem tohoto příspěvku je informovat odbornou veřejnost o různých možnostech, ale i úskalích použití statistických metod, zejména pak regresní analýzy, při zkoumání nehodovosti. Uplatnění získaných poznatků je především v oblasti prognózování – při predikci dopravních nehod.

Jako výchozí databáze byly vybrány záznamy o nehodách na pražských světelně řízených křižovatkách za posledních 10 let. Tyto údaje spolu s dalšími daty (které se vztahovaly ke stavebnímu uspořádání jednotlivých křižovatek, uspořádání radících pruhů a především k dopravnímu zatížení jednotlivých vjezdů) byly analyzovány různými statistickými metodami. Cílem bylo jednak získat konkrétní užitečné výstupy, týkající se nehodovosti pro danou databázi (např. objektivní ohodnocení „nebezpečnosti“ jednotlivých uzlů), kromě toho jsem se též snažil posoudit vhodnost jednotlivých použitých statistických metod, tj. zjistit, do jaké míry se výstupy, získané jednotlivými metodami liší a proč tomu tak je, jaké nároky na předzpracování dat jednotlivé metody kladou, apod.

### **2. Analyzovaná data**

Pracoval jsem s údaji o nehodách na pražských čtyřramenných světelně řízených křižovatkách za časové období od roku 1995 do roku 2004 včetně. Tyto záznamy byly získány od Policie ČR v podobě datových souborů pro všechny pražské sledované uzly zvláště pro jednotlivé roky a jednotlivé obvody. V těchto

---

Martin Hála, RNDr., CSc.

ČVUT v Praze, Fakulta stavební, Katedra matematiky

Thákurova 7, 166 29 Praha 6 - Dejvice

tel.: 224355449, e-mail: hala@mat.fsv.cvut.cz

souborech v zásadě jedné nehodě odpovídá jeden řádek s podrobnými údaji mimo jiné o místě a čase nehody, pohybech jednotlivých zúčastněných vozidel, hmotných škodách, případném zranění osob, o funkčnosti světelné signalizace v době nehody, atd. Kromě těchto údajů jsem čerpal z databáze pražských křižovatek zpracované Ústavem dopravního inženýrství (ÚDI) v rámci programu IDIS, kde jsou uloženy v grafické podobě podrobné údaje o stavebním uspořádání křižovatek, počtu a uspořádání řadících pruhů a hlavně údaje o zatížení jednotlivých vjezdů podle druhů vozidel a směrů jízdy v podobě 16-ti hodinových intenzit získaných při posledním sčítání dopravy na daném uzlu.

Hlavním úkolem, který jsem si vytýčil, bylo hledání vztahů mezi nehodovostí a zatížením daného uzlu, ověřování různých pracovních hypotéz o této závislosti, popř. i zkoumání vlivu časového faktoru. Za tímto účelem bylo třeba programově analyzovat data z obou databází. Určitým problémem však byla absence digitální podoby údajů uložených v IDISu. Proto byla v průběhu řešení vytvořena nová databáze sledovaných uzlů, do které byly v podstatě „ručně“ přepisovány údaje z IDISu, tj. čísla jednotlivých vjezdů, počty a druhy řadících pruhů a příslušné intenzity dopravy. Je třeba otevřeně říci, že spolehlivost těchto dat rozhodně není stoprocentní. IDIS je sice rok od roku aktualizován, přesto však velmi pomalu a nespolehlivě reaguje na případné změny stavebního uspořádání uzlů. Největší slabinou jsou jistě intenzity. IDIS v podstatě vychází vždy z posledního sčítání dopravy na daném uzlu a toto sčítání bylo leckdy realizováno v době dávno před zkoumaným časovým obdobím (není výjimkou začátek 90. nebo i konec 80. let).

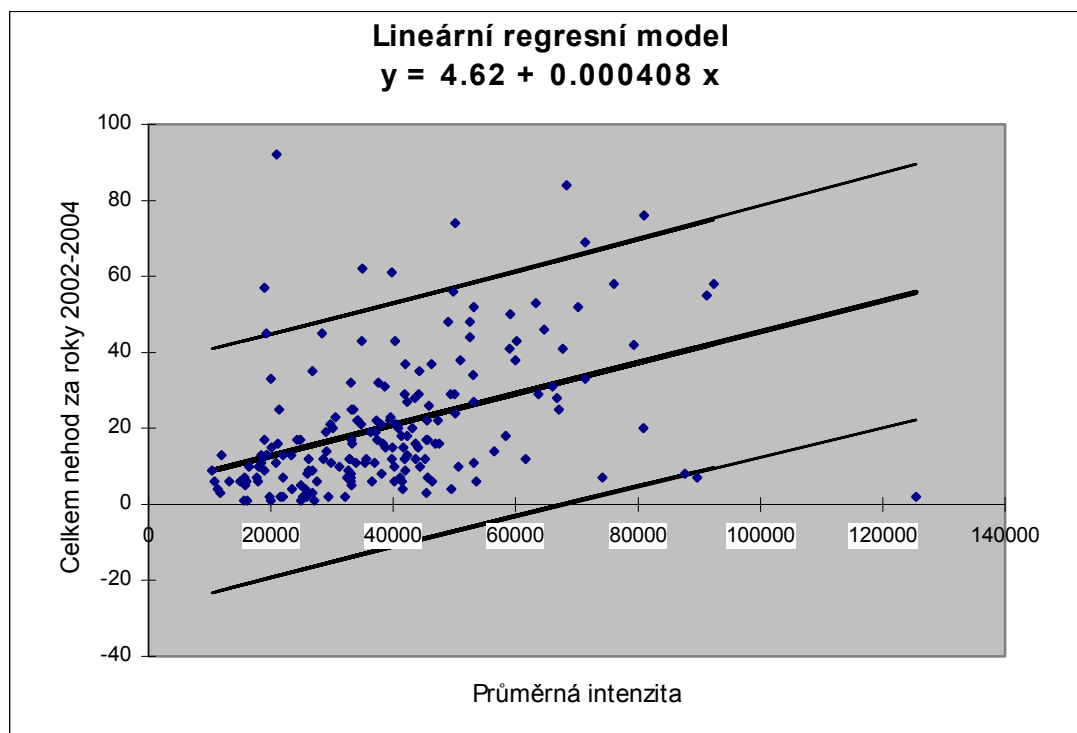
Co se týče řádkových výpisů nehod získaných od Policie ČR, bylo třeba především automaticky (pomocí vhodného počítačového programu) vytržít jen nehody na zkoumaných uzlech, tj. čtyřramenných světelně řízených křižovatkách. Při třídění jsem uvažoval jen nehody, které se odehrály za provozu světelné signalizace, kvůli srovnatelnosti údajů a dále kvůli předpokladům používaných metod jsem se omezil jen na srážky dvou nekolejových vozidel (tj. nezpracovával jsem hromadné nehody, nehody s účastí chodců, s pevnou překážkou, s tramvají, atd.). Konečně jsem se též omezil jen na čtyřramenné křižovatky, o kterých jsem měl z IDISu k dispozici údaje o intenzitách.

Při výše vysvětleném předzpracování dat se ukázala určitá nekompatibilita policejní databáze a databáze IDIS. Např. lišící se číslování vjezdů do křižovatek, některé údaje o směrech jízdy zúčastněných vozidel byly ve zjevném rozporu s uspořádáním křižovatky, apod. To, co bylo v mých silách, jsem zpětně opravil, nehody se zásadními nejasnostmi jsem ze zpracování buďto vyloučil nebo jsem je zpracovával v rámci oddělené kategorie „nezařazených nehod“.

### 3. Použití regresní analýzy, závislost nehodovosti na celkovém zatížení křižovatky

Objektivní možnosti pro posouzení nebezpečnosti jednotlivých uzlů se současným zahrnutím vlivu dopravního zatížení dává regresní analýza. V první fázi jsem odvozoval lineární i nelineární regresní modely závislosti nehodovosti na jednotlivých uzlech (vyjádřené prostým počtem nehod daného druhu za zvolené období) na průměrném celkovém dopravním zatížení daného uzlu za toto období. Jako druh nehod jsem uvažoval všechny nehody, resp. jen nehody se zraněním a jako rozhodné období byly voleny nejprve jednotlivé roky a potom jsem zkoušel tyto roky slučovat do delších období – za účelem získání věrohodnějších modelů, založených na větším počtu dat. Při tomto slučování bylo třeba najít vhodný kompromis. Modely odvozené jen z dat jednoho roku jsou nespolehlivé, neboť příliš mnoho uzlů vykazuje během jednoho roku jen nepatrné množství nehod. Na druhou stranu model založený na celém desetiletém období pracuje s průměrnými intenzitami a tak nemůže reflektovat případné velké změny v dopravním zatížení (např. trvalý odklon dopravy) a samozřejmě ani změny v uspořádání křižovatky. Jako nejrozumnější volba se zdá být 3 roky. V obr. 1 jako ukázkou uvádím lineární regresní model pro všechny nehody za poslední tři roky sledovaného období.

Obr. 1 Ukázka regresního modelu



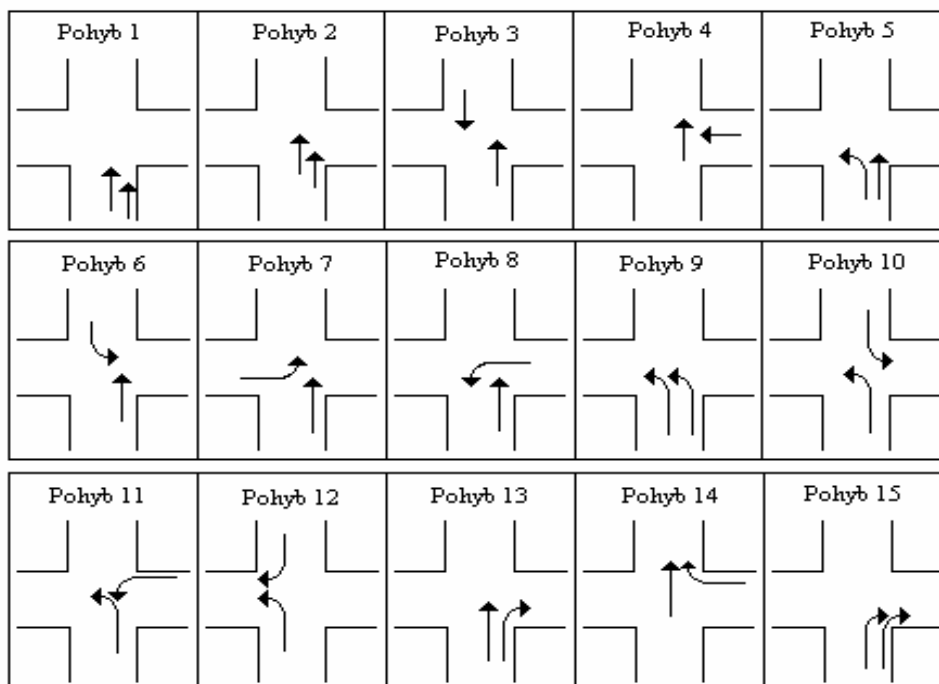
Směrnice přímky je významně různá od nuly, zatímco konstanta by bylo možno vypustit. Uveďme pro zajímavost interpretaci směrnice nalezené přímky: Pro „průměrnou“ křižovatku z naší databáze by trvalému nárůstu celkového dopravního zatížení o např. 10000 vozidel odpovídalo očekávané zvýšení počtu nehod o  $10000 * 0.000408 \cong 4$  nehody (ovšem za tříleté období!).

Křivky nad a pod regresní přímkou vyjadřují hranice 95% predikčního pásu pro závisle proměnnou. Tento pás lze interpretovat mimo jiné tak, že uzly ležící nad ním se jeví jako statisticky významně nebezpečné (na hladině spolehlivosti 95%) a naopak uzly ležící pod tímto pásem se analogicky jeví jako neobvykle bezpečné.

#### 4. Závislost nehodovosti na zatížení vjezdů vozidel zúčastněných na nehodě

V další části výzkumu jsem se zabýval detailní analýzou nehodovosti podle křižovatkových pohybů vozidel zúčastněných na nehodě. Navazoval jsem tím na statistické práce E. Hauera, který se touto problematikou zabýval v 90. letech při analýze nehodovosti v Torontu. Hauer zavedl klasifikaci nehod dvou vozidel na čtyřramenných křižovatkách podle jejich zamýšleného manévru před nehodou do následujících 15 typů – viz obr. 2.

Obr. 2 Dělení nehod podle druhu konfliktního pohybu



Nejčastějším typem nehody ve zkoumané databázi je náraz zezadu (kód 2), z hlediska nebezpečnosti jsou ovšem důležitější kódy 4 a 6 (srážka z boku, resp. srážka dvou protijedoucích vozidel, při které jedno z nich odbočuje vlevo).

Rozhodně velmi důležitou částí analýzy je zkoumání nehodovosti v uzlech databáze podle jednotlivých vjezdů a výše uvedených křižovatkových pohybů. I jenom mechanické roztřídění počtu nehod s určitým kódem podle jednotlivých vjezdů a uzlů za zvolené období přinese užitečné informace o nebezpečnosti křižovatek. Určitě je cennější informace, že na určité křižovatce a určitém vjezdu došlo za dané období k neobvykle vysokému počtu nehod typu 6, než pouhá informace, že tato křižovatka vykazuje velký celkový počet nehod.

Pro srovnatelnost počtu nehod na jednotlivých uzlech bylo třeba zohlednit řazení pruhů na inkriminovaných vjezdech a též zatížení těchto vjezdů. Hypotézou je, že řekněme počet nehod typu 6 podle Hauerovy klasifikace na dané křižovatce a vjezdu by měl záviset na počtu pruhů (přesněji na počtu všech kombinací pruhů), ve kterých se daná nehoda mohla odehrát a dále pak na průměrných intenzitách dopravy v těchto pruzích. Relativní nehodovostí v dalších výpočtech rozumím počet nehod (daného typu za dané časové období na určitém vjezdu) dělený počtem kombinací jízdních pruhů, ve kterých se tyto nehody mohly odehrát. Pomocí dvourozměrné lineární regresní analýzy byly hledány modely závislosti této relativní nehodovosti na průměrných intenzitách připadajících za dané období na jeden pruh dovolující manévr vozidla A (*prumintenzita1*) a na jeden pruh dovolující manévr vozidla B (*prumintenzita2*).

Jako ukázkou uvádím výsledek jednoho z mnoha regresních výpočtů.

VÝSLEDEK	Všechny nehody	Kód 4	Období: 2001-2004	
<i>Regresní statistika</i>				
Násobné R	0.173309			
Hodnota spolehlivosti R	0.030036			
Nastavená hodnota spolehlivosti R	0.022488			
Chyba stř. hodnoty	2.135766			
Pozorování	260			
ANOVA				
	<i>Rozdíl</i>	<i>SS</i>	<i>MS</i>	<i>F</i>
Regrese	2	36.3019	18.1509	3.979167

Rezidua	257	1172.3043	4.5615	
Celkem	259	1208.6062		

	<i>Koeficienty</i>	<i>Chyba stř. hodnoty</i>	<i>t stat</i>	<i>Hodnota P</i>
Konstanta	0.893023	0.276229	3.232907	0.001386
Prumintenzita1	0.000029	0.000017	1.710177	0.088440
Prumintenzita2	0.000042	0.000017	2.437359	0.015474

Kvalita výše uvedeného modelu není sice velká, ale prokázalo se, že relativní nehodovost typu 4 statisticky významně roste s rostoucím dopravním zatížením pruhu, ve kterém se pohybuje vozidlo B (tj. vozidlo, které při nehodě naráží zprava do druhého vozidla – viz Hauerovo schéma na obr. 2).

Vedlejším, avšak z praktického hlediska velmi užitečným výstupem je seznam nejnebezpečnějších uzlů a vjezdů, který lze snadno odvodit analýzou reziduí nejen pro tento, ale pro jakýkoliv regresní model.

Poděkování: Tento příspěvek je podporován výzkumným záměrem Udržitelná výstavba MSM 684070005.

### **Použitá literatura**

- ČARSKÁ, Z., 2003: Optimalizace stavebních a dopravních podmínek na křižovatkách z hlediska nehodovosti, disertační práce, ČVUT-FD
- HÁLA, M., 2003: Statistical analysis of traffic accidents, sborník konf.: Mathematical and computer modeling in science and engineering, Praha, str. 137-141
- HÁLA, M., SLABÝ, P., 1999: Statistická analýza dopravních nehod – 3. část, Silniční obzor **60-7**, Praha, str. 186-190