

ČESKÉ VYSOKÉ UČENÍ TECHNICKÉ V PRAZE
Fakulta dopravní

Bc. Alina Troitskaya

MODELOVÁNÍ DOPRAVNÍCH NEHODOVÝCH DAT

Diplomová práce

2019



K617..... Ústav logistiky a managementu dopravy

ZADÁNÍ DIPLOMOVÉ PRÁCE
(PROJEKTU, UMĚLECKÉHO DÍLA, UMĚLECKÉHO VÝKONU)

Jméno a příjmení studenta (včetně titulů):

Bc. Alina Troitskaya

Kód studijního programu a studijní obor studenta:

N 3710 – LA – Logistika a řízení dopravních procesů

Název tématu (česky): **Modelování dopravních nehodových dat**

Název tématu (anglicky): Modeling of Transport Accident Data

Zásady pro vypracování

Při zpracování diplomové práce se řiďte následujícími pokyny:

- Popis dopravy a nehodovosti, základní pojmy
- Analýza dopravních nehod a jejich druhů
- Popis matematických modelů
- Popis hl. města Praha, vstupní data
- Praktická část, odhad diskretních dat
- Závěr, zhodnocení





- Rozsah grafických prací: podle pokynů vedoucího diplomové práce
- Rozsah průvodní zprávy: minimálně 55 stran textu (včetně obrázků, grafů a tabulek, které jsou součástí průvodní zprávy)
- Seznam odborné literatury: Šachl, J.: Analýza nehod v silničním provozu. ČVUT v Praze, 2010
Zeltermán, D.: Models for Discrete Data. Oxford University Press, 2006
Chemlík, J. a kolektiv: Dopravní nehody. Vydavatelství a nakladatelství Aleš Čeněk s.r.o., 2009

Vedoucí diplomové práce: **doc. Ing. Zdeněk Říha, Ph.D.**
doc. Ing. Ivan Nagy, CSc.

Datum zadání diplomové práce: **30. června 2018**
(datum prvního zadání této práce, které musí být nejpozději 10 měsíců před datem prvního předpokládaného odevzdání této práce vyplývajícího ze standardní doby studia)

Datum odevzdání diplomové práce: **2. prosince 2019**
a) datum prvního předpokládaného odevzdání práce vyplývající ze standardní doby studia a z doporučeného časového plánu studia
b) v případě odkladu odevzdání práce následující datum odevzdání práce vyplývající z doporučeného časového plánu studia


doc. Ing. Tomáš Horák, Ph.D.
vedoucí
Ústavu logistiky a managementu dopravy


doc. Ing. Pavel Hrubeš, Ph.D.
děkan fakulty

Potvrzuji převzetí zadání diplomové práce.


Bc. Alina Troitskaya
jméno a podpis studenta

V Praze dne6. června 2019

Poděkování

Ráda bych na tomto místě poděkovala svým vedoucím, panu doc. Ing. Ivan Nagymu, CSc. a panu doc. Ing. Zdenku Říhovi, Ph.D. za vedení, vstřícnost při konzultacích a vypracování práce, cenné rady a pomoc při získání potřebných informací a podkladů.

Také bych ráda poděkovala své rodině a blízkým přátelům za pomoc a neustálou podporu během studia.

Prohlášení

Prohlašuji, že jsem předloženou práci vypracovala samostatně a že jsem uvedla veškeré použité informační zdroje v souladu s Metodickým pokynem o dodržování etických principů při přípravě vysokoškolských závěrečných prací.

Nemám žádný závažný důvod proti užití školního díla ve smyslu § 60 Zákona č. 121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon).

V Praze, dne

ČESKÉ VYSOKÉ UČENÍ TECHNICKÉ V PRAZE

Fakulta dopravní

MODELOVÁNÍ DOPRAVNÍCH NEHODOVÝCH DAT

diplomová práce

listopad 2019

Alina Troitskaya

ABSTRAKT

Předmětem diplomové práce „Modelování dopravních nehodových dat“ je porovnat a zhodnotit dopravní nehody podle různých kritérií na území Evropské unie a území České republiky. Druhá polovina práce se věnuje modelování nehodových dat s použitím statistických rozdělení. Snahou je najít vhodný model, který by mohl nahradit nejčastěji použitelný model s kategoričným rozdělením.

ABSTRACT

The aim of the diploma thesis „Modeling of traffic accident data“ is to compare and evaluate traffic accidents according to various criteria in the territory of the European Union and the territory of the Czech Republic. The second half of the thesis deals with modeling of accident data using statistical distributions. The main goal is to find a suitable model that could replace the most commonly used model with categorical distribution.

Obsah

Seznam použitých zkratk	6
Úvod	7
1. Popis dopravy a nehodovosti, základní pojmy	9
1.1. Historie dopravní nehodovosti. Vývoj pasivní bezpečnosti	9
1.2. Základní pojmy	10
1.3. Kritéria a druhy dopravní nehodovosti	11
1.4. Hlavní příčiny vzniku nehod	12
2. Analýza dopravních nehod a jejich druhů	14
2.1. Analýza nehodovosti v EU	14
2.1.1. Závislost nehodovosti na HDP	15
2.1.2. Závislost nehodovosti na dopravní infrastruktuře	17
2.2. Nehody v ČR	19
2.2.1. Porovnání nehodovosti mezi kraji ČR	19
2.2.2. Dopravní nehodovost v celé ČR	21
2.2.3. Závislost nehodovosti na dopravních investicích	23
3. Popis nehod v hlavním městě Praha. Vstupní data	26
3.1. Praha z pohledu dopravy	26
3.2. Plán udržitelné mobility Prahy jako řešení zvýšení dopravní bezpečnosti	29
3.3. Vstupní data	31
4. Popis matematických modelů	32
4.1. Binomické rozdělení	32
4.2. Kategorické rozdělení	33
4.3. Geometrické rozdělení	34
4.4. Negativně binomické rozdělení	35
4.5. Poissonovo rozdělení	35
4.6. Rovnoměrné rozdělení	36
4.7. Alternativní rozdělení	37
5. Praktická část, odhad diskrétních dat	38
5.1. Odhad simulovaných diskrétních dat	46
5.2 Odhad reálných diskrétních dat	59
Závěr, zhodnocení	66
Použité zdroje	69
Seznam obrázků	70
Seznam tabulek	72

Seznam použitých zkratk

ČR-Česká republika

EU-Evropská Unie

HDP-Hrubý domácí produkt

PPS-páteř kupní síly

SFDI – Státní fond dopravní infrastruktury

ČSÚ – Český statistický úřad

TSK-Technická správa komunikací

Úvod

V současné době je pro většinu lidí náročné si představit svůj život bez denního využití dopravy. S rozvojem měst a populace dochází k celkovému růstu využití různých druhů dopravy a s tím spojeným růstem dopravní sítí. Na druhou stranu s rozvojem sítě a pravidelným využitím dopravy se objevuje riziko vzniku dopravní nehodovosti.

Tato diplomová práce se zabývá automobilovou dopravou, konkrétně analýzou dopravní nehodovosti a modelováním nehodových dat. Danou práci lze rozdělit na několik základních částí.

První kapitola diplomové práce se bude věnovat stručnému popisu vývoje dopravy a úvodu do problematiky dopravní nehodovosti včetně souvisejících definic a základních pojmů.

V další kapitole bude provedena analýza dopravních nehod včetně jejich druhů. První část této kapitoly se bude zabývat analýzou nehod v zemích EU. Ke zpracování bude vybráno několik států EU s různou ekonomickou situací a bude provedeno porovnání úrovně nehodovosti mezi těmito státy. Dále budou hledána závislosti dopravní nehodovosti ve vybraných státech EU na HDP a na kvalitě dopravní infrastruktury. Druhá část této kapitoly se bude věnovat analýze a hodnocení dopravní nehodovosti na území ČR. Kapitola bude obsahovat porovnání úrovně nehodovosti mezi kraji a analýzu vývoje dopravní nehodovosti v určitém období. Potom bude provedeno porovnání nehodovosti podle dnů v týdnu. Na konci kapitoly bude zkoumána závislost mezi vznikem dopravních nehod na území České republiky a investicemi do dopravní infrastruktury.

V třetí kapitole bude popsán vývoj intenzit dopravy a nehodovosti v Praze a také bude plán udržitelné mobility Prahy a okolí jako projektu, který by měl zlepšit dopravní situaci v Praze a Středočeském kraji a mít vliv na snížení počtu dopravních nehod. Kapitola bude obsahovat i popis vstupních dat, potřebných pro praktickou část, týkající se modelování nehodových dat.

V další kapitole bude reprezentován popis matematických modelů, potřebných pro zpracování praktické části. Tento popis zahrnuje seznam statistických rozdělení včetně teorie a matematických vzorků.

Dále bude následovat praktická část, ve které prostřednictvím počítačového programu Scilab a s použitím statistických rozdělení (kategorického, rovnoměrného, binomického, negativně binomického, geometrického a Poissonova) budou modelována simulovaná a reálná diskrétní data. Cílem práce je pokusit se pro reálná data použít nějaký „kompaktní model“, který by byl

nastaven pomocí malého počtu parametrů (podobně jako je tomu u spojitých modelů). Tedy je potřebné najít vhodný model, který by mohl nahradit pro tyto účely nejvíce používaný model s kategoričným rozdělením a zkoumat, jak se simulovaná a reálná data budou modelovat.

Na konci bude provedeno porovnání a hodnocení výsledků. Cílem práce je udělat analýzu a zhodnocení dopravní nehodovosti v státech Evropské unie a území ČR, dát data o dopravní nehodovosti do souladu s ekonomickými ukazateli a s ukazateli kvality dopravní sítě a vyhledat možnou závislost. Tato diplomová práce je v souladu s řešením problematiky analýzy dopravních nehod. Základem analýzy nehodových dat je model, který této data popisuje. Práce byla zaměřena na hledání modelu vhodného pro odhad diskretních nehodových dat. Cílem daného výzkumu je najít okolnosti (způsob jízdy, stavební úpravy, komunikace atd.), které vedou k vážným dopravním nehodám. Výzkum byl započítán na žádost Magistrátu hlavního města Praha.

1. Popis dopravy a nehodovosti, základní pojmy

1.1. Historie dopravní nehodovosti. Vývoj pasivní bezpečnosti

Doprava je nedílnou součástí lidského života, která se zaměřuje na přemísťování osob a věcí pomocí dopravních prostředků po dopravních cestách.

Z historie víme, že vznik a vývoj dopravy je přímo spojen s rozvojem společenských, ekonomických a politických vztahů. Lze připustit, že doprava je přibližně stejně stará jako lidské pokolení samo, její vznik je přímo spjatý s přirozenou lidskou potřebou směřovat zboží.

[1]

S postupem času se doprava vyvinula od primitivní přepravy nákladů až po vysokorychlostní a pohodlnou přepravu osob a nákladů, se kterou se setkáváme dnes. To bylo způsobeno primárně rozvojem dopravních prostředků, kterými je možno uskutečnit tuto přepravu. Na druhou stranu rozvoj dopravy a rapidní nárůst dopravních prostředků na cestách způsobil i zvýšení dopravní nehodovosti.

K první dopravní nehodě se smrtelným zraněním došlo v roce 1896 v Londýně během prezentace nového automobilu, kdy byla sražena 44letá žena, která navzdory všem zákazům vstoupila na vozovku (viz. obr. 1). [2]



Obrázek 1: Silniční nehoda XIX století [2]

S rozvojem dopravy bylo nutné zavést pravidla, které by měl každý dodržovat a tím zvyšovat bezpečnost provozu. Za jeden z primárních důvodů snížení počtu zraněných nebo usmrcených osob v dopravní nehodovosti lze uvést vznik a vývoj pasivní bezpečnosti a její rozšíření v silniční přepravě.

Prvním prvkem pasivní bezpečnosti v silniční dopravě byl tříbodový pás, který v XX. století zavedla švédská společnost Volvo. Dále díky společnosti Volvo byla zavedena čalouněná přístrojová deska. V roce 1973 se objevily první airbagy, které v případě zájmu mohli kupovat vlastníci automobilů Oldsmobile Tornado. V Evropě se první airbagy objevily v roce 1980 díky automobilové společnosti Mercedes. V roce 1978 společnost Bosch uvedla první protiblokovací brzdový systém ABC, který je dosud běžnou výbavou aut. [13]

1.2. Základní pojmy

Definice dopravy byla již uvedena v předchozí kapitole č.1.1. Dopravu rozdělujeme na několik základních druhů: silniční, železniční, potrubní, vodní a leteckou. Bohužel v každém druhu dopravy občas dochází k dopravní nehodě.

Tato diplomová práce se věnuje nehodovosti v silničním provozu, proto je potřeba uvést pojmy související především s nehodami na pozemních komunikacích.

Pozemní komunikace je dopravní cestou, která je určena k užití silničními a jinými vozidly a chodci, včetně pevných zařízení nutných pro zajištění tohoto využití a jeho bezpečnosti.

Silničním vozidlem definujeme motorové nebo nemotorové vozidlo, které je vyrobené s cílem provozu na pozemních komunikacích a přepravy osob, věcí nebo zvířat.

Provozovatelem silničního vozidla je osoba, která je uvedena v registru silničních vozidel jako vlastník daného vozidla, nesmí být v registru zapsaná jako provozovatel jiná osoba. Provozovatel musí udržovat vozidlo v náležitém stavu podle požadavků pro obsluhu a údržbu, stanovenými výrobcem. [3]

Dopravní nehodou rozumíme událost v provozu, která nastala nebo byla započata na pozemní komunikaci. Výsledkem dopravní nehody zpravidla je zranění nebo usmrcení osoby nebo hmotná škoda majetku v přímé souvislosti s provozem vozidla v pohybu.

Místem nehody se nazývá místo, u kterého došlo ke vzniku dopravní nehody.

Účastníkem nehody je osoba, která se přímým způsobem účastní na nehodě. K účastníkům nehody patří: přepravovaná osoba, řidič, cyklista, chodec, osoba přibraná k zajištění bezpečnosti provozu atd.

V případě vzniku dopravní nehody účastníci nehody musí provést bezpečnostní opatření v místě dopravní nehody, poskytnout první pomoc postiženým a přivolat zdravotnickou záchrannou službu. V případě dopravní nehody je její účastník povinen informovat policii a označit místo nehody, sdílet informace o poškození komunikace, zařízení nebo životního prostředí. Navíc účastník dopravní nehody by měl poskytnout údaje o vozidle, sepsat záznam o dopravní nehodě včetně místa a času její vzniku. [4]

1.3. Kritéria a druhy dopravní nehodovosti

Závažnost poranění je hlavním kritériem ve statistice dopravní nehodovosti. Toto kritérium má 3 úrovně: zranění, těžké zranění a úmrtí. Stupeň zranění při dopravní nehodě zpravidla určuje lékař. Za usmrčenou se počítá osoba, která zemřela při dopravní nehodě přímo na místě, nebo do 30 dnů od data nehody.

K dalším kritériím v dopravní nehodovosti patří:

- počasí,
- druh vozidla,
- druh pozemní komunikace,
- lokalita nehody
- stav řidiče
- kategorie řidiče
- viditelnost atd.

Dopravní nehody lze rozdělit na několik druhů:

- srážka s jedoucím nekolejovým vozidlem
- srážka s vozidlem zaparkovaný, odstaveným
- srážka s pevnou překážkou

- srážka s chodcem
- srážka s lesní zvěří
- srážka s domácím zvířetem
- srážka s tramvají
- havárie
- převrácení vozidel v důsledku smyku, ztráty kontroly, nepříznivých silničních podmínek,
- převrácení vozidel v důsledku použití náhlé nebo nesprávné techniky řízení řidičem
- jiný druh nehody [5]

1.4. Hlavní příčiny vzniku nehod

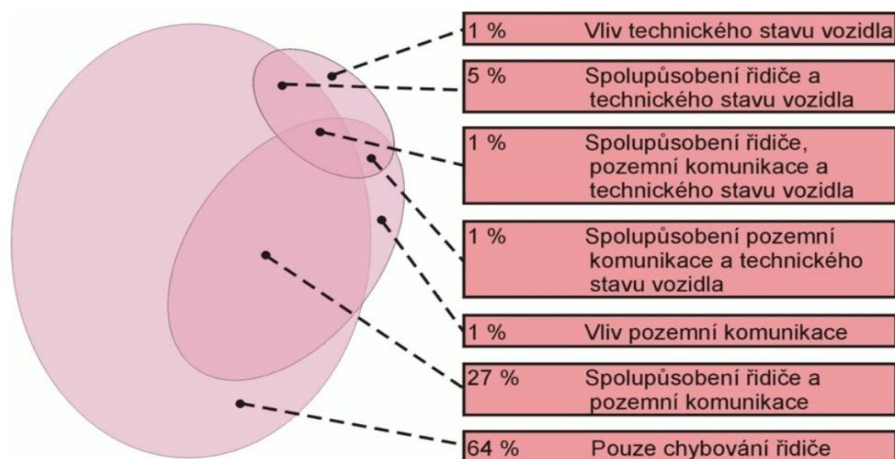
Nejčastější příčinou dopravních nehod je lidský faktor. Vinou ze strany řidiče může být překročení dovolené rychlosti, řízení pod vlivem alkoholu nebo drogu, nepozornost, obecně nepřizpůsobení jízdy vozidla stavu vozovky, resp. viditelnosti. Pozornost řidiče během jízdy nejčastěji narušuje telefonování, psaní SMS, stravování, mluvení s cestujícími, nastavení rádia, klimatizace a další.

Druhou častou příčinou je technická porucha vozidla. Vliv technického stavu dopravního prostředku na nehodovost může být způsoben selháním brzdového systému, opotřebením pneumatik, poruchou řízení, nedostatečným osvětlením a jinými příčinami spojenými s technickou poruchou dopravního prostředku.

Neuspokojivý stav dopravní infrastruktury na úsecích je také častou příčinou vzniku dopravních nehod. Za podprůměrný stav dopravní infrastruktury počítáme na příklad chybné dopravní značení, nedostatečné nebo vadné silniční osvětlení, nevhodné řešení křižovatek a další.

Negativní vliv přírodních faktorů též ovlivňuje růst dopravní nehodovosti na pozemních komunikacích. Déšť, sníh nebo mlha a jiné přírodní faktory omezují viditelnost, zhoršují pohyb vozidla a tím snižují bezpečnost jízdy. [6]

Na obrázku č. 2 lze vidět porovnání několika faktorů, které vedou ke vzniku dopravní nehod. Z obrázku je viditelné, že většina nehod vzniká kvůli chybám řidiče. Menší, ale výrazný vliv v porovnání s ostatními faktory, má spolupůsobení řidiče a pozemní komunikace. Nejmenší procentuální hodnotou na obrázku jsou označeny takové faktory, jako vliv technického stavu vozidla a jeho spolupůsobení s řidičem a s pozemní komunikací.



Obrázek 2: Podíl jednotlivých činitelů na vzniku dopravních nehod za období 2011-2015 [6]

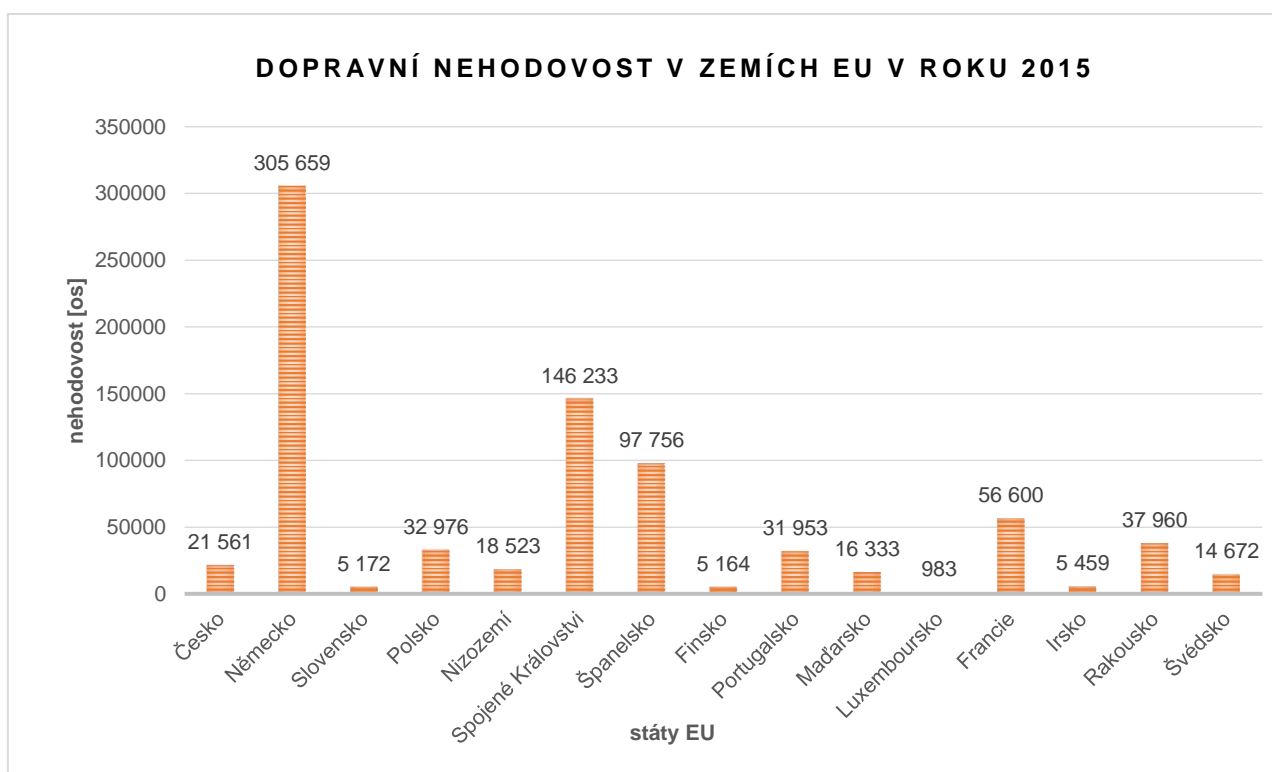
2. Analýza dopravních nehod a jejich druhů

Růst nebo pokles dopravní nehodovosti je ovlivněn různými faktory. Tato kapitola se bude věnovat analýze dopravní nehodovosti ve světě a krajích České republiky a hledání závislosti nehodovosti na několika faktorech: HDP, dopravní infrastruktury a dopravních investic v oblasti silniční infrastruktury.

2.1. Analýza nehodovosti v EU

Dopravní nehodovost v různých státech Evropské Unie se výrazně liší. Pro náhled a analýzu dopravní nehodovosti bylo vybráno 15 států Evropské unie: Česká republika, Německo, Slovensko, Polsko, Nizozemí, Spojené Království, Španělsko, Finsko, Portugalsko, Maďarsko, Lucembursko, Francie, Irsko, Rakousko a Švédsko.

Na obrázku č. 3 je zobrazena evidence nehodovosti ve vybraných státech v roce 2015. Zdrojem dat je Ministerstvo dopravy ČR.



Obrázek 3 Počet nehod v silničním provozu, při kterých došlo ke zranění nebo usmrcení osob [7]

Z obrázku je patrné, že největší počet nehod byl zaznamenán v Německu. Jedním z důvodů je vysoký počet obyvatel v porovnání s ostatními sledovanými státy. Nejnižší dopravní nehodovost je zaznamenána v Lucembursku. Důvodem může být malý počet obyvatel a menší území. V tabulce č. 1 je uveden celkový počet obyvatel ve vybraných zemích ve stejném časovém období, který potvrzuje výše uvedený předpoklad.

Tabulka 1: Dopravní nehodovost a obyvatelstvo v sledovaných zemích EU v roce 2015 [9]

zemí EU	obyvatelstvo	nehodovost
Česko	10 538 275	21 561
Německo	81 197 537	305 659
Slovensko	5 421 349	5 172
Polsko	38 005 614	32 976
Nizozemí	16 900 726	18 523
Spojené Království	64 853 393	146 233
Španělsko	46 449,565	97 756
Finsko	5 471 753	5 164
Portugalsko	10 374 822	31 953
Maďarsko	9 855 571	16 333
Lucembursko	562 958	983
Francie	66 458 153	56 600
Irsko	4 677 627	5 459
Rakousko	8 584 926	37 960
Švédsko	9 995 153	14 672

Pro podrobnější odůvodnění je třeba prozkoumat závislost nehodovosti na vybraných faktorech.

2.1.1. Závislost nehodovosti na HDP

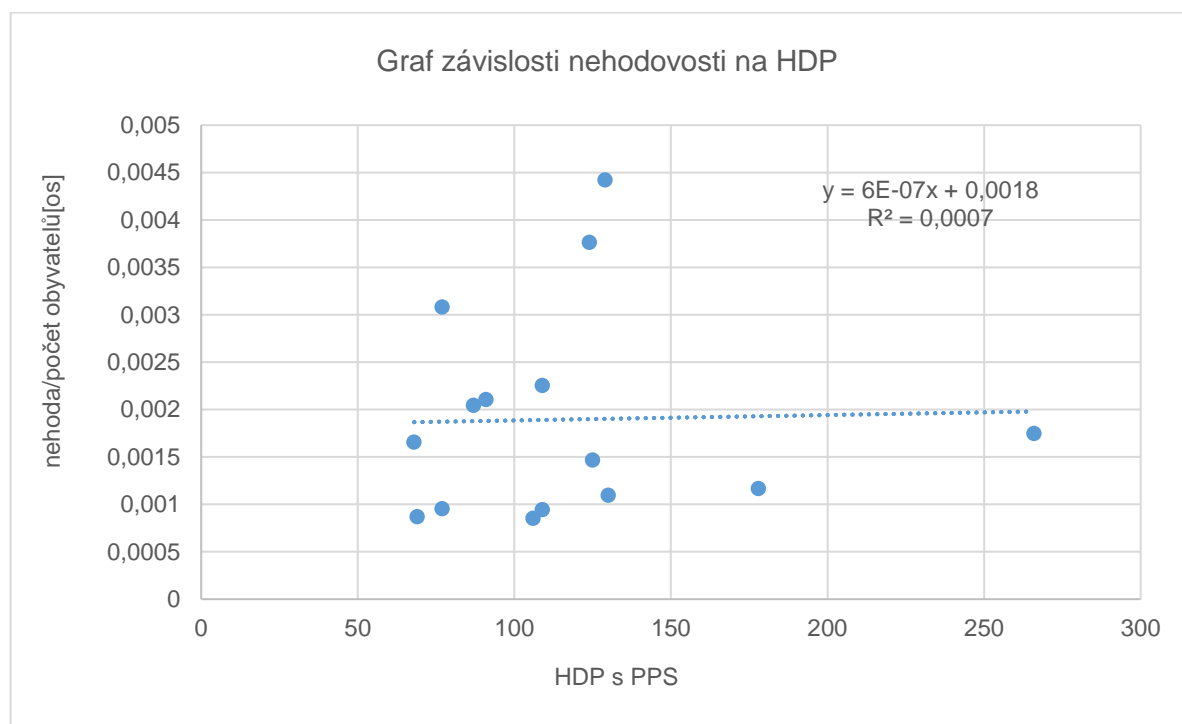
HDP je jedním ze základních makroekonomických ukazatelů. Hrubý domácí produkt představuje hodnotu veškerých finálních statků a služeb vyprodukovaných v dané ekonomice za dané časové období. [8]

Předpokladem je závislost dopravní nehodovosti na makroekonomickém ukazateli HDP. Pro analýzu závislosti HDP na počtu dopravních nehod ve vybraných 15 zemích EU byla získána informace o HDP v paritě kupní síly, počtu obyvatel a již zmíněna informace o dopravní nehodovosti. Data týkající se HDP s PPS a počtu obyvatel ve vybraných zemích byla získána za rok 2015 z webových stránek Eurostatu.

Tabulka 2: nehodovost, počet obyvatelů, HDP v PPS v roku 2015[zdroj: Eurostat [9]

Zemí EU	Nehodovost[os]	Počet obyvatelů[os]	HDP v PPS
Česko	21 561	10 538 275	87
Německo	305 659	81 197 537	124
Slovensko	5 172	5 421 349	77
Polsko	32 976	38 005 614	69
Nizozemí	18 523	16 900 726	130
Spojené království	146 233	64 853 393	109
Španělsko	97 756	46 449 565	91
Finsko	5 164	5 471 753	109
Portugal	31 953	10 374 822	77
Maďarsko	16 333	9 855 571	68
Lucembursko	983	562 958	266
Francie	56 600	66 458 153	106
Irsko	5 459	4 677 627	178
Rakousko	37 960	8 584 926	129
Švédsko	14 672	9 995 153	125

Na obrázku č.4 je reprezentován graf závislosti HDP v paritě kupní síly na osobu a dopravních nehod na osobu v jednotlivých zemích.



Obrázek 4 Závislost HDP v PPS na podílu dopravních nehod na počtu obyvatelů [Zdroje: Ministerstvo dopravy, Eurostat, vlastní výpočty]

Z grafu a z nízkého korelačního koeficientu plyne, že závislost HDP v PPS na podílu počtu dopravních nehod na celkovém počtu obyvatelů je velmi slabá. Nicméně linie trendu na grafu zobrazuje slabě viditelnou závislost mezi dvěma vybranými typy dat. Paradoxně z těchto dat

slabě vyplývá, že čím více HDP, tím více nehodovosti a naopak. Ale tato závislost je na hranici významnosti. Výsledek je velice překvapivý, ale dá se odůvodnit tím, že v zemích s vyšší úrovní HDP je větší počet registrovaných vozidel, což by mohlo ovlivnit zvýšení počtu dopravních nehod.

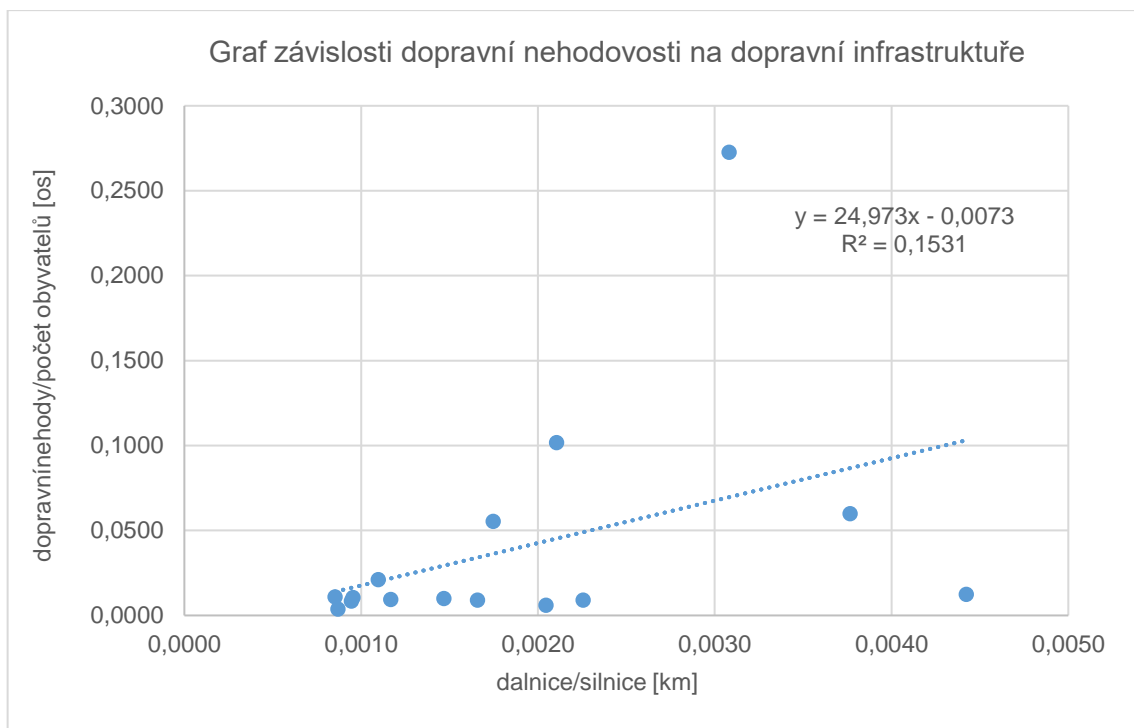
2.1.2. Závislost nehodovosti na dopravní infrastruktuře

Pro hledání závislosti nehodovosti na dopravní infrastruktuře byly získány údaje o délce silnic a dálnic za rok 2015. Zdrojem poskytování této informace je Eurostat.

Tabulka 3: Dopravní infrastruktura ve vybraných zemích EU v roce 2015 [9]

Dopravní infrastruktura v zemích EU 2015		
zemí	Dálnice (km)	Silnice(km)
Česko	776	129 880
Německo	12 993	216 894
Slovensko	463	4 4014
Polsko	1 559	422 303
Nizozemí	2 730	129 532
Spojené Království	3 768	417 491
Španělsko	15 336	150 667
Finsko	881	105 118
Portugalsko	3 065	11 245
Maďarsko	1 884	208 879
Lucembursko	161	2 908
Francie	11 599	1 074 619
Irsko	916	97 982
Rakousko	1 719	138 696
Švédsko	2 119	212 972

Na obrázku č. 5 je představen graf, který reprezentuje závislost dopravní nehodovosti na dopravní infrastruktuře. Analýza byla provedena na základě porovnání podílu dálnic z celkové délky silnic a podílu počtu dopravních nehod na celkovém počtu obyvatel. Linie trendu a koeficient R na grafu potvrzují závislost mezi dvěma vybranými typy dat. Z toho lze usoudit, že čím vyšší úroveň dopravní infrastruktury tím vyšší počet dopravních nehod a naopak. Opět jde o paradoxní závěr, který lze odůvodnit pouze tím, že na větších úsecích je větší intenzita dopravy, což zvyšuje riziko vzniku dopravní nehody.



Obrázek 5 Graf závislosti dopravní nehodovosti na dopravní infrastruktuře [Ministerstvo dopravy, Eurostat vlastní výpočty]

Stejným způsobem byla provedena analýza závislosti dopravní infrastruktury na nehodovosti, ke které došlo k usmrcení. Byla použita stejná data týkající se dopravní infrastruktury (podíl délky dálnic z celkové délky silnic). Místo celkového počtu nehod byl použit počet nehod v jednotlivých zemích, při kterých došlo k usmrcení.

Data o nehodách s usmrcením byla získána z webové stránky Ministerstva dopravy. Získaná data o nehodovosti s usmrcením byla vztažena počet obyvatel v jednotlivých zemích.

Tabulka 4: Podíl nehod s usmrcením na počet obyvatelů a dopravní infrastruktura v roce 2015 [zdroje: Ministerstvo dopravy, Eurostat, vlastní výpočty]

zemí	nehody s usmrcením /počet obyvatelů [os]	dálnice/silnice [km]
Česko	0,000070	0,005974746
Německo	0,000043	0,059904838
Slovensko	0,000057	0,010521676
Polsko	0,000077	0,003691662
Nizozemí	0,000037	0,021075873
Spojené Království	0,000028	0,009026369
Španělsko	0,000036	0,101787385
Finsko	0,000049	0,008381057
Portugalsko	0,000060	0,272565585
Maďarsko	0,000065	0,009019097

Lucembursko	0,000064	0,055364512
Francie	0,000052	0,010793593
Irsko	0,000035	0,009348656
Rakousko	0,000056	0,012394013
Švédsko	0,000026	0,009949665

Závislost se projevila slabší, než při porovnání dopravní infrastruktury na celkovém počtu nehod. Výsledek je stejný jako v předchozí analýze. Tady lze říct, že s růstem dopravní infrastruktury roste počet dopravních nehod s usmrcením a naopak. Lze to odůvodnit tím, že rozvinutá komunikační síť má obvykle větší intenzitu dopravy, což zvyšuje riziko vzniku dopravních nehod s vážnými důsledky.

2.2. Nehody v ČR

Dopravní nehodovost v různých krajích České republiky se liší. Z dat získaných z Ministerstva dopravy ČR je možné odvodit korelaci mezi počtem obyvatel v kraji, hustotou obyvatelstva silniční infrastruktury s nehodovostí.

2.2.1. Porovnání nehodovosti mezi kraji ČR

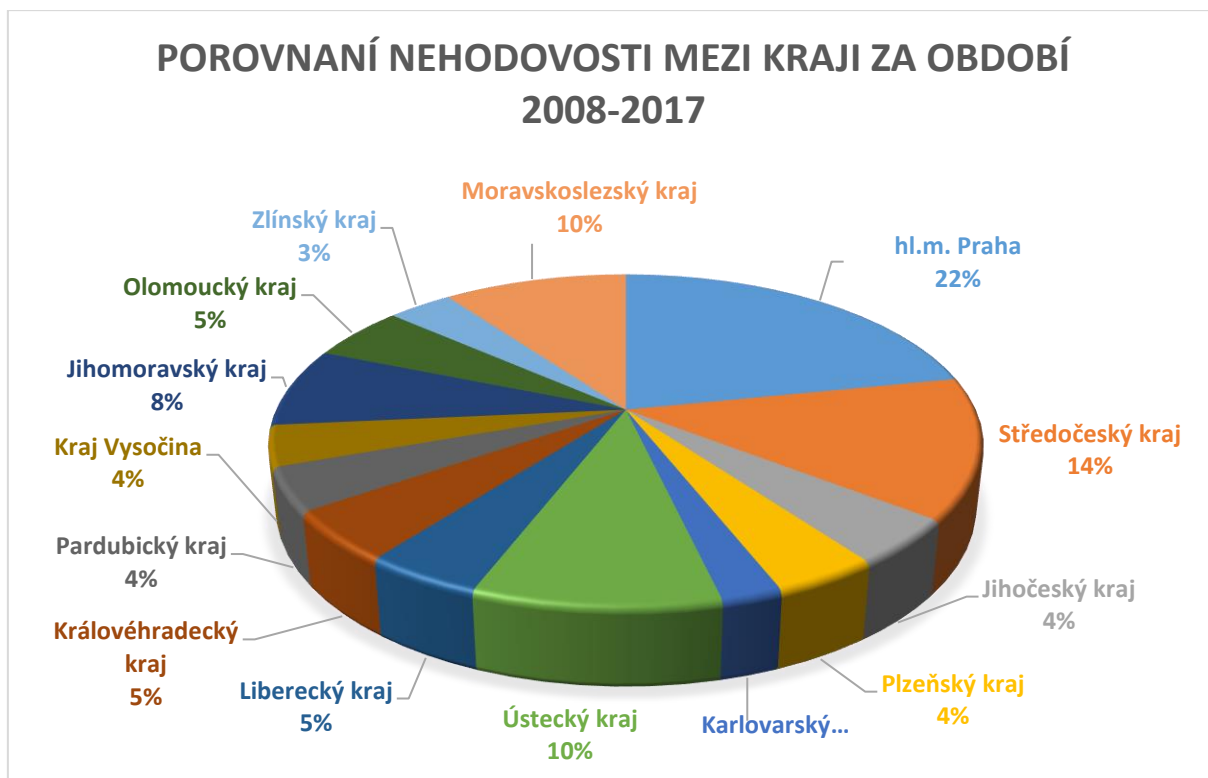
Tabulka č. 5 reprezentuje vývoj silniční nehodovosti v České republice podle krajů za období 2008-2012. Všechny hodnoty byly získány z webových stránek Ministerstva dopravy ČR. Na základě uvedených tabulek byl vytvořen graf, pomocí kterého lze celkově zhodnotit silniční nehodovost v ČR. Graf je znázorněn na obrázku č. 6.

Největší počet dopravních nehod ve vybraném období byl zaznamenán v hlavním městě Praha. Hlavním důvodem je vysoká intenzita dopravy, zapříčiněná především povahou hlavního města, které je logicky centrálním bodem země, kam směřují lidé za celou řadou aktivit. Druhou příčinou je vysoká úroveň turismu v porovnání s ostatními kraji.

Z obrázku č. 6 je vidět že nejmenší počet nehod má Karlovarský kraj. Důvodem může být nízký počet obyvatel v kraji. Karlovarský kraj ve srovnání s ostatními kraji je podle počtu obyvatel na posledním místě. Druhým důvodem může být nízká úroveň dopravní infrastruktury.

Tabulka 5: Počet nehod v silničním provozu, při kterých došlo ke zranění nebo usmrcení osob za období 2008-2017 [7]

kraj/rok	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
hl.m. Praha	30 251	15 583	18 190	16 572	17 795	18 593	19 306	21 462	22 876	23 032
Středočeský kraj	22 053	11 183	9 870	9889	10 595	11 266	11 604	12 463	13 833	14 707
Jihočeský kraj	9 623	3 206	2 899	3015	3 207	3 557	3 753	3 916	4 223	4 301
Plzeňský kraj	8 968	3 217	2 813	3107	3 453	3 121	2 905	3 205	3 352	3 590
Karlovarský kraj	4 658	1 772	1 737	1489	1 396	1 626	1 732	1 996	2 675	2 967
Ústecký kraj	12 294	8 033	7 217	7126	7 551	8 230	8 372	9 707	10 002	10 638
Liberecký kraj	7 088	4 366	3 864	3620	3 859	3 788	3 572	3 889	4 094	4 443
Královéhradecký kraj	7 382	3 692	3 697	3843	4 281	4 164	4 254	4 460	4 774	5 163
Pardubický kraj	6 814	3 501	3 357	3 582	3 726	3 622	3 451	3 482	3 695	3 970
Kraj Vysočina	6 623	1 843	2 390	2 594	3 295	3 696	3 709	4 114	4 151	4 423
Jihomoravský kraj	14 174	3 642	5 650	5941	6 670	6 701	6 950	7 056	7 094	7 587
Olomoucký kraj	8 386	4 407	4 156	4274	4 406	4 432	4 450	4 738	4 979	5 161
Zlínský kraj	5 596	1 798	1 780	2 014	3 025	3 314	3 484	3 680	4 044	4 215
Moravskoslezský kraj	16 466	8 572	7 902	8071	8 145	8 288	8 317	8 899	9 072	9 624



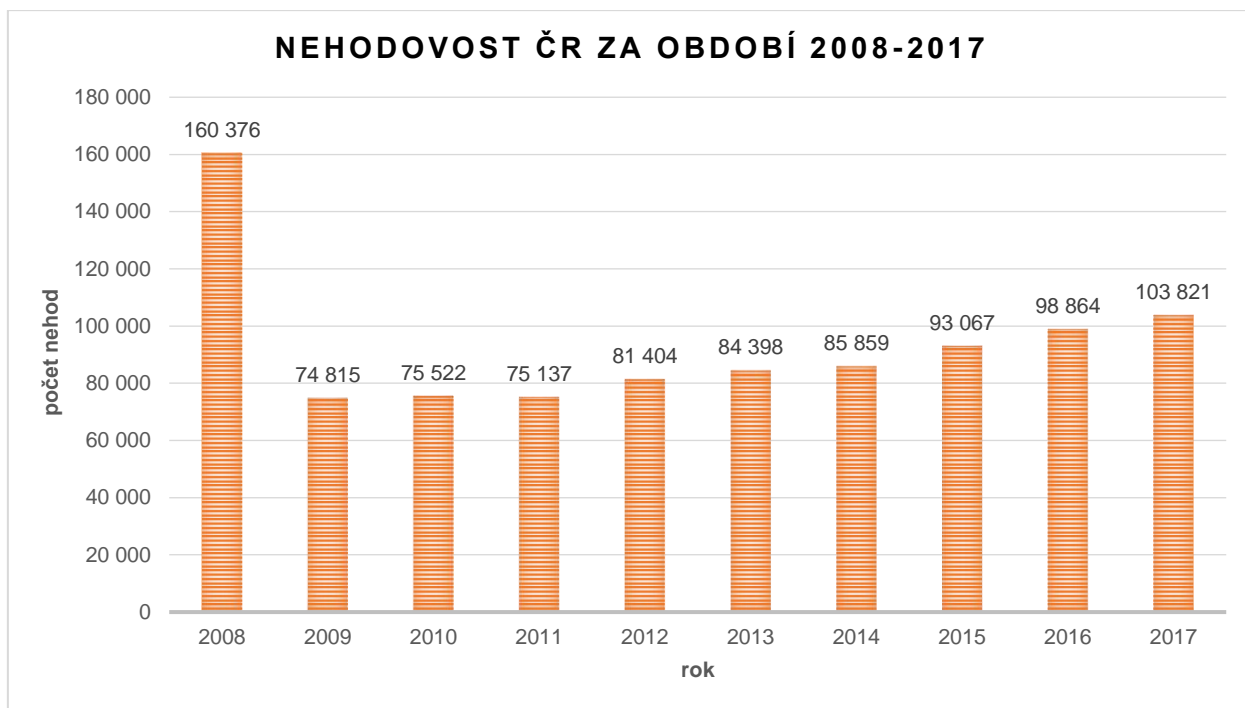
Obrázek 6 Graf nehodovost ČR podle krajů za období 2008-2017 [7]

2.2.2. Dopravní nehodovost v celé ČR

Vývoj dopravních nehod v celé České republice v sledovaném období je uveden na obrázku č. 7. Z obrázku je vidět, že největší počet dopravních nehod v sledovaném období byl zaznamenán v roce 2008 a dosáhl hodnoty 160376. Počet nehod v silničním provozu dle oficiálních statistik po roce 2008 rapidně klesl. V roce 2009 se počet dopravních nehod snížil o hodnotu 85561.

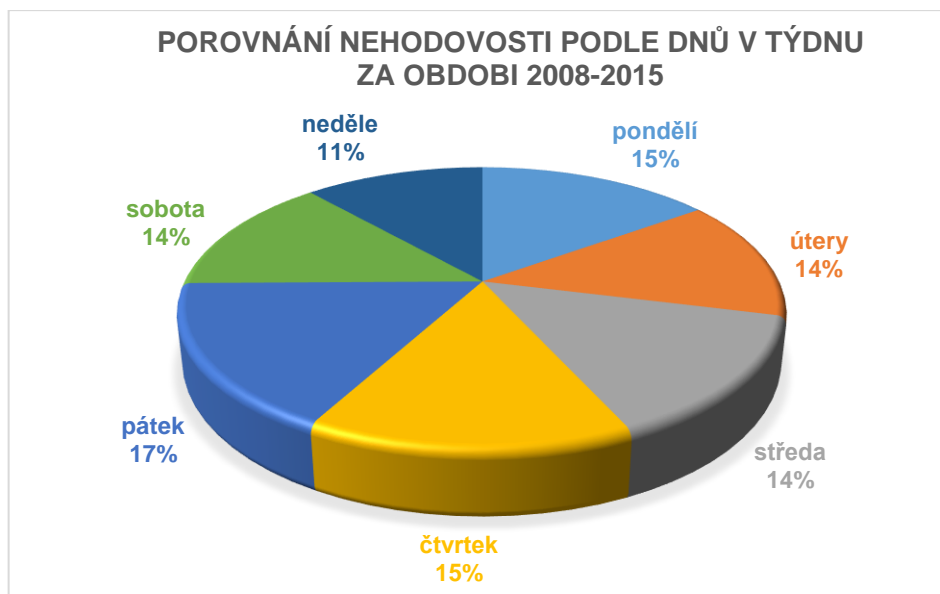
Tento výrazný pokles je především ovlivněn změnou hranice hmotné škody. Od roku 2009 je hranice pro povinné hlášení škody zvýšena na 100000 českých korun. [11]

Od roku 2012 do 2017 dopravní nehodovost měla poměrně rostoucí trend. To mohlo být způsobeno rozvojem dopravní infrastruktury a dlouhodobým nárůstem celkového počtu vozidel na území ČR.



Obrázek 7 Graf nehodovosti v silničním provozu v ČR za období 2008-2017 [7]

Dále bylo provedeno porovnání nehodovosti v ČR podle dnů v týdnu. Údaje byly zajištěny z webových stránek Policie ČR za období 2008-2015. Výsledky reprezentuje graf, znázorněný na obrázku č. 8. Z grafu je vidět, že nehodovost podle dnů v týdnu má celkem stabilní tendence. Nejvíce nehod v silničním provozu obvykle vzniká v pátek. Naopak nízká nehodovost byla vždy zaznamenána v neděli.



Obrázek 8 Graf nehodovosti v silničním provozu podle dnů v týdnu za období 2008-2015 [zdroj Policie ČR, vlastní výpočty]

2.2.3. Závislost nehodovosti na dopravních investicích

Je předpokladem, že míra dopravních investic má souvislost s růstem nebo poklesem dopravní nehodovosti. Pro hledání závislosti dopravní nehodovosti na dopravních investicích byly vybrány hodnoty reprezentující podíl dopravních nehod na počet obyvatel za období 2008-2017 a podíl dopravních dopravních investic na HDP.

Údaje o dopravních investicích v ČR za období 2008-2017 byly získány z webových stránek SFDI. Data zachycují součet celkových investičních výdajů do dopravní infrastruktury a výdajů na opravy a údržbu dopravní infrastruktury. Jednotlivé údaje o dopravních investicích jsou uvedeny v tabulce č. 6.

Tabulka 6: Investiční výdaje do dopravní infrastruktury v sledovaném období [12]

rok	Celkové neinvestiční výdaje do dopravní infrastruktury (mil. Kč)	Celkové výdaje na opravy a údržbu dopravní infrastruktury (silniční) (mil. Kč)
2008	50 962	15 257,2
2009	52 524	15 300,4
2010	43 494	16 941,9
2011	31 799	14 009,4
2012	22 036	14 350,9
2013	16 827	13 334,3

2014	16 632	16 166,5
2015	24 157	18 674,8
2016	22 958	20 742,6
2017	25 911	18 988,6

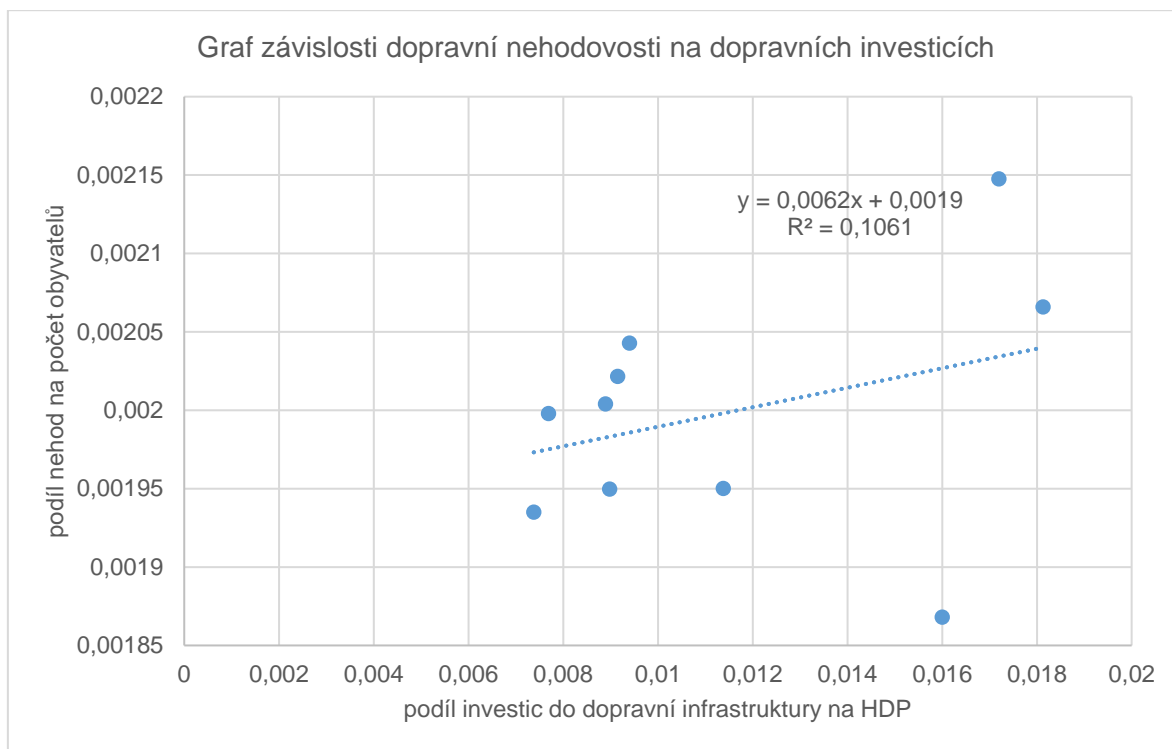
Údaje o HDP v sledovaném období byly získány z webových stránek ČSÚ. Data byla stanovena součtem HDP jednotlivých krajů v příslušných rocích. V tabulce č. 7 je uvedena informace o HDP jednotlivých krajů ČR v sledovaném období.

Tabulka 7: Hrubý domácí produkt podle krajů za období 2008-2017 [11]

	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
hl.m. Praha	977 136	945 796	972 551	1 002 326	999 916	1 016 162	1 037 351	1 112 788	1 193 240	1 283 415
Středočeský kraj	415 250	395 557	405 837	441 181	446 066	447 830	483 511	527 868	552 470	599 821
Jihočeský kraj	195 803	195 195	195 732	202 677	208 741	210 509	218 981	230 508	238 620	247 332
Plzeňský kraj	178 474	176 553	175 398	200 303	197 809	205 986	220 471	232 894	243 908	255 226
Karlovarský kraj	81 467	81 213	79 845	82 883	82 239	83 461	83 049	86 956	89 461	93 482
Ústecký kraj	246 541	250 301	250 267	252 289	249 915	246 103	255 325	275 214	274 254	283 381
Liberecký kraj	120 680	116 297	117 992	128 136	132 309	132 939	138 318	147 208	155 081	160 808
Královéhradecký kraj	173 544	171 660	171 247	182 129	183 998	186 846	196 438	208 560	221 053	243 500
Pardubický kraj	154 384	148 528	146 614	166 279	159 560	161 404	169 049	180 168	186 151	200 687
Kraj Vysočina	152 415	151 101	150 666	162 089	165 427	167 028	170 849	180 155	190 141	199 430
Jihomoravský kraj	397 332	387 479	387 309	417 751	430 768	445 225	465 032	498 757	513 666	531 374
Olomoucký kraj	176 763	172 249	173 882	189 667	191 744	190 587	200 042	215 650	219 892	236 619
Zlínský kraj	185 264	180 885	177 989	189 224	190 571	193 226	210 520	222 918	228 601	237 885
Moravskoslezský kraj	393 358	366 413	369 908	405 476	408 612	398 954	411 950	434 971	466 702	474 307
ČR	3 850 419	3 741 234	3 777 248	4 024 421	4 049 687	4 088 273	4 262 900	4 556 630	4 775 256	5 049 284

Graf závislosti dopravní nehodovosti na dopravních investicích je znázorněn na obrázku č. 9. Jak bylo řečeno dříve, hledání závislosti bylo provedeno na základě porovnání podílu počtu dopravních nehod na celkovém počtu obyvatel a podílu celkových investičních výdajů na HDP. Linie trendu a koeficient R na grafu potvrzují závislost mezi dvěma vybranými typy dat. Na

základě daného výzkumu lze říci, že s růstem investic do dopravní infrastruktury roste dopravní nehodovost. Platí to i naopak, se snižováním investic do dopravní infrastruktury se snižuje úroveň dopravní nehodovosti. Takový výsledek není v souladu s předpokladem. Možným důvodem jsou vyšší intenzity dopravy na nových a kvalitnějších komunikacích. Tento vyšší provoz může vést i k vyššímu počtu dopravních nehod.



Obrázek 9 Graf závislosti dopravní nehodovosti na dopravních investicích [zdroj: SFDI, ČSU, vlastní výpočty]

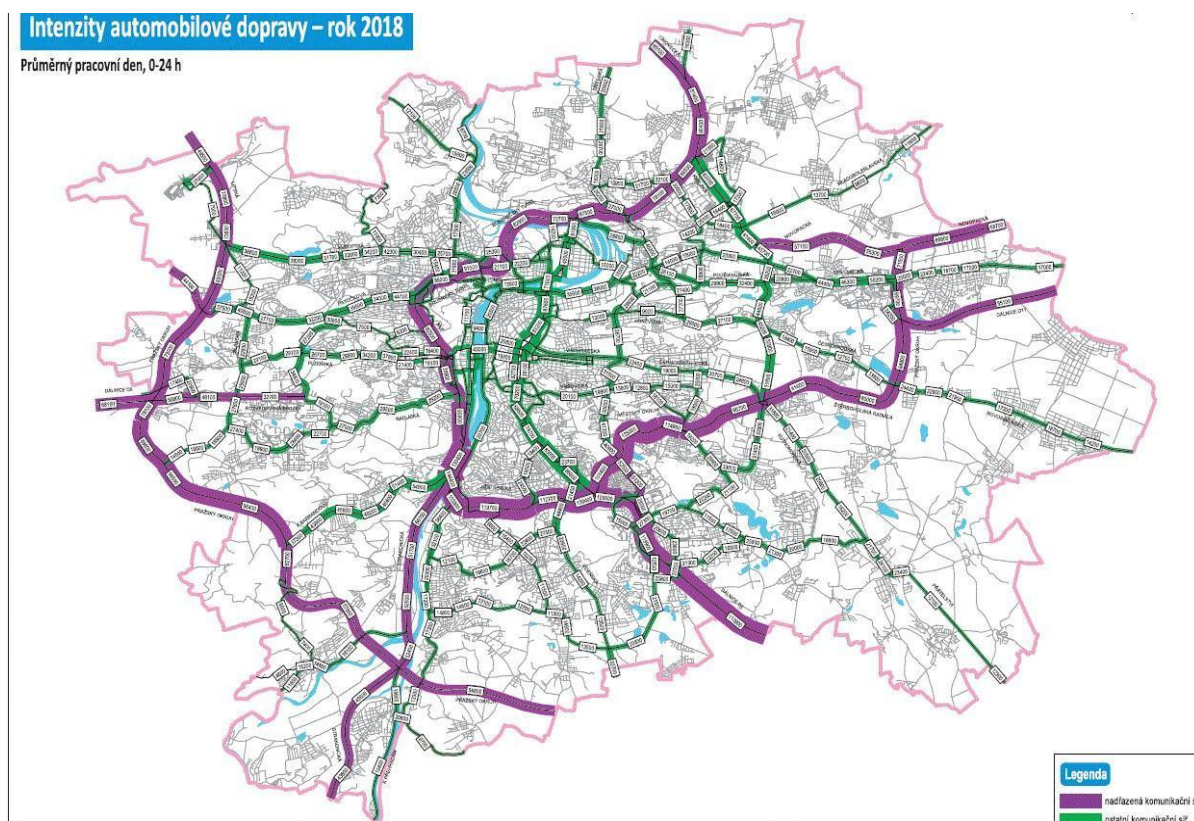
3. Popis nehod v hlavním městě Praha. Vstupní data

3.1. Praha z pohledu dopravy

Z hlediska silniční dopravy se hl. město Praha člení na 2 hlavní okruhy: Pražský (vnější) a Městský (vnitřní). Dané rozdělení plyne z historie města. Navíc, Praha je městem s vysoce nadprůměrnými intenzitami dopravy a dopravními výkony v porovnání s jinými městy v ČR.

Podle dat z Technické správy komunikací dosáhl dopravní výkon v období jednoho pracovního dne v roce 2018 hodnoty 23,006 milionů vozokilometrů. Z toho 91 % je podíl individuální automobilové dopravy. Ranní špičkou je období od 8 do 9 ráno. Odpolední špičku lze pozorovat mezi 16. a 17. hodinou.

Neméně důležitým ukazatelem vývoje automobilové dopravy v Praze je kordonové sledování. Kordonové sledování spočívá v dopravním sčítání všech vstupních komunikací určitého úseku (kordonu). Stejně jako okruhy Prahy, kordony se rozdělují na centrální kordon a kordon vnější. Přehled vnější a vnitřní oblasti je znázorněn na obrázku č. 3



Obrázek 10 Intenzity automobilové dopravy v Praze rok 2018 [16]

Rozloha hlavního města se rovná 498 km². Podle statistiky TSK měla v roce 2018 Praha registrováno celkem 1 308 632 obyvatel. Délka pražské komunikační sítě činí 3 989 km. Míra automobilizace v Praze je jedna z nejvyšších v porovnání s Evropou. Kvalita Pražské hromadné dopravy má vysoké hodnocení na mezinárodní úrovni. Dělna přepravy v městě se sestává z 41% individuální automobilové dopravy a 51% hromadné dopravy.

Informace o současné situaci na komunikační síti, respektive o dopravních nehodách, stupních provozu a jakýchkoli mimořádných situacích a dlouhodobě plánovaných uzavírkách na úsecích v Praze poskytuje dopravní informační centrum.

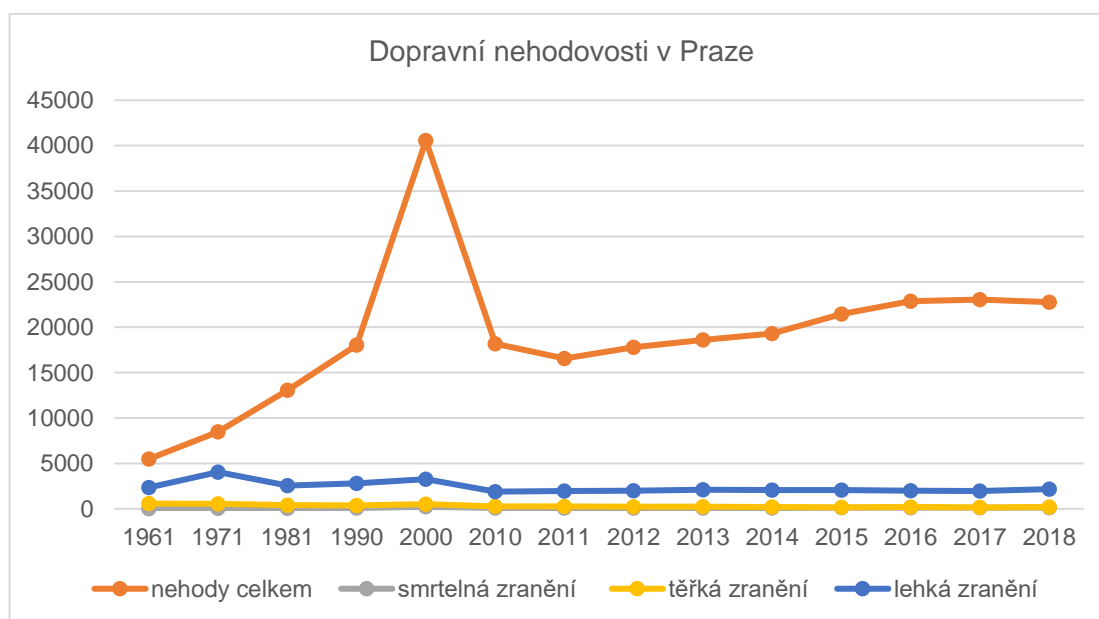
V každém městě občas vznikají dopravní nehody. Bohužel Praha není výjimkou. Nehodovost v Praze je vyšší než v jiných městech ČR. Je to odůvodněno vysokou intenzitou dopravy v porovnání s ostatními městy ČR.

Podle statistik TSK bylo v roce 2018 v Praze zaznamenáno 22 767 nehod. Z nich došlo k úmrtí 31 osob a zranění 2 374 osob. Tyto nehody byly převážně způsobeny chybou řidiče. Mezi nejčastějšími důvody lze pozorovat nedodržení bezpečné vzdálenosti za vozidlem, nedání přednosti při přejíždění z jednoho jízdního pruhu do druhého a alkohol.

V tabulce č. 8 je představena evidence o celkových dopravních nehodách a počtů nehod podle jejich následků. Na základě těchto údajů byl vytvořen graf, který je znázorněn na obrázku č. 11. Z toho je patrné, že počet celkových nehod v období 1961-2000 měl rostoucí tendenci. V roce 2000 byl evidován nejvyšší počet dopravních nehod v porovnání s ostatními roky v tabulce. Dále celková nehodovost až do roku 2011 výrazně klesala. Tento pokles lze odůvodnit ustanovením zákona o provozu na pozemních komunikacích a jeho následných novelizacích. Pozitivem je to, že růst intenzity dopravy nezpůsobuje výrazné zvýšení nehodovosti.

Tabulka 8: Vývoj dopravní nehodovosti v Praze [16]

rok	nehody celkem	smrtelná zranění	těžká zranění	lehká zranění
1961	5 495	30	580	2 361
1971	8 496	47	567	4 046
1981	13 064	72	401	2 572
1990	18 024	100	369	2 806
2000	40 560	225	521	3 260
2010	18 190	101	279	1 893
2011	16 572	92	279	1 955
2012	17 795	99	236	2 009
2013	18 593	103	228	2 116
2014	19 306	107	206	2 070
2015	21 462	119	179	2 078
2016	22 876	129	194	1 983
2017	23 032	128	156	1 951
2018	22 767	126	182	2 165



Obrázek 11 Graf dopravní nehodovosti za období 1961-2018 [16]

Snížit dopravní nehodovost umožňují systémy pro zvýšení bezpečnosti provozu. Jsou to například měřiče rychlostí, systémy kontrolující výšku vozidel, LED návěstidla, atd.

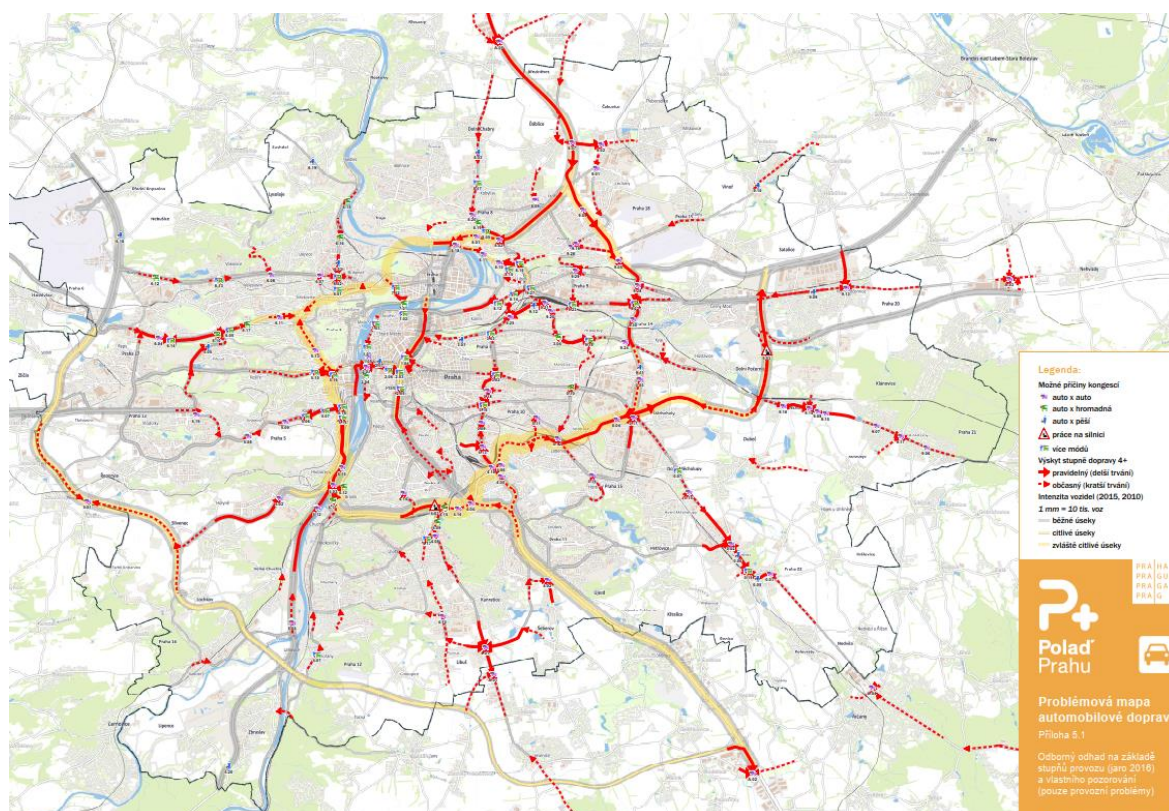
Na zvýšení bezpečnostních opatření v Praze v roce 2008 bylo poskytnuto 68,5 mil. Kč. Navíc ke zlepšení dopravní situace na komunikačních sítích v ČR se otevírají různé projekty financované veřejným sektorem. [16]

3.2. Plán udržitelné mobility Prahy jako řešení zvýšení dopravní bezpečnosti

V současné době Praha v spolupráci se Středočeským krajem přepravuje plán udržitelné mobility, který řeší dopravu jako celek.

Mezi hlavní cíle projektu patří snižování závislosti na fosilních palivech, zlepšení území pro všechny městské funkce, zvýšení účinnosti v přepravě osob a zboží, snížení znečištění životního prostředí a zvýšení bezpečnosti v dopravě. Pod zvýšením bezpečnosti v dopravě je převážně myšleno snížení dopravní nehodovosti.

Z vysokého stupně automobilizace vyplývají vysoké hodnoty intenzit dopravy na pražské komunikační síti v porovnání s jinými městy České republiky. Z důvodu vysokého podílu automobilizace také vychází potřeba efektivně regulovat dopravu tak, aby byla zajištěna únosná míra nehodového rizika. Na obrázku č. 4 jsou vyznačeny problematické úseky hlavního města Prahy.



Obrázek 12: Problémová mapa automobilové dopravy [17]

Jedním ze strategických cílů plánu udržitelné mobility je do roku 2020 snížit počet usmrcených v silničním provozu na úroveň průměru evropských zemí a dále pak o 40 % snížit počet těžce zraněných. Splněním stanoveného cíle se sníží následky dopravních nehod.

Zavedením plánu udržitelné mobility Praha dosáhne pozitivního trendu snížení dopravní nehodovosti několika způsoby:

- likvidací nehodových míst,
- zvýšením dohledu nad provozem na pozemních komunikacích a revitalizací věřených prostranství s prvky zklidnění dopravy

Samotný přesun částí cest na kolejovou síť spolu s návazností pěších a cyklistických tras se také podstatně ovlivní snížení dopravní nehodovosti v městě. Dopravní nehodovost se rovněž sníží pomoci většího nasazení informačních a komunikačních technologií včetně asistenčních systémů ve vozidlech a v dopravní infrastruktuře. Na tato opatření, jejichž cílem je zvýšení bezpečnosti dopravy, je v plánu do roku 2030 vynaložit 600 mil. Kč

V návrhové části plánu udržitelné mobility Prahy se rozlišují standardní opatření a opatření rozvojová.

Cílem standardních opatření je udržení alespoň nynějšího stupně dopravního systému, který je nutné realizovat na nynější infrastruktuře. Avšak tyto opatření nejsou postačující k tempu růstu města. Bezpečnost dopravy v Praze bude zlepšena pomocí takových standardních opatření jako například odstranění billboardů od místních komunikací s rychlostí 70 km/h a více a bezpečnostní úpravy nehodových míst. Tyto opatření mají plošný charakter a kvůli tomu lze očekávat snížení dopravní nehodovosti. Zvláštní pozornost si zaslouží železniční přejezdy, protože díky jejich vysokému počtu nebo nedostatečné míře zabezpečení, tyto přejezdy představují vysoké riziko vzniku dopravních nehod. Bezpečnost dopravy by také mohly ovlivnit takové standardní opatření jako rozšíření parkovacích zón placeného stání, rekonstrukce silniční infrastruktury, řízení dopravy prostřednictvím vlastnictví a správy světelných signalizačních zařízení.

Rozvojová opatření slouží k realizaci zřetelného růstu kvality a udržitelnosti dopravního systému. Návrh rozvojových opatření je potřebný v oblastech s problémy, které nejde vyřešit opravou nebo zlepšením.

Řízení dopravy jako navržené rozvojové opatření bude mít pozitivní vliv na bezpečnost dopravy. V plánu udržitelné mobility jsou uvedeny příklady řízení dopravy, jsou to liniové řízení Pražského okruhu, oblastní dopravně závislé řízení a detekce incidentů a navádění na volná parkovací místa. Tady jde o rozvoj nadřazených adaptovaných systémů, které by měli mít vliv na plynulost a bezpečnost silničního provozu. Vzat v úvahu je také rozvoj a uplatnění kooperativních systémů potřebných ke zvýšení informovanosti a bezpečnosti účastníků silničního provozu, k zefektivnění průjezdů vozidel hromadné dopravy a k zvýšení plynulosti a

bezpečnosti jízdy. Tato opatření rozšiřují oblast řízení dopravy o rozvoj nadřazených adaptivních systémů, které by měly mít příznivý vliv na plynulost a bezpečnost silničního provozu. [17]

3.3. Vstupní data

Pro modelování reálných diskrétních dat byla získána data o nehodovosti v Praze. Poskytovatelem těchto dat je Policie ČR.

Získaná data specifikují nehody podle různých kritérií: časových údajů o dopravní nehodě, datu nahlášení nehody a její lokality, druhu, charakteru, dělení komunikace, chování chodce, druhu vozidla, zavinění nehody, charakteristiky vozidla, následku ve vozidle, následku na životní zdraví chodce atd.

Poskytovaná data jsou příliš rozsáhlá. Proto ke zpracování byla vybrána jen nejlépe vyhovující data podle několika určitých kritérií:

1. 18 DĚLENÍ KOMUNIKACE
2. 24 POČET ZÚČASTNĚNÝCH VOZIDEL

Hlavní příčiny nehody se uvádějí podle různých specifik. Jednou z nich je dělení komunikace. Dělení komunikace člení se na 7 skupin. Podle těchto skupin se ke každému záznamu přiřazuje číslo.

DĚLENÍ KOMUNIKACE:

- 1 dvoupruhová
- 2 tří pruhová
- 3 čtyř pruhová s dělicím pásem
- 4 čtyř pruhová s dělicí čarou
- 5 vícepruhová
- 6 rychlostní komunikace
- 0 žádná z uvedených

Počet zúčastněných vozidel uvádí se jako skutečný počet vozidel.

4. Popis matematických modelů

Analýza dat je důležitým zdrojem informace pro úlohy, ve kterých je potřeba poznávat nebo ovlivňovat reálné procesy, které jsou předmětem lidského zájmu.

Takových úloh je celá řada v lékařství, průmyslů a převážně v dopravě. V řadě případů jsou měřená a analyzovaná data spojitá, jejich hodnoty jsou reálná čísla. Někdy jsou ale data diskrétní. Diskrétní data mají konečný počet různých hodnot a ty se opakují. Taková data pochází typicky z dotazníků, průzkumů nebo záznamů o určité události. Diskrétní data vyžadují odpověď buď ano/ne (1/0) nebo odstupňovaně 1-2-3...n, kde n je konečné číslo. Tato diskrétní data většinou pochází z různých dotazníků. S těmito daty je možno se potkat i v záznamech o dopravních nehodách pořizovaných policií.

Zatímco pro analýzu spojitých dat existuje řada metod, které lze při práci použít, pro diskrétní data jsou statistické metody podstatně chudší. [14]

Tato diplomová práce se věnuje testování diskrétních dat podle několika statistických rozdělení:

- Binomické rozdělení
- Kategorické rozdělení
- Geometrické rozdělení
- Negativní binomické rozdělení
- Poissonovo rozdělení
- Rovnoměrné rozdělení
- Alternativní rozdělení

4.1. Binomické rozdělení

Pravděpodobnost, že v n alternativních pokusech bude y_t úspěchů.

Podmínkou binomického rozdělení je konstantní pravděpodobnost úspěchu p v jednom pokuse.

Model binomického rozdělení:

$$f(p) = \binom{n}{y_t} p^{y_t} (1-p)^{n-y_t}$$

Kde $y_t = 0, 1, 2, \dots, n$; $p \in (0, 1)$ je pravděpodobnost úspěchu v jednom pokusu.

Likelihood ve statistice vyjadřuje pravděpodobnost, že daný soubor pozorování je pravděpodobný pro různé hodnoty statistických parametrů, která indexuje rodinu těchto pravděpodobností. Likelihood je definován jako součin modelů s dosazenými daty z výběru. [15]

Likelihood:

$$L = \prod_{\tau=1}^t \binom{n}{y_\tau} p^{S_\tau} (1-p)^{nk_\tau - S_\tau}$$

Kde $S_t = \sum_{i=1}^t y_i$ a $k_t = t$.

Přepočítání statistik:

$$S_t = S_{t-1} + y_t$$

$$k_t = k_{t-1} + 1$$

Odhad se uskuteční pomocí dat ze simulace. Simulace v programu Scilab je definovaná níže uvedeným vzorkem:

`Y = grand(nr, nc, "bin", n, p)`

Střední hodnotu lze vypočítat podle vzorku:

$$E[y] = np$$

4.2. Kategorické rozdělení

U kategorického rozdělení veličina y_t má n hodnot a každá z nich nastává s konstantní pravděpodobností $p = [p_1, p_2, \dots, p_n]$.

Model kategorického rozdělení:

$$f(y_t|p) = p_{y_t}$$

Kde $y_t = 1, 2 \dots n$ a $p = [p_1, p_2, \dots, p_n]$ je vektor pravděpodobností hodnot y_t .

Likelihood kategorického rozdělení:

$$L = \prod_{i=1}^n p^{V_{i,t}}$$

Kde $V_{i,t}$ je i -tý prvek statistiky $V_t = [V_1, V_2, \dots, V_3]_t$. V_i počítá kolikrát nastala hodnota $y = i$. [15]

V přepočtu statistiky kategorického rozdělení k aktuální složce statistiky se připočítává jednička:

$$V_{t,y_t} = V_{t-1,y_t} + 1$$

Odhadem kategorického rozdělení je normalizace statistiky na součet jedna:

$$\hat{p}_{i,t} = \frac{V_{i,t}}{\sum_j V_{j,t}}$$

Střední hodnotu lze vypočítat podle níže uvedeného vzorku:

$$E[y] = \sum y_i p_i$$

4.3. Geometrické rozdělení

Pro výpočet geometrického rozdělení je potřeba uvažovat alternativní pokusy. Rozdělení určuje pravděpodobnost, že prvnímu úspěchu předchází y_t neúspěchů. Podmínkou geometrického rozdělení je konstantní pravděpodobnost úspěchu p . [15]

Model kategorického rozdělení:

$$f(y_t|p) = p(1-p)^{y_t}$$

$$y_t = 0, 1, 2, \dots, n; \quad p \in (0,1)$$

Likelihood:

$$L = p^{k_t}(1-p)^{S_t}$$

Kde $S_t = \sum_{i=1}^t y_i$ a $k_t = t$.

Přepočet statistik geometrického rozdělení:

$$S_t = S_{t-1} + y_t$$

$$k_t = k_{t-1} + 1$$

Odhad geometrického rozdělení:

$$\hat{p}_t = \frac{k_t}{S_t + k_t} = \frac{1}{(\bar{y} + 1)}$$

Střední hodnota:

$$E[y] = \frac{1-p}{p}$$

4.4. Negativně binomické rozdělení

U negativního binomického rozdělení se provádějí alternativní pokusy. Rozdělení určuje pravděpodobnost, že r -tému úspěchu předchází y_t neúspěchů. Podmínkou negativního binomického rozdělení je konstantní pravděpodobnost úspěchu p . [15]

Model negativně binomického rozdělení:

$$f(y_t|p) = p^r(1-p)^{y_t}$$

Kde r je známa zadaná konstanta. Jinak se model podobá modelu geometrického rozdělení.

Střední hodnota:

$$E[y] = \frac{p^r}{(1-p)}$$

4.5. Poissonovo rozdělení

Poissonovo rozdělení určuje pravděpodobnost, že za jednotku času nastane y_t událostí, jestliže tok událostí λ je konstantní. Událost lze chápat na příklad jako dopravní nehodu.

Model Poissonova rozdělení:

$$f(y_t|\lambda) = \exp\{-\lambda\} \frac{\lambda^{y_t}}{y_t!}$$

Likelihood:

$$L = \exp\{-\lambda k_t\} \frac{\lambda S_t}{\prod_i y_t}$$

Kde $S_t = \sum_{i=1}^t y_t$ a $k_t = t$.

Přepočítání statistik:

$$S_t = S_{t-1} + y_t$$

$$k_t = k_{t-1} + 1$$

Odhad:

$$\hat{\lambda}_t = \frac{k_t}{S_t} = \frac{1}{\bar{y}}$$

Střední hodnota:

$$E[y] = \lambda$$

4.6. Rovnoměrné rozdělení

Rovnoměrné rozdělení, někdy také známé jako pravoúhlé, je rozdělení, které má konstantní pravděpodobnost v hodnotách na pevném intervalu.

Model rovnoměrného rozdělení:

$$f(y_t) = \frac{1}{n}$$

Pro $y_t = L, L + 1, L + 2, \dots, R - 1, R$, kde L, R jsou dolní, horní mez pravděpodobnosti funkce.

- Toto rozdělení modeluje situaci, kdy hodnoty náhodné veličiny jsou zdola i zhora tvrdě omezeny
- V daném intervalu přípustných hodnot nemáme žádné preference.

Statistiky jsou \mathcal{L} a \mathcal{R} s přepočtem:

$$\mathcal{L}_t = \min(\mathcal{L}_{t-1}, y_t)$$

$$\mathcal{R}_t = \max(\mathcal{R}_{t-1}, y_t)$$

Odhad rovnoměrného rozdělení:

$$\bar{L}_t = \mathcal{L}_t = \min(y)$$

$$\bar{R}_t = \mathcal{R}_t = \max(y)$$

Střední hodnota:

$$E[y] = \frac{R+L}{2} \quad [15]$$

4.7. Alternativní rozdělení

Alternativní rozdělení $A(p)$ představuje diskrétní rozdělení pravděpodobnosti náhodné proměnné, která uvažuje hodnotu 1 s pravděpodobností p a hodnotu 0 s pravděpodobností $q = 1 - p$.

Alternativní rozdělení představuje úspěch pokusu (nepřímo i neúspěch) s pravděpodobností $0 < p < 1$. Dané rozdělení nabývá pouze dvou hodnot: úspěch, který je označen 1, a neúspěch, který je označen 0.

$$E[y] = p$$

Náhodná veličina sledující alternativní rozdělení je speciálním případem binomického rozdělení s parametrem $n = 1$. [15]

5. Praktická část, odhad diskrétních dat

Pro modelování diskrétních dat, které jsou základem pro statistickou analýzu, se standardně používá model s kategoričným rozdělením. Tento model je velice obecný – každé diskrétní hodnotě náhodné veličiny přiřazuje její vlastní pravděpodobnost. Lze jím modelovat jak statické, tak i dynamické veličiny.

Nicméně, je tu jeden problém. Model je reprezentován pravděpodobnostní tabulkou, která pro více veličin s více hodnotami má obrovské rozměry. Takovou tabulku nelze ani udržet v paměti počítače. Navíc, pro odhad jejích prvků (což jsou parametry modelu) je potřeba obrovské množství dat.

Proto cílem této kapitoly je pokusit se pro reálná a simulovaná data použít nějaký vhodný model, který by byl nastaven pomocí malého počtu parametrů. Každá z vybraných veličin bude odhadnuta všemi modely. Dále bude vyhodnocena kvalita odhadu (tj. jak dobře daný model vyhovuje popisu této veličiny).

Modelování simulovaných a reálných dat se provádí prostřednictvím počítačového programu Scilab.

Níže je uveden kód, který je základem k potřebnému modelování.

```
// Testování diskrétních veličin
```

```
// -----
```

```
[u,t,n]=file();           // find working directory
```

```
chdir(dirname(n(2)));     // set working directory
```

```
clc, clear, xdel(winsid()); mode(0); getd();
```

```
cat=0; // jestli v odhadu je kategoričké rozdělení ano=1/ne=0
```

```
simul=1; // jestli chceme simulovaná=1 nebo reálná=0 data
```

```
if simul==1
```

```
    disp(' SIMULOVANÁ DATA')
```

```
    // simulovaná data
```

```
    nd=500;
```

```

// generování pomocí kategorických dat

pS=[5 40 50 70 30 5 90];           // TADY SE NASTAVÍ TVAR PR. FUNKCE

pS=pS/sum(pS);                     // normalizace

for t=1:nd

    y(1,t)=sum(rand(1,1,'u')>cumsum(pS))+0; // generování

end

nh=length(pS);                     // počet hodnot testované veličiny

else

    disp(' REÁLNÁ DATA')

    // reálná data

    dt=csvRead('2012_I_L.csv',';'); // načtení reálných dat

    nd=size(dt,1);                  // délka datz

    velicina=18;                   // 7 a 11 nezkoušet

    y=dt(:,velicina);

    y=y-min(y);                    // aby se začínalo od nuly

end

my=mean(y);                        // průměr veličiny y

// počáteční statistiky (ka - počet apriorních dat pro statistiku)

V1.ka=100; V1.S=my*V1.ka;

V2.ka=100; V2.n=6; V2.S=my*V2.ka; // (n - pevné)

V3.v=1e1*ones(1,max(y)+1);

V4.ka=100; V4.S=my*V4.ka;

V5.ka=10; V5.r=8; V5.S=my*V5.ka;

V6.ka=100; V6.S=my*V6.ka;

```



```

V7.L=mean(y); V7.R=mean(y);

// počáteční parametry

th1.p=V1.S/V1.ka;

th2.p=V2.S/(V2.ka*V2.n); th2.n=V2.n;

th3.p=V3.v/sum(V3.v);

th4.p=V4.S/V4.ka;

th5.p=V5.S/V5.ka; th5.r=V5.r;

th6.La=V6.S/V6.ka;

th7.L=1; th7.R=6;

L=zeros(7,1);

w=ones(7,1);

tx=['1 Alt';'2 Bin';'3 Categ';'4 Geo';'5 NegB';'6 Poiss';'7 Unif'];

// cyklus odhadování

for t=1:nd,

    if y(t)<my, y1=0; else y1=1; end

    p(1)=1e-3;           // alternativní rozdělení vynecháme

    p(2)=ProxBinN(y(t),th2);

    if cat==1

        p(3)=ProxCateg(y(t),th3);

    else

        p(3)=1e-3;

    end

    p(4)=ProxGeoS(y(t),th4);

    p(5)=ProxNegB(y(t),th5);

```

```

p(6)=ProxPoiss(y(t),th6);

p(7)=ProxUniD(y(t),th7);

wp=p/sum(p);           // výpočet vah

wt(:,t)=wp(:);        // pamatování vah

[th2,V2]=EstBinN(w(2),y(t),V2); Th2(t)=th2.p;

if cat==1

    [th3,V3]=EstCateg(w(3),y(t),V3);

end

[th4,V4]=EstGeoS(w(4),y(t),V4);

[th5,V5]=EstNegB(w(5),y(t),V5);

[th6,V6]=EstPoiss(w(6),y(t),V6);

[th7,V7]=EstUniD(w(7),y(t),V7);

end

// Výsledky

// komponenta, která měla nejčastěji maximální váhu

[wm,m]=max(wt,'r');

vL=vals(m);

[xxx,jj]=max(vL(2,:));

j=vL(1,jj);

printf('nejlepší shoda: %s\n',tx(jj))

disp '-----'

disp 'Četnosti maximálních vah'

disp(tx')

disp(vL)

```

```

// časový průběh vah komponent

set(scf(1),'position',[20 400 500 300])

plot(wt')

title 'Váhy'

legend(tx,3);

// Predikce

// -----

N=th2.n; p=th2.p;

x2 = grand(1, nd, "bin", N, p);    // binomické

p=th3.p;

x3=zeros(1,nd);

for i=1:nd

    x3(i)=sum(rand(1,1,'u')>cumsum(p))+1; // kategorické

end

p=th4.p;

x4 = grand(1, nd, "geom", p);    // geometrické

r=th5.r; p=th5.p;

x5 = genNegB(p,r,1,nd);        // negativně binomické

La=th6.La;

x6 = grand(1, nd, "poi", La);    // poissonovo

L=th7.L; U=th7.R;

x7=(U-L)*rand(1,nd,'u')+L;    // rovnoměrné

disp ''

// Odhad středních hodnot

```

```

printf(' Odhadnuté stredni hodnoty\n')

E2bin=th2.p*V2.n

E4geo=(1-th4.p)/th4.p

E5nbi=th5.p*th5.r/(1-th5.p)

E6poi=th6.La

E7uni=(th7.L+th7.R)/2

Edata=mean(y)

// histogramy: data vs. predikce

nh=max(y)+1;

set(scf(),'position',[100 300 400 200])

histplot(nh,y);

title data

set(scf(),'position',[250 340 400 200])

histplot(nh,x2);

title binomicke

set(scf(),'position',[400 380 400 200])

histplot(nh,x3);

title kategoricke

set(scf(),'position',[550 420 400 200])

histplot(nh,x4);

title geometricke

set(scf(),'position',[700 460 400 200])

histplot(nh,x5);

title negBinomicke

```

```
set(scf(),'position',[850 500 400 200])
```

```
histplot(nh,x6);
```

```
title Poissonovo
```

```
set(scf(),'position',[1000 540 400 200])
```

```
histplot(nh,x7);
```

```
title rovnomerné
```

Funkce odhadu binomického rozdělení:

```
function [th, V]=EstBinN(w, y, V)
```

```
// estimation of Neg.binom. distr.
```

```
V.S=V.S+w*y;
```

```
V.ka=V.ka+w;
```

```
th.p=V.S/(V.ka*V.n);
```

```
th.n=V.n;
```

```
endfunction
```

Funkce odhadu kategorického rozdělení:

```
function [th, V]=EstCateg(w, y, V)
```

```
// estimation of Categorical distr.
```

```
if (y+1>=1) & (y+1<=length(V.v))
```

```
V.v(y+1)=V.v(y+1)+w;
```

```
else disp('Error in EstCateg'); end
```

```
th.p=V.v/sum(V.v);
```

```
endfunction
```

Funkce odhadu geometrického rozdělení:

```
function [th, V]=EstGeoS(w, y, V)
```

```
// estimation of Geometric distr.
```

```
V.S=V.S+w*y;
```

```

V.ka=V.ka+w;
yb=V.S/V.ka;
th.p=1/(1+yb);
endfunction

```

Funkce odhadu negativně binomického rozdělení:

```

function [th, V]=EstNegB(w, y, V)
// estimation of Neg.binom. distr.
V.S=V.S+w*y;
V.ka=V.ka+w;
th.p=V.S/(V.r*V.ka+V.S);
th.r=V.r;
endfunction

```

Funkce odhadu Poissonova rovnoměrného rozdělení:

```

function [th, V]=EstPoiss(w, y, V)
// estimation of Poiss. distr.
V.S=V.S+w*y;
V.ka=V.ka+w;
th.La=V.S/V.ka;
endfunction

```

Funkce odhadu rovnoměrného rozdělení:

```

// estimation of discrete uniform distr.
function [th, V]=EstUniD(w, y, V)
if y<V.L
V.L=V.L+w*(y-V.L);
end
if y>V.R
V.R=V.R+w*(y-V.R);
end

```

```
th.L=V.L;
```

```
th.R=V.R;
```

```
endfunction
```

Tento program je obecný. Je napsán pro modely: alternativní, binomický, kategorický, geometrický, negativně binomický, poissonovský a rovnoměrný. V provedených experimentech byly ale modely alternativní a kategorický vynechány. Důvodem je to, že alternativní model má jen dvě hodnoty 0 a 1 a testované veličiny jsou vícehodnotové. Kategorickému modelu se vyhýbáme záměrně, protože, jak již bylo dříve řečeno, je reprezentován tabulkou, která pro více hodnot má obrovské rozměry. Úkolem je najít vhodné řešení k modelování nehodových dat, které nebude vyžadovat využití kategorického rozdělení.

Program je napsán jak pro simulovaná, tak i pro reálná data. Volbou parametru **simul** rozhodujeme, v jakém režimu program pracuje. simul=1 pracuje se simulovanými daty, simul=0 s reálnými nehodovými daty.

V testech se mění počet dat „nd“ a tvar pravděpodobnostní funkce „pS“. Tyto proměnné mají nezanedbatelný vliv na modelování, což by mělo se odrazit na následujících výsledcích zkoumání.

5.1. Odhad simulovaných diskrétních dat

Prvními pro modelování a odhad diskrétních dat byla použita simulovaná data. Zvolení parametru **simul** „0“ nebo „1“ v podmínce programového kódu definuje typ dat použitých pro modelování. Simulovaná data v kódu byla volena pomocí simul=1. Na základě toho bylo provedeno několik testů, ve kterých došlo k modelování simulovaných diskrétních dat.

V prvním testu byla zkoumaná data s počtem hodnot 500. Tvar pravděpodobnostní funkce v simulaci byl nastaven vzestupně. V závislosti na střední hodnotě „y“ byly nastaveny počáteční parametry a statistiky.

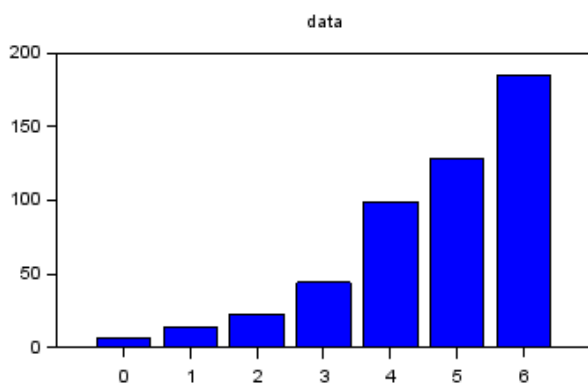
Tvar nenormované pravděpodobnostní funkce rostoucích simulovaných dat:

```
pS=[5 10 20 40 60 90 150]
```

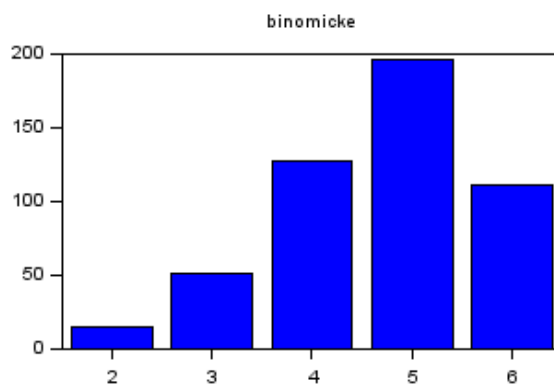
Na základě výsledků provedeného testu bylo vybráno binomické rozdělení. Alternativní a kategorické rozdělení byly vyloučeny záměrně v průběhu kódování. Odhadnutá střední hodnota u většiny modelu se rovná 4,676. Střední hodnota rovnoměrného rozdělení

v porovnání s ostatními modely je velmi odlišná a rovná se hodnotě 3. Z toho důvodu vylučujeme dané rozdělení z výběru. Geometrické rozdělení odpovídá simulovaným datům hůře než binomický model, nicméně dané rozdělení by také mohlo být vhodné pro použití. Stejná situace nastává u modelu s negativně binomickým rozdělením. Níže jsou uvedeny obrázky, které reprezentují histogramy vybraných simulovaných dat a také jsou zobrazeny modely, které byly vytvořeny na základě využití těchto dat.

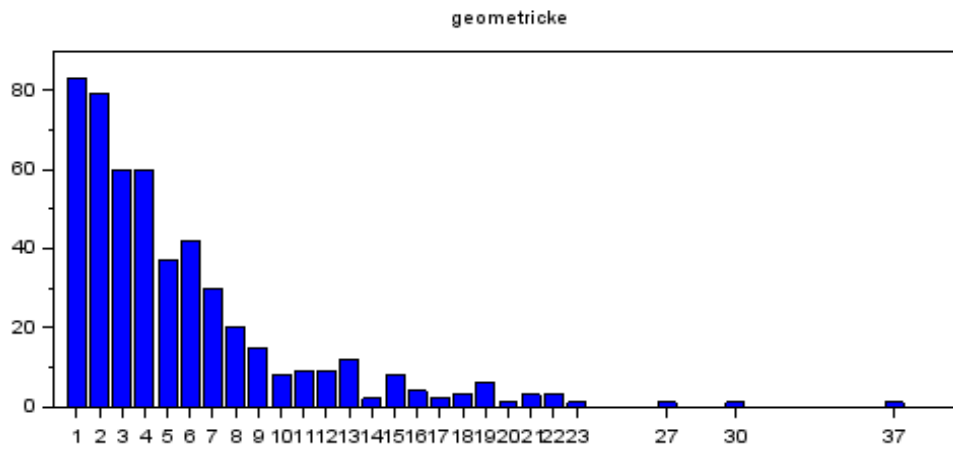
Z histogramu na obrázku č. 13 je vidět že data se generují v rostoucím uspořádání. Na obrázku č. 14 je zobrazen histogram binomického modelu, který je ve výsledku daného pokusu nejlepším. Je vidět, že tvar histogramu binomického rozdělení co nejvíce napodobuje hodnoty simulované pravděpodobnostní funkce. Na obrázku č. 15 a 16 jsou představeny modely geometrického a negativně binomického rozdělení, jejich výsledky jsou horší než u binomického modelu, ale také by tyto modely mohly být vhodné pro použití. Z histogramů je vidět, že výsledné hodnoty modelu slabě napodobují tvar simulované pravděpodobnostní funkce. Výsledkem modelování Poissonova rozdělení je histogram, znázorněný na obrázku č. 17. Poissonův model v tomto pokusu nesplňuje podmínky modelování. Obrázky č. 18 a 19 představují histogramy kategorického a rovnoměrného rozdělení, které se jednoznačně vylučují z výběru. Histogramy těchto rozdělení vůbec neodpovídají tvaru modelovaných dat.



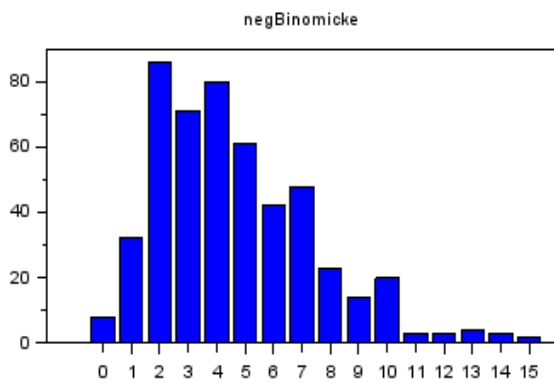
Obrázek 13 Histogram posloupnosti simulovaných dat, vlastní výpočty, Scilab



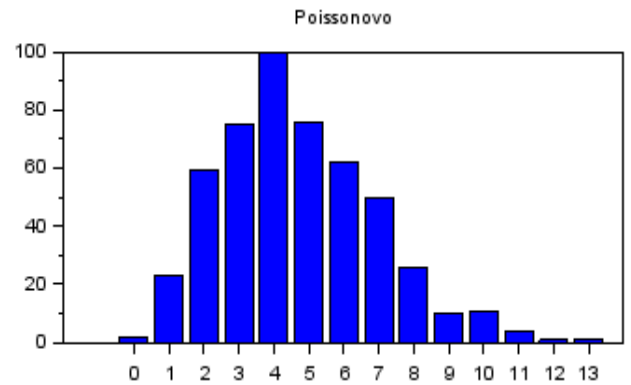
Obrázek 14 Histogram binomického modelu, vlastní výpočty, Scilab



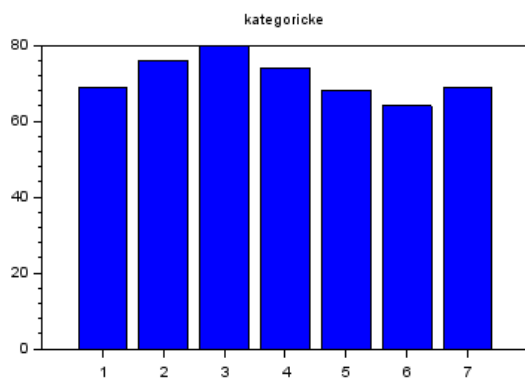
Obrázek 15 Histogram geometrického modelu, vlastní výpočty, Scilab



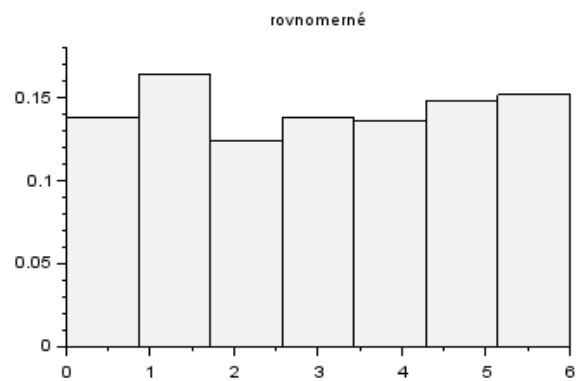
Obrázek 16 Histogram negativně binomického modelu, vlastní výpočty, Scilab



Obrázek 17 Histogram Poissonova rozdělení, vlastní výpočty, Scilab



Obrázek 18 Histogram kategorického modelu, vlastní výpočty, Scilab



Obrázek 19 Histogram modelu s rovnoměrným rozdělením u, vlastní výpočty, Scilab

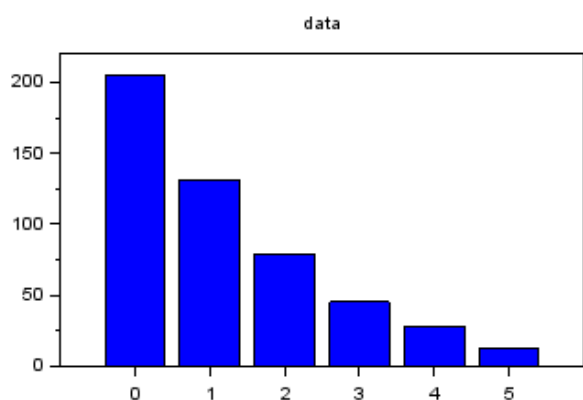
V druhém testu byla vybrána sestupná posloupnost pravděpodobnostní funkce pro simulaci dat. Počet dat „nd“ byl nastaven na hodnotu 500. V závislosti na střední hodnotě byly nastavené počáteční parametry a statistiky.

Data byla generována pomocí kategoričkého modelu s (nenormovanou) pravděpodobnostní funkcí ve tvaru:

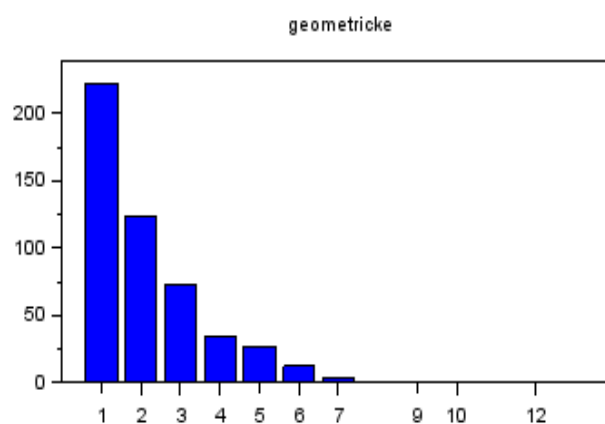
$$pS = [150 \ 90 \ 60 \ 40 \ 20 \ 10 \ 1]$$

Ve výsledku daného pokusu lze říci, že použitým datům více odpovídá geometrické rozdělení. Binomické, geometrické, negativně binomické a Poissonovo rozdělení mají stejnou střední hodnotu, která je shodná se střední hodnotou nastavených simulovaných dat. Tato střední hodnota se rovná 1,192. Střední hodnota rovnoměrného rozdělení je 2,5. Tato hodnota je velice odlišná od veličin ostatních modelů. Z toho důvodu rovnoměrné rozdělení v tomto testu nevyhovuje. Model s binomickým rozdělením je druhou vhodnou možností pro modelování nastavených dat, jelikož jeho výsledek se blíží výsledku binomického rozdělení, které je nejvíce vhodné pro modelování. V porovnání s geometrickým a binomickým modely, výsledky negativně binomického a Poissonova rozdělení jsou nesrovnatelně slabší, proto dané rozdělení nebudou vzaty v úvahu při výběru vhodného modelu.

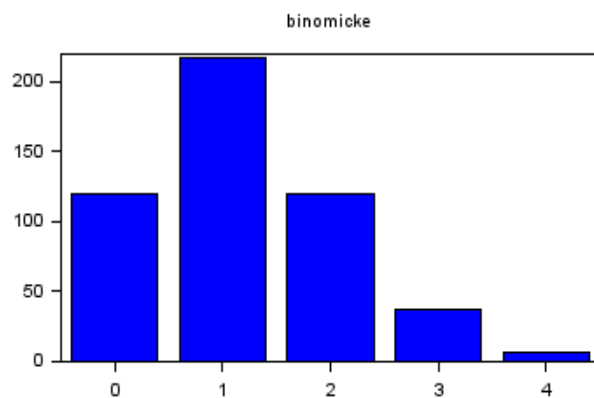
Na obrázku č. 20 jsou představeny histogramy klesajících modelovaných dat. Obrázek č. 21 představuje histogram nejvíce vhodného v daném pokusu modelu s geometrickým rozdělením. Z obrázku je vidět, že histogram geometrického modelu je nejlepším způsobem napodobuje tvaru modelované funkce. Na obrázku č. 22 je znázorněn histogram binomického modelu, který se méně shoduje s tvarem vybraných dat, ale také může být použité v daném modelování.



Obrázek 20 Histogram klesajících simulovaných dat, vlastní výpočty, Scilab



Obrázek 21 Histogram geometrického modelu, vlastní výpočty, Scilab



Obrázek 22 Histogram binomického modelu

V následujících pokusech byla použita data, která měla rostouce-klesající tendence. Počet dat daného modelu se rovná 500. Počáteční parametry a statistiky byly nastavené s ohledem na střední hodnotu y .

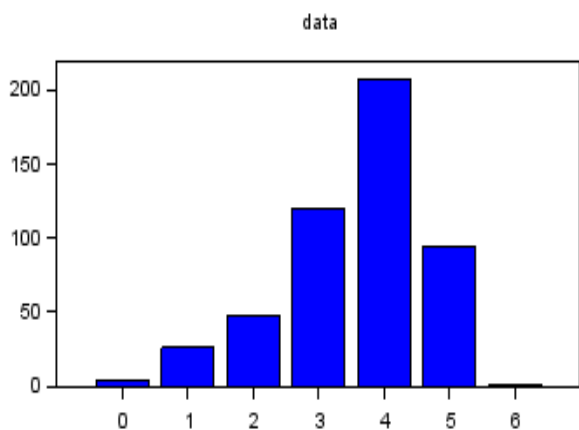
U modelování třetího testu byla data generována pomocí pravděpodobnostní funkce ve tvaru kopečku s vrcholem vpravo. V závislosti na střední hodnotě funkce byly nastavené počáteční parametry a statistiky.

Řádkový vektor tvaru pravděpodobnostní funkce:

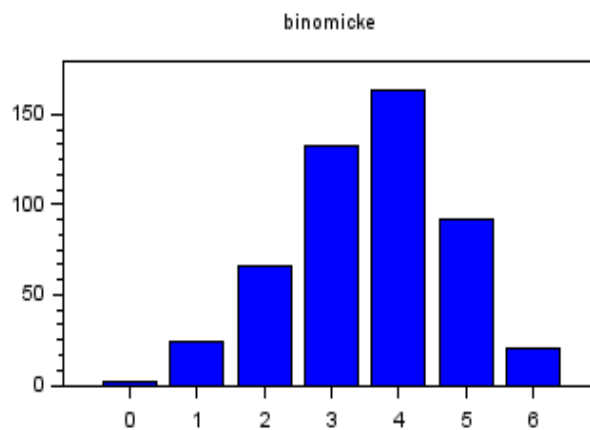
$$pS=[1 \ 10 \ 20 \ 40 \ 70 \ 40 \ 1]$$

Ve výsledku třetího pokusu bylo vybráno binomické rozdělení jako nejlepší možnost pro využití daného modelu. Rovnoměrné, alternativní a kategorické rozdělení byla eliminována. Alternativní a kategorické rozdělení byla vyloučena záměrně v průběhu kódování. Rovnoměrné rozdělení bylo vyloučeno z důvodu odlišnosti jeho střední hodnoty od hodnot ostatních modelů. Střední hodnota většiny modelů je rovna 3,578, zatímco střední hodnota rovnoměrného modelu je 3. Poissonovo a geometrické rozdělení mají horší výsledky než binomický model, nicméně také by mohly být vhodné pro použití.

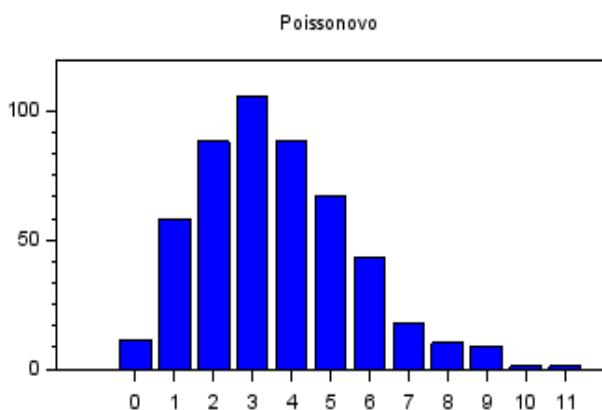
Níže jsou zobrazeny histogramy nastavených dat a modelů. Obrázek č. 23 reprezentuje histogram simulovaných dat. Na obrázku č. 24 je představen histogram binomického modelu, který nevíce vhodný pro daný pokus. Obrázky č. 25 a 26 zobrazují modely, které by mohly nahradit binomický model.



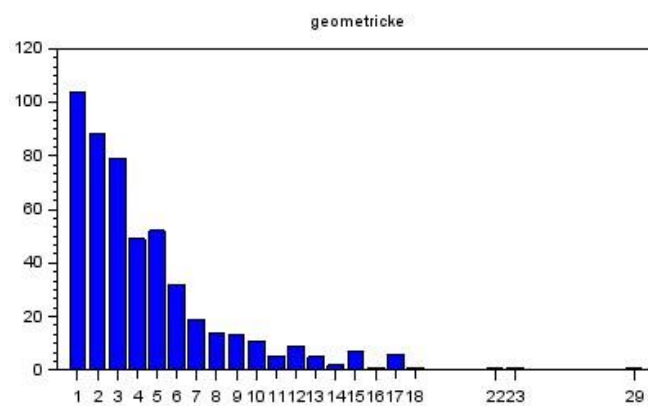
Obrázek 23 Histogram simulovaných dat, vlastní výpočty, Scilab



Obrázek 24 Histogram binomického modelu, vlastní výpočty, Scilab



Obrázek 25 Histogram Poissonova modelu, vlastní výpočty, Scilab



Obrázek 26 Histogram geometrického modelu, vlastní výpočty, Scilab

V čtvrtém pokusu simulovaná data byly zvolené tak, aby model pravděpodobnostní funkce měl podobu kopečku s vrcholem vlevo.

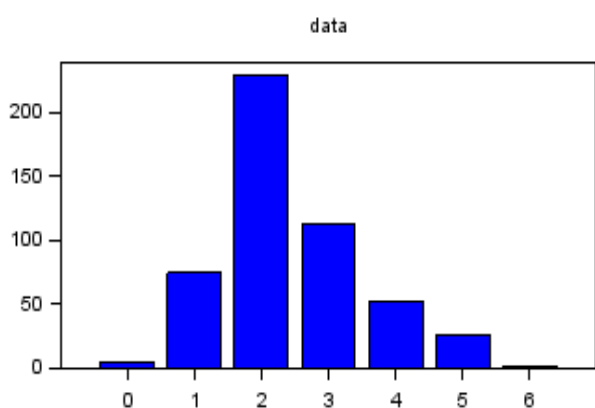
Řádkový vektor tvaru pravděpodobnostní funkce:

$pS = [1 \ 30 \ 70 \ 40 \ 20 \ 10 \ 1]$

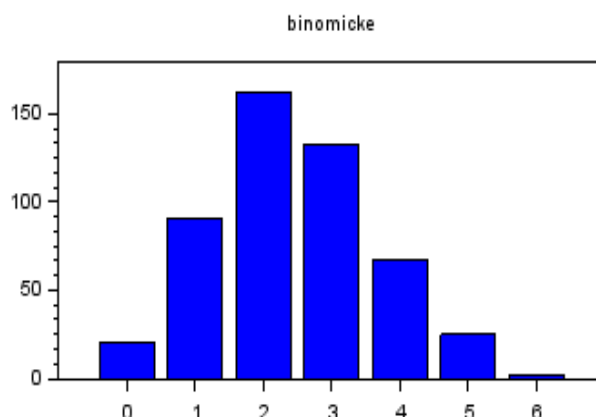
Ve výsledku testu zvítězilo binomické rozdělení. Poissonův model má horší výsledky než model s binomickým rozdělením, ale také by mohl být vhodný pro použití. V důsledku testu také lze poznamenat možnost využití geometrického modelu, ale jeho výsledky jsou horší než výsledky Poissonova rozdělení. Negativně binomické rozdělení v daném testu nemá významný vliv na fungování modelu. Modely popsané výše mají střední hodnotu 2,434. Tato

veličina je rovna střední hodnotě tvaru pravděpodobnostní funkce. Střední hodnota rovnoměrného rozdělení je odlišná od ostatních, její hodnota se rovná 3.

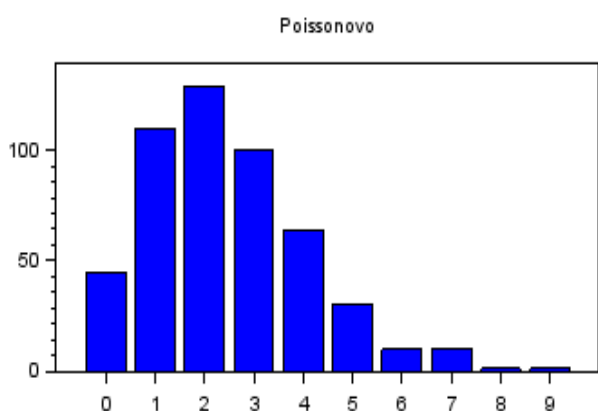
Histogram tvaru pravděpodobnostní funkce je představen na obrázku č. 27. Na obrázku č. 28 je zobrazen histogram binomického rozdělení. Náhradní možnosti (modely s Poissonovým a geometrickým rozdělením) jsou znázorněny na obrázcích č. 29 a 30.



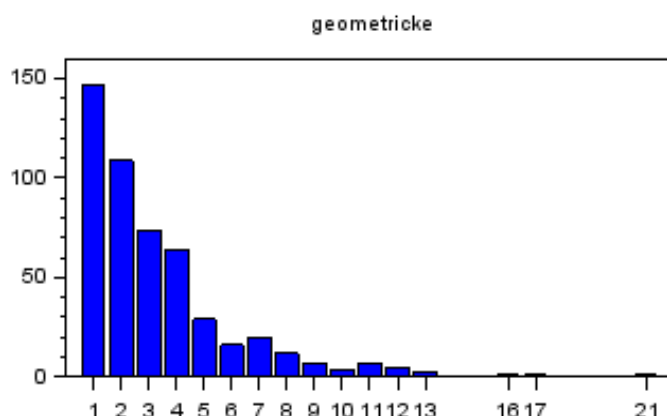
Obrázek 27 Histogram simulovaných dat, vlastní výpočty, Scilab



Obrázek 28 Histogram binomického modelu, vlastní výpočty, Scilab



Obrázek 29 Histogram Poissonova modelu, vlastní výpočty, Scilab



Obrázek 30 Histogram geometrického modelu, vlastní výpočty, Scilab

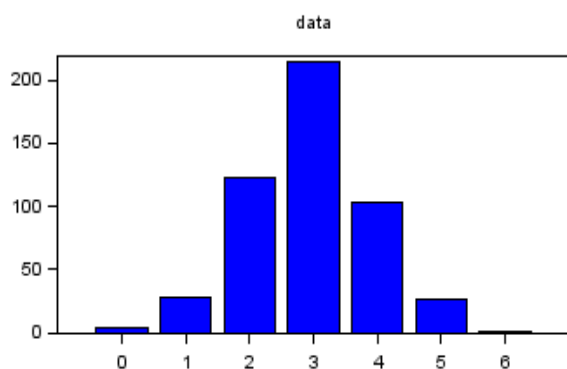
V pátém testu simulovaná data byla generována modelem s pravděpodobnostní funkcí ve tvaru kopečku s vrcholkem uprostřed. Tedy vektor pravděpodobnostní funkce má tvar:

$$pS=[1 \ 10 \ 40 \ 70 \ 40 \ 10 \ 1]$$

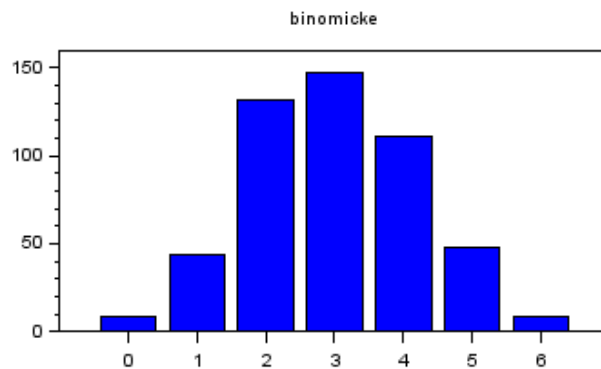
Výsledky ukázaly, že pro modelování pátého testu je nejvíce vhodný model s binomickým rozdělením. Význam modelů s geometrickým a Poissonovým rozdělením v daném testu je

relativně stejný. Tyto modely odpovídají simulovaným datům hůře než binomický model, ale také by mohly být vhodné pro použití. V testu se nevyklučují ani modely se střední hodnotou 2,934. Střední hodnota rovnoměrného rozdělení neodpovídá průměru předchozích modelů. Z toho důvodu bylo konstatováno, že rovnoměrné rozdělení není vhodné pro dané modelování.

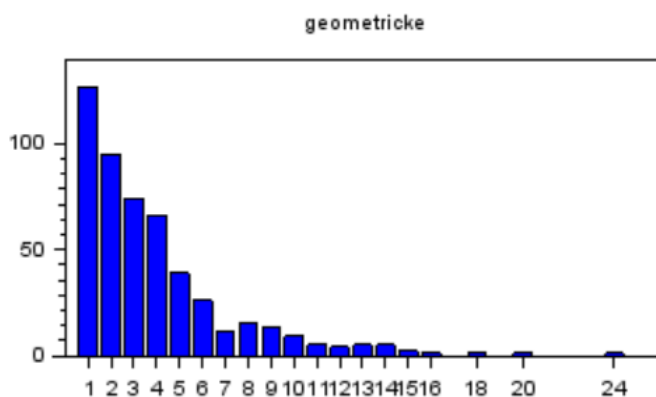
Histogram modelovaných dat je představen na obrázku č. 31. Je dobře vidět, že histogram binomického rozdělení imituje podobu předchozího histogramu simulovaných dat (obr. 32). Histogramy s Poissonovým a geometrickým rozdělení mají slabší souvislost s formou uspořádání simulovaných dat. (obr. 33, 34)



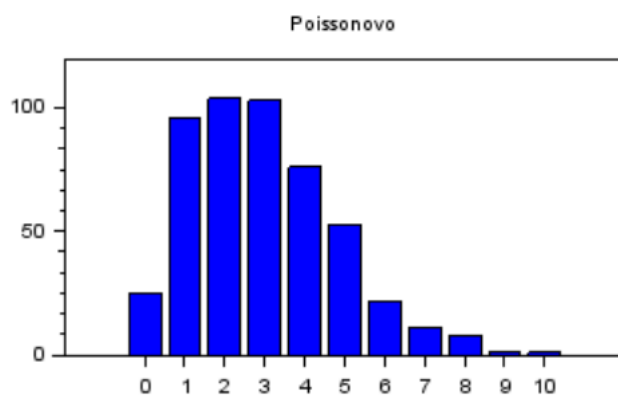
Obrázek 31 Histogram simulovaných dat, vlastní výpočty, Scilab



Obrázek 32 Histogram binomického model, vlastní výpočty, Scilab



Obrázek 33 Histogram geometrického modelu, vlastní výpočty, Scilab



Obrázek 34 Histogram Poissonova modelu, vlastní výpočty, Scilab

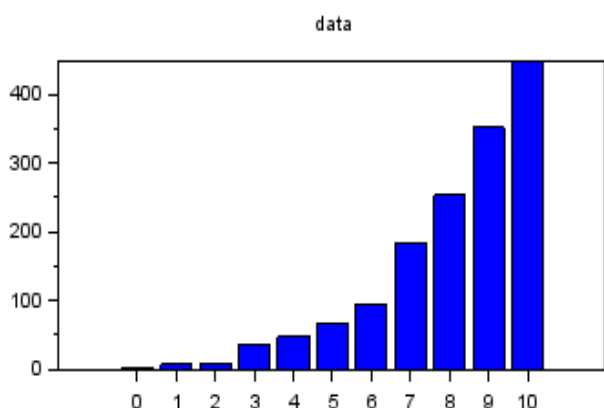
Dále byly provedeny testy s výběrem většího počtu hodnot. Tentokrát byl počet hodnot u simulovaných dat nastaven na hodnotu 1500. Dále byl u testů změněn tvar pravděpodobnostní funkce a na základě programového kódu byl určen nejvhodnější model.

V prvním testu s počtem hodnot "nd" nastaveným na 1500, simulovaná data byla generována v níže uvedené posloupnosti:

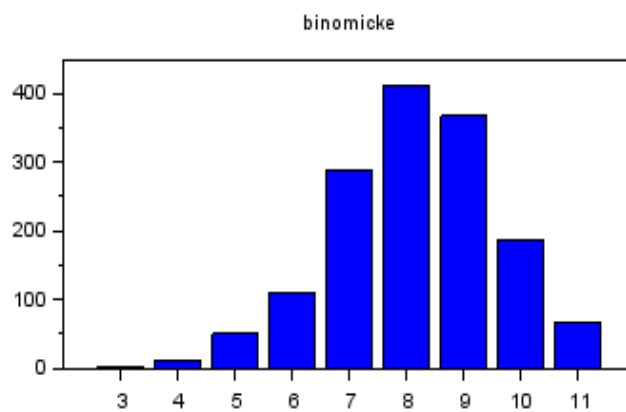
pS=[1 5 10 25 40 60 100 150 220 300 400]

Výsledky testu ukázaly, že binomické rozdělení je nejvíce vhodné pro modelování vybraných simulovaných dat. Dále se v modelování dobře projevilo Poissonovo rozdělení. Jeho výsledky jsou horší než u binomického modelu, ale dané rozdělení by také mohlo být použitelné. V daném pokusu bylo rozhodnuto o vyloučení kategorického a alternativního rozdělení, jejich výsledky neodpovídají tvaru modelovaných dat. Rovnoměrné rozdělení bylo vyloučeno z výběru, protože jeho střední hodnota je velmi odlišná od hodnot ostatních rozdělení. Geometrické a negativně binomické rozdělení se v daném modelování projevily slabě. Proto při výběru vhodného modelu bylo toto rozdělení odmítnuto.

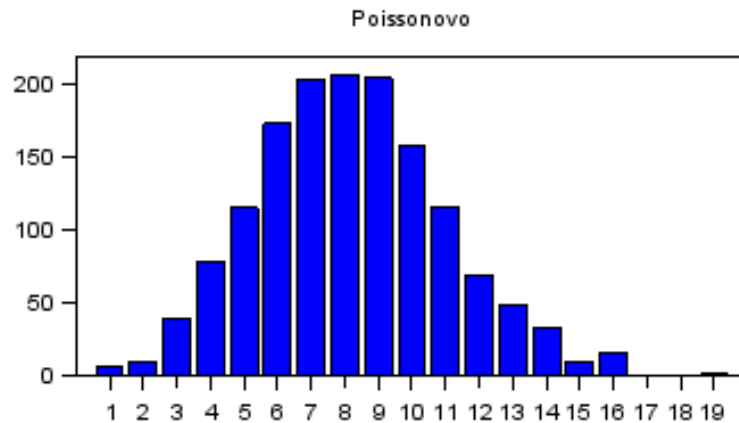
Níže jsou představeny histogramy dat a rozdělení, která byla použita pro dané modelování. Simulovaná diskretní data v daném modelu mají rostoucí tendenci (obr. 35). Z histogramu binomického rozdělení je vidět, že se jeho výsledné hodnoty co nejvíce přizpůsobují hodnotám simulovaných dat (obr. 36). Histogram Poissonova rozdělení je představen na obrázku č. 37. Jeho tvar je horší odpovídá hodnotám posloupnosti modelované pravděpodobnostní funkce, nicméně je vidět že mezi Poissonovým modelem a modelovanými daty se projevuje zákonitost.



Obrázek 35 Histogram simulovaných dat, vlastní výpočty, Scilab



Obrázek 36 Histogram binomického rozdělení, vlastní výpočty, Scilab



Obrázek 37 Histogram Poissonova rozdělení, vlastní výpočty, Scilab

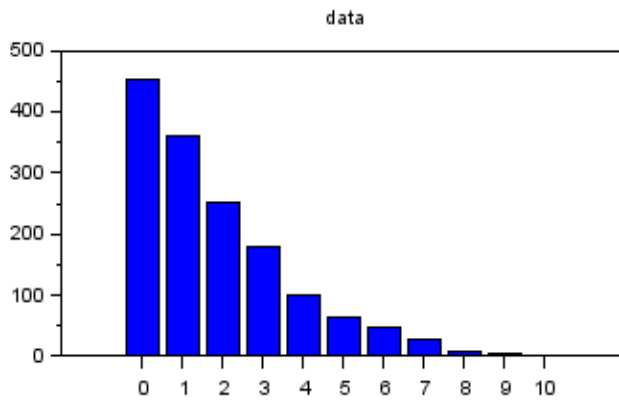
Sedmý pokus byl také proveden s nastavením hodnot $nd=1500$. V daném testu byla simulovaná data zvolena v klesající posloupnosti. Řádkový vektor „pS“ tedy má tvar:

$pS=[400. 300. 220. 150. 100. 60. 40. 25. 10. 5. 1.]$

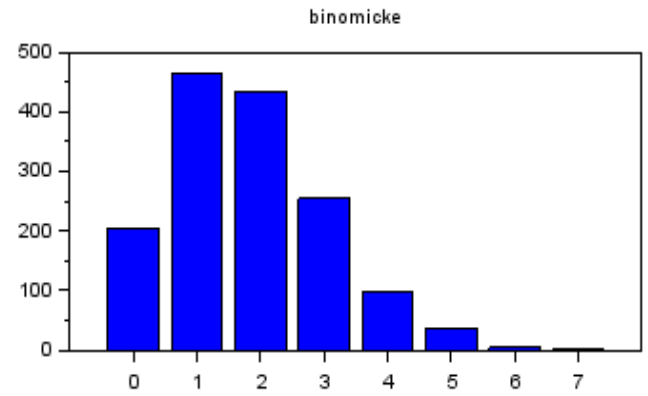
Ve výsledku modelování simulovaných dat bylo vybráno binomické rozdělení. Geometrický model se také může hodit pro modelování, ale jeho výsledky jsou horší než u binomického. Proto v daném testu lze hodnotit model s geometrickým rozdělením jako náhradní možnost k použití. Poissonovo a negativně binomické rozdělení mají výsledky na hranici významnosti. Z toho důvodu bylo rozhodnuto daná rozdělení nezohlednit. V daném pokusu došlo k vyloučení kategorického a alternativního rozdělení, jelikož jejich výsledky neodpovídají tvaru modelovaných dat. Rovnoměrné rozdělení bylo rovněž vyloučeno z výběru, protože jeho střední hodnota je velmi odlišná od hodnot ostatních rozdělení. Průměr rovnoměrného rozdělení je 5, zatímco střední hodnota ostatních je 1,814.

Na obrázku č. 38 je ukázán histogram modelovaných dat. Z histogramu je vidět že simulovaná data jsou seřazena od největšího k nejmenšímu. Histogram binomického modelu je představen na obrázku č. 39. Z histogramu je vidět, že hodnoty binomického modelu nejlépe napodobují tvar simulovaných hodnot. Histogram geometrického rozdělení je znázorněn na obrázku č. 40. Je vidět, že hodnoty histogramu odpovídají tvaru simulovaných dat. Obrázky č. 41 a 42 představují histogramy Poissonova a negativně binomického rozdělení. Na obrázcích je vidět, že obě rozdělení se snaží přizpůsobit posloupnosti modelovaných dat, ale vychází to mnohem hůře, než u modelu s binomickým a geometrickým rozdělením. Na obrázcích č. 43 a 44 jsou představeny modely kategorického a rovnoměrného rozdělení. Histogramy daných

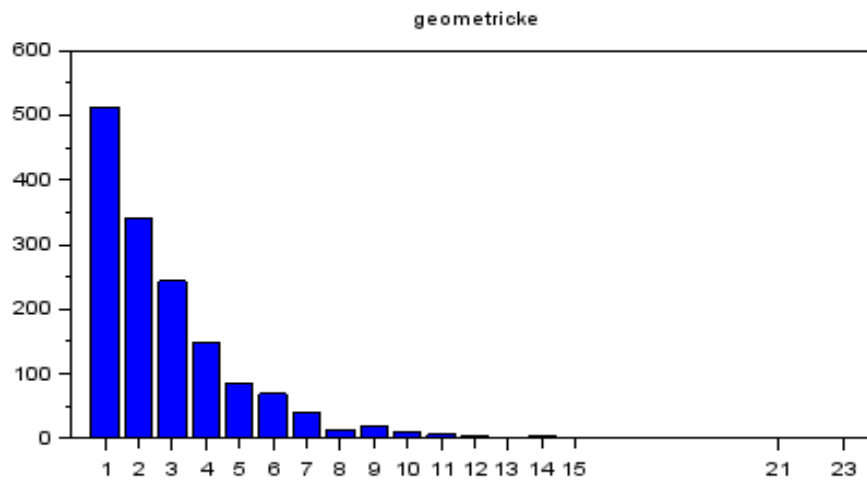
rozdělení mají úplně jiný tvar než histogramy předchozích modelů. Hodnoty histogramů kategoričkého a rovnoměrného rozdělení vůbec nesouvisí z posloupnosti simulovaných dat.



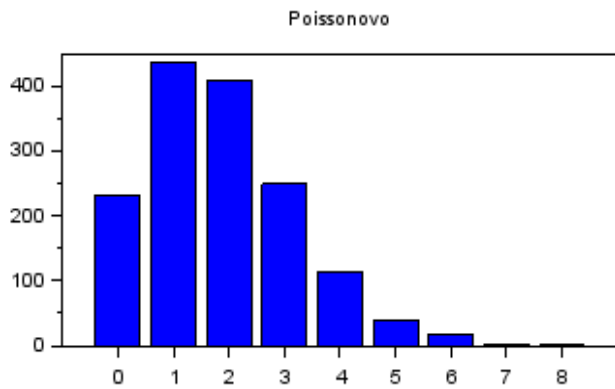
Obrázek 38 Histogram simulovaných dat, vlastní výpočty, Scilab



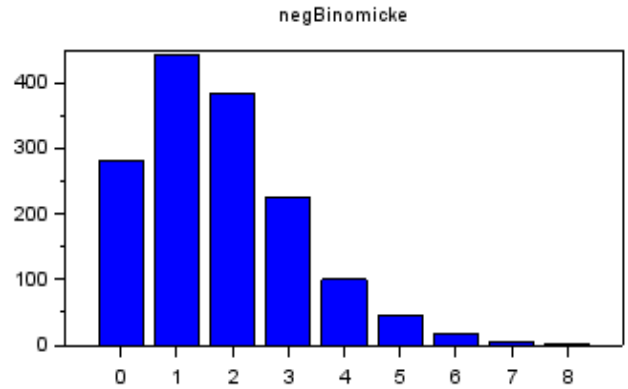
Obrázek 39 Histogram binomického modelu, vlastní výpočty, Scilab



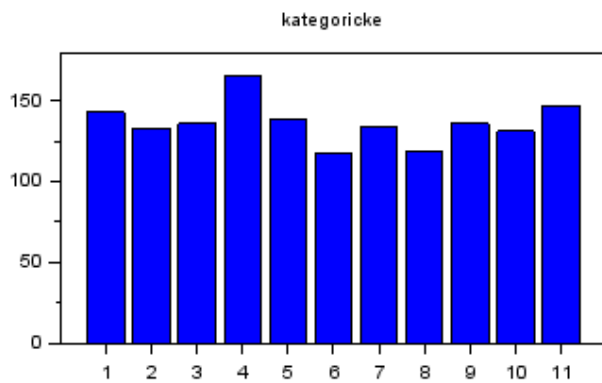
Obrázek 40 Histogram geometrického modelu, vlastní výpočty, Scilab



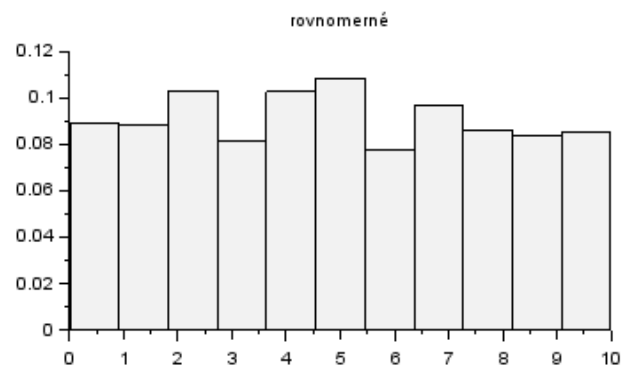
Obrázek 41 Histogram Poissonova modelu, vlastní výpočty, Scilab



Obrázek 42 Histogram negativně binomického modelu, vlastní výpočty, Scilab



Obrázek 43 Histogram modelu s kategoričným rozdělením, vlastní výpočty, Scilab



Obrázek 44 Histogram modelu s rovnoměrným rozdělením, vlastní výpočty, Scilab

Dále byl proveden osmý test. U daného pokusu byla vybrána data, napodobující tvar kopečku s vrcholkem vpravo. V závislosti na střední hodnotě byly nastaveny počáteční parametry a statistiky.

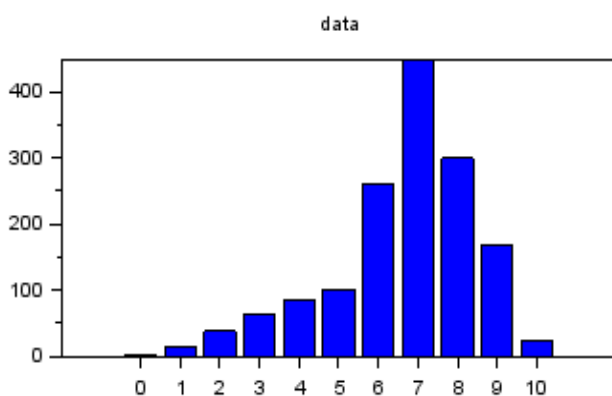
Posloupnost určující pravděpodobnostní funkci generovaných dat:

$pS=[1. 5. 10. 20. 30. 40. 90. 150. 100. 60. 10.]$

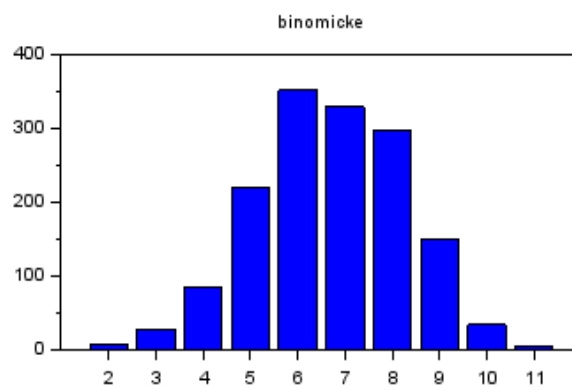
Stejně jako v předchozím pokusu, se k modelu simulovaných dat nejvíce hodí binomické rozdění. Poissonovo rozdění má horší výsledky než binomický model, nicméně dané rozdění by také mohlo být vhodné pro použití. Model s geometrickým rozdělením by také mohl být vhodný pro použití, avšak je třeba upřednostnit binomické a Poissonovo rozdění, protože jejich výsledky jsou vhodnější. Negativně binomické rozdění se v daném modelování projevilo slabě. Proto při výběru vhodného modelu bylo rozhodnuto ho zanedbat. Kategoričké rozdění bylo záměrně vyloučeno pomoci nastavení „cat“ na hodnotu „0“ ve výše

uvedeném kódu programu. Rovnoměrné rozdělení není vhodné pro daný model. Jeho střední hodnota je odlišná od střední hodnoty simulovaných dat.

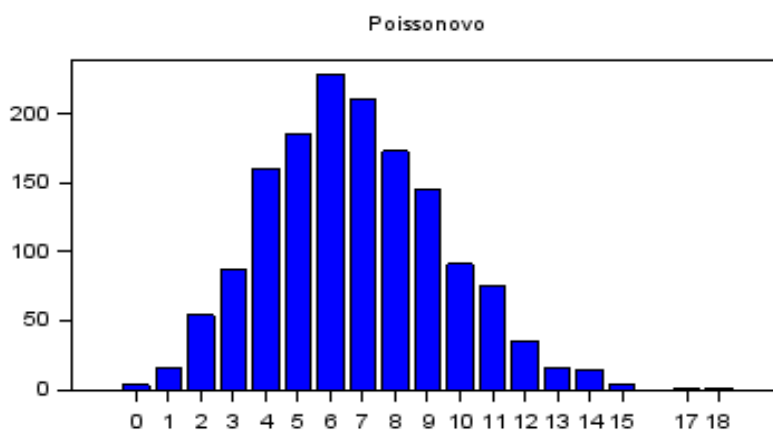
Na obrázku č. 45 představen histogram posloupnosti simulovaných dat. Na obrázku č. 46 je zobrazen histogram binomického modelu, který je ve výsledku daného pokusu nejlepším. Z histogramu binomického rozdělení je vidět, že jeho výsledné hodnoty co nejvíce se přizpůsobují hodnotám simulovaných dat. Výsledkem modelování Poissonova rozdělení je histogram, znázorněný na obrázku č. 47. V porovnání s binomickým modelem, histogram modelu s Poissonovým rozdělením se méně podobá hodnotám z histogramu simulovaných dat. Histogram geometrického rozdělení, znázorněný na obrázku č. 48, se shoduje s tvarem vybraných dat, ještě méně než Poissonův model. Nicméně je také vhodný pro dané modelování.



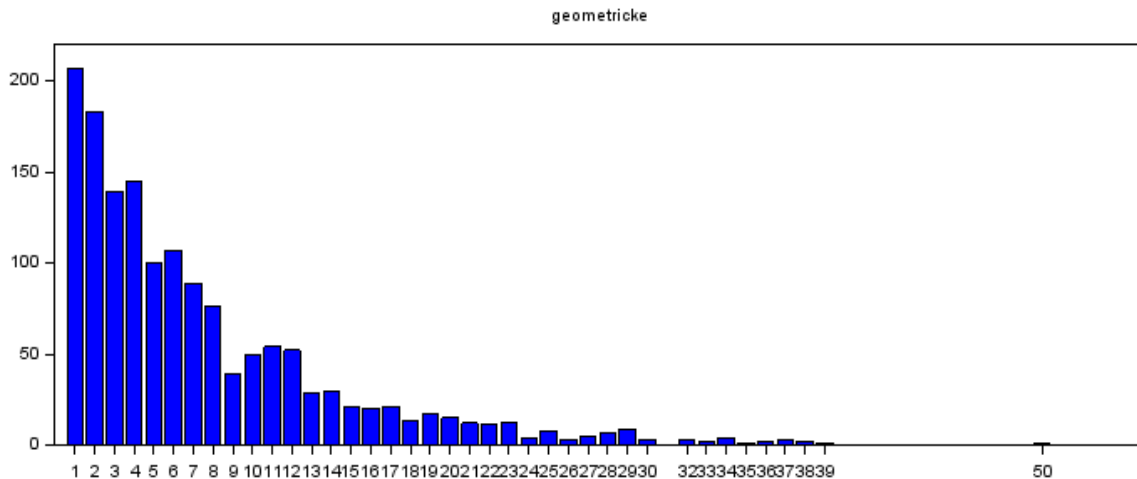
Obrázek 45 Histogram simulovaných dat, vlastní výpočty, Scilab



Obrázek 46 Histogram binomického modelu, vlastní výpočty, Scilab



Obrázek 47 Histogram Poissonova modelu, vlastní výpočty, Scilab



Obrázek 48 Histogram geometrického modelu, vlastní výpočty, Scilab

5.2 Odhad reálných diskrétních dat

V této části kapitoly byly provedeny pokusy s využitím reálných dat. Pro modelování byla použita reálná nehodová diskrétní data, získaná prostřednictvím Policie ČR.

Jak bylo již dříve řečeno (viz. Kapitola 3.3) pro modelování byly vybrány veličiny: dělení komunikace, počet zúčastněných vozidel.

Pro práci s reálnými daty je třeba v kódu programu nastavit veličinu „simul“ na 0. Cílem je najít jiné vhodné rozdělení než kategorické, které by mohlo být vhodné pro modelování. Proto veličina „cat“ v kódu programu byla nastavena na 0. Tím bylo kategorické rozdělení eliminováno z odhadu.

Výše řečene nastavení v kódu Scilab:

```
cat=0; // jestli v odhadu je kategorické rozdělení ano=1/ne=0
```

```
simul=0; // jestli chceme simulovaná=1 nebo reálná=0 data
```

```
if simul==1
```

Dále v průběhu práce s reálnými diskrétními daty bylo nutné si vybrat veličinu, která by se mohla hodit některému nekategorickému rozdělení. Pomocí zadání určitých čísel do položky „veličina“ v kódu programu, Scilab vybírá vhodnou veličinu. Zadáním do kódu čísla 18 dojde k modelování nehodových dat podle dělení komunikace. Zadáním čísla 24 do položky „veličina“ v Scilabu dojde k modelování reálných nehodových dat podle počtu zúčastněných vozidel.

Níže jsou uvedeny veličiny, které byly pro modelování k dispozici (Tab. 9).

Tabulka 9 Seznam veličin kódu, které definují kritérium modelovaných nehod

přiřazená veličina, Scilab	ref. Policie ČR	kritérium dopravní nehody
1	6	DRUH NEHODY
2	7	DRUH SRÁŽKY JEDOUCÍCH VOZIDEL
3	8	DRUH PEVNÉ PŘEKÁŽKY
4	9	CHARAKTER NEHODY
5	10	ZAVINĚNÍ NEHODY
6	11	ALKOHOL U VINÍKA NEHODY PŘÍTOMEN
7	12	HLAVNÍ PŘÍČINY NEHODY
8	13a	usmrceno NÁSLEDKY NEHODY - stav do 24
9	13b	těžce zraněno NÁSLEDKY NEHODY - stav do 24
10	13c	lehce zraněno NÁSLEDKY NEHODY - stav do 24
11	14	CELKOVÁ HMOVNÁ ŠKODA
12	15	DRUH POVRCHU VOZOVKY
13	16	STAV POVRCHU VOZOVKY V DOBĚ NEHODY
14	17	STAV KOMUNIKACE
15	18	POVĚTRNOSTNÍ PODMÍNKY V DOBĚ NEHODY
16	19	VIDITELNOST
17	20	ROZHLEDOVÉ POMĚRY
18	21	DĚLENÍ KOMUNIKACE
19	22	SITUOVÁNÍ NEHODY NA KOMUNIKACI
20	23	ŘÍZENÍ PROVOZU V DOBĚ NEHODY
21	24	MÍSTNÍ ÚPRAVA PŘEDNOSTI V JÍZDĚ
22	27	SPECIFICKÁ MÍSTA A OBJEKTY V MÍSTĚ NEHODY
23	28	SMĚROVÉ POMĚRY
24	34	POČET ZÚČASTNĚNÝCH VOZIDEL
25	35	MÍSTO DOPRAVNÍ NEHODY

V prvním pokusu bylo provedeno modelování reálných nehodových dat podle dělení komunikace. Do položky „veličina“ v kódu programu bylo zadáno číslo 18, které definuje danou veličinu. Při výběru veličiny „dělení komunikace“ nehodová data nabývají hodnot od 0 do 6.

Natažení modelovaných dat:

```
dt=csvRead('2012_I_L.csv',';');
```

```
nd=size(dt,1);
```

```
velicina=18;
```

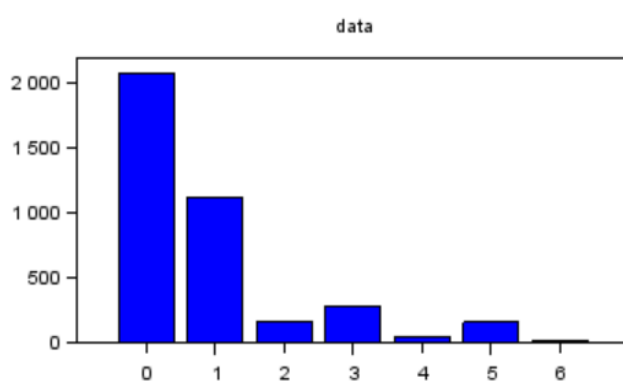
```
y=dt(:,velicina);
```

$y = y - \min(y)$;

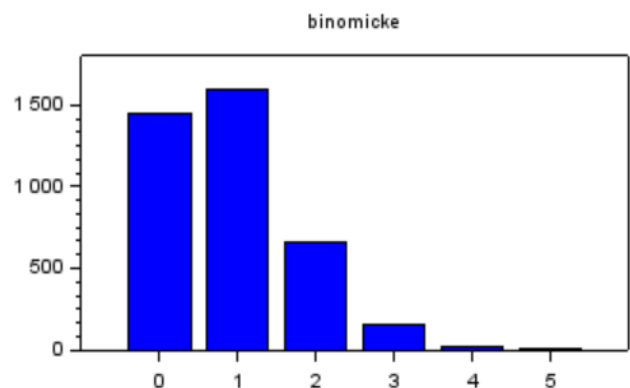
$my = \text{mean}(y)$;

Ve výsledku testu bylo vybráno geometrické rozdělení jako nejlepší možnost pro modelování zvolených dat. Na druhém místě bylo binomické rozdělení. Dané rozdělení modeluje vybraná data hůře než geometrické, nicméně také by mohlo být vhodné pro modelování. Střední hodnota modelovaných dat se rovná 0,8913713. Rovnoměrné rozdělení má odlišnou střední hodnotu, rovnou 3. Z toho důvodu rovnoměrné rozdělení bylo vyloučeno z výběru. Průměrná hodnota modelů s ostatními rozdělení se shoduje s hodnotou modelovaných dat. Ostatní rozdělení v průběhu modelování se velmi neprojevily.

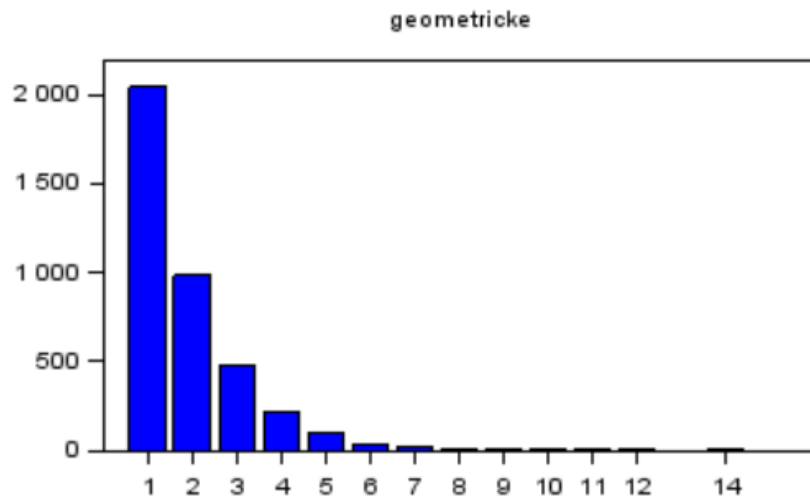
Obrázek č. 49 reprezentuje histogram modelovaných reálných dat. Je vidět, že reálná data na rozdíl od simulovaných dat nemají „spojité“ uspořádání, tedy lze je hodnotit jako náhodné veličiny s multimodálním rozdělením. Na obrázku č. 51 je představen histogram modelu s geometrickým rozdělením. Je vidět, že hodnoty geometrického modelu se přizpůsobují hodnotám modelovaných dat. Binomické rozdělení, které také vyhovuje danému modelu, je znázorněno na obrázku č. 50. V porovnání s geometrickým modelem hodnoty binomického modelu méně napodobují hodnoty reálných dat. Na obrázcích č. 52 a 53 jsou představeny modely Poissonova a negativně binomického rozdělení. Z histogramů je vidět, že výsledné hodnoty modelu neodpovídají tvaru histogramu reálných modelovaných dat. Obrázky č. 54 a 55 reprezentují histogramy kategoričného a rovnoměrného rozdělení, jejich výsledné hodnoty neodpovídají tvaru reálných modelovaných dat.



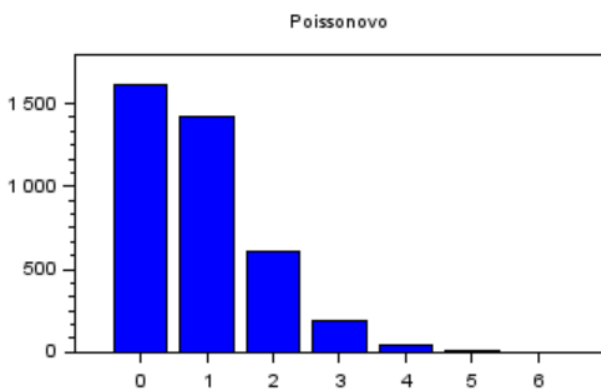
Obrázek 49 Histogram reálných dat, vlastní výpočty, zdroj: Policie ČR



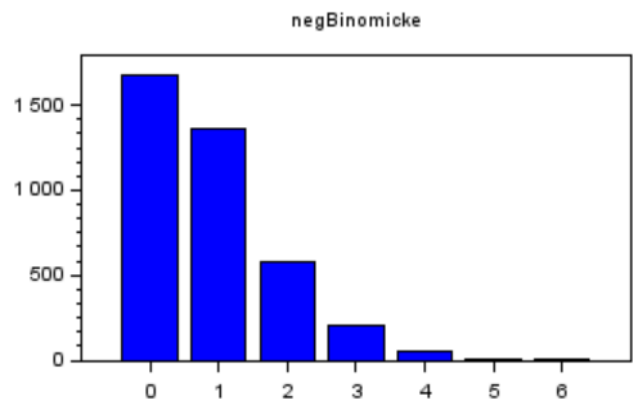
Obrázek 50 Histogram binomického modelu, vlastní výpočty, zdroj: Policie ČR



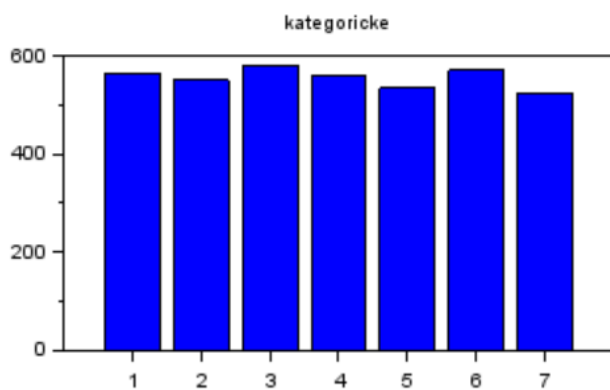
Obrázek 51 Histogram geometrického modelu, vlastní výpočty, zdroj: Policie ČR



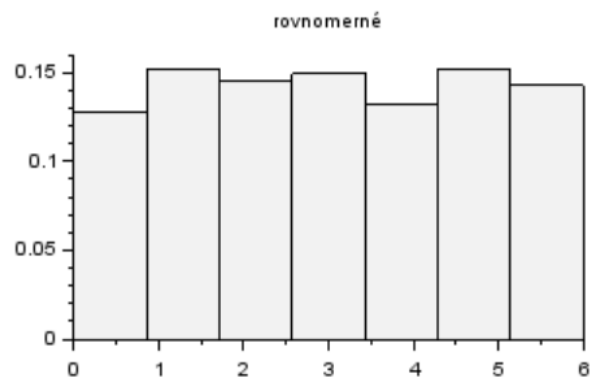
Obrázek 52 Histogram Poissonova rozdělení, vlastní výpočty, zdroj: Policie ČR



Obrázek 53 Histogram negativně binomického rozdělení, vlastní výpočty, zdroj: Policie ČR



Obrázek 54 Histogram modelu s kategoričným rozdělením, vlastní výpočty, zdroj: Policie ČR



Obrázek 55 Histogram modelu s rovnoměrným rozdělením, vlastní výpočty, zdroj: Policie ČR

V dalším pokusu byla modelována nehodová data podle počtu zúčastněných vozidel. Do položky „veličina“ v kódu programu bylo zadáno číslo 24, které definuje veličinu nehodových dat. Tady hodnoty nehodových dat kolísají od 1 do 7.

Natažení modelovaných dat:

```
dt=csvRead('2012_I_L.csv',';');
```

```
nd=size(dt,1);
```

```
velicina=24;
```

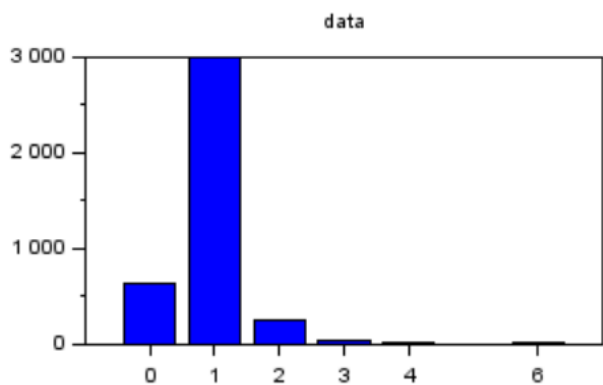
```
y=dt(:,velicina);
```

```
y=y-min(y);
```

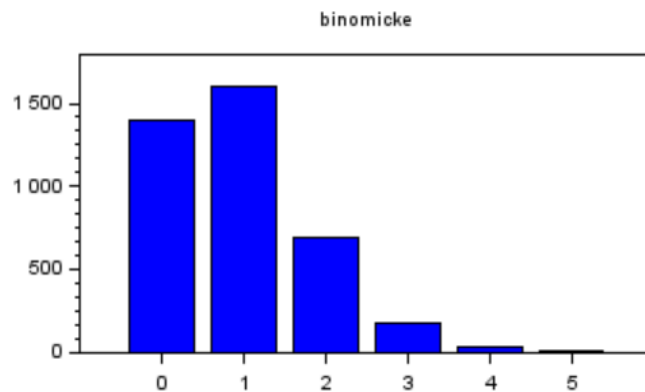
```
my=mean(y);
```

Výsledku testu ukázaly, že binomické rozdělení nejlépe vyhovuje danému modelování reálných dat. Binomické rozdělení je preferující pro dané modelování, nicméně geometrické rozdělení by mohlo nahradit binomický model. Výsledky geometrického rozdělení jsou horší, ale toto rozdělení také splňuje požadavky daného modelování. Alternativní a kategorické rozdělení byly vyloučeny záměrně v průběhu kódování. Rovnoměrné rozdělení bylo vyloučeno z důvodu odlišnosti jeho střední hodnoty od hodnot ostatních modelů. Střední hodnota většiny modelu se rovná 0,920904, zatímco střední hodnota rovnoměrného modelu je 3. Ostatní vyzkoušené rozdělení neprokázaly dobré výsledky.

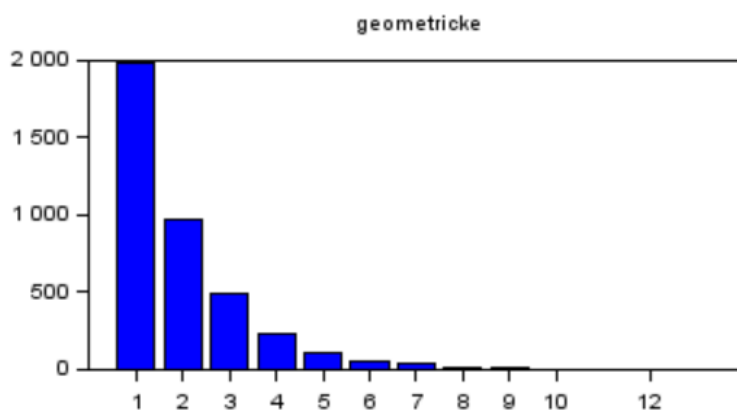
Vybraná reálná diskrétní data v daném modelu jsou představena na obrázku č. 56. Z histogramu dat z binomického rozdělení je vidět, že jeho výsledné hodnoty imitují podobu předchozího histogramu reálných dat (Obr. 57). Histogram geometrického rozdělení je představen na obrázku č. 58. V porovnání s binomickým modelem se histogram modelu s geometrickým rozdělením méně podobá histogramu reálných dat. Na obrázcích č. 59 a 60 znázorněny histogramy Poissonova a negativně binomického rozdělení, jejich hodnoty neshodují s hodnotami reálných dat. Histogramy kategorického a rovnoměrného rozdělení mají úplně jiný tvar než histogramy předchozích modelů (Obr. 61,62). Hodnoty histogramů nesouvisí se simulovanými daty.



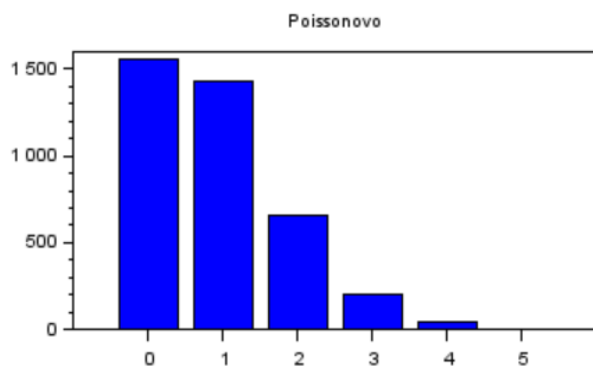
Obrázek 56 Histogram reálných dat, zdroj: Policie ČR, vlastní výpočty



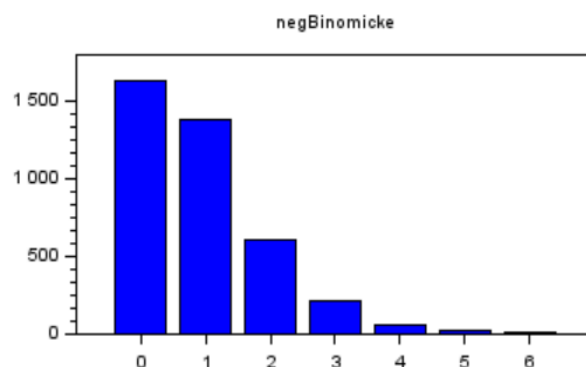
Obrázek 57 Histogram binomického modelu, zdroj: Policie ČR, vlastní výpočty



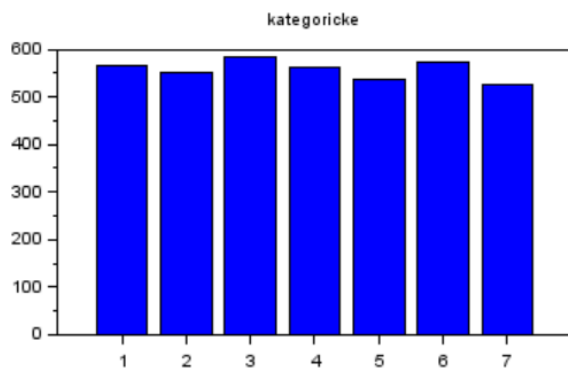
Obrázek 58 Histogram geometrického modelu, vlastní výpočty, zdroj: Policie ČR



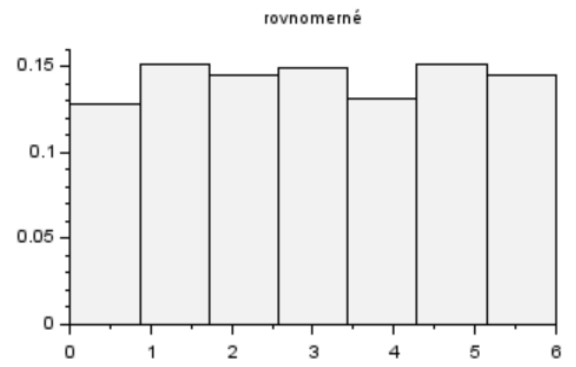
Obrázek 59 Histogram Poissonova modelu, vlastní výpočty, zdroj: Policie ČR



Obrázek 60 Histogram negativně binomického modelu, vlastní výpočty, zdroj: Policie ČR



Obrázek 61 Histogram modelu s kategoričným rozdělením, vlastní výpočty, zdroj: Policie ČR



Obrázek 62 Histogram modelu s rovnoměrným rozdělením, vlastní výpočty, zdroj: Policie ČR

Závěr, zhodnocení

V první části práce byly popsány základní pojmy a historický vývoj automobilismu s ohledem na dopravní nehody. Dopravní nehody byly popsány a rozděleny zejména s ohledem na příčiny jejich vzniku.

V další kapitole práce byla provedena analýza dopravní nehodovosti v EU a ČR. Pro porovnání nehodovosti v EU bylo vybráno několik států s odlišným ekonomickým výkonem: Česká republika, Německo, Slovensko, Polsko, Nizozemí, Spojené Království, Španělsko, Finsko, Portugalsko, Maďarsko, Lucembursko, Francie, Irsko, Rakousko a Švédsko. Z webových stránek Ministerstva dopravy byly zjištěny údaje o dopravních nehodách za rok 2015. Cílem bylo ověřit předpoklad, že s v zemích s vyšším HDP a rozvinutější dálniční sítí dochází k menšímu počtu nehod a k nehodám s menšími dopady na lidské zdraví.

Pro tato zkoumání byla využita lineární regrese, která zkoumala závislost HDP v paritě kupní síly na podílu nehodovosti na počet obyvatel. Data týkající se HDP s PPS a počtu obyvatel ve vybraných zemích byla získána za rok 2015 z webových stránek Eurostatu. Paradoxně z těchto dat (byť slabě) vyplynulo, že s rostoucím HDP roste nehodovost. Výsledek je velice překvapivý, ale dá se odůvodnit tím, že v zemích s vyšší úrovní HDP je větší počet registrovaných vozidel, což by mohlo ovlivnit zvýšení počtu dopravních nehod.

Pro hledání závislosti nehodovosti na dopravní infrastruktuře byly získány údaje o délce silnic a dálnic za rok 2015 na základě dat z Eurostatu. Analýza byla provedena na základě porovnání podílu dálnic z celkové délky silnic a podílu počtu dopravních nehod na celkovém počtu obyvatel. Výsledky opět ukázaly překvapivý závěr, že s rostoucím podílem dálnic roste počet dopravních nehod. Opět jde o paradoxní závěr, který lze odůvodnit pouze tím, že na větších úsecích je větší intenzita dopravy, což zvyšuje riziko vzniku dopravní nehody. Stejným způsobem byla provedena analýza závislosti dopravní infrastruktury na nehodovosti, ke které došlo k usmrcení. Závislost se projevila slabší, než při porovnání dopravní infrastruktury na celkovém počtu nehod. Výsledek byl stejný jako v předchozí analýze.

V další části kapitoly bylo provedeno porovnání nehodovosti mezi kraji ČR. Údaje o nehodách byly zjištěny z webových stránek Ministerstva ČR za období 2008-2012. Největší počet dopravních nehod byl ve vybraném období zaznamenán v hlavním městě Praha. Hlavním důvodem je vysoká intenzita dopravy, zapříčiněná především povahou hlavního města, které je logicky centrálním bodem země, kam směřují lidé za celou řadou aktivit. Druhou příčinou je vysoká úroveň turismu v porovnání s ostatními kraji. Nejmenší počet nehod měl Karlovarský kraj.

V práci byl také sledován vývoj nehodovosti v celé ČR za stejné období. Největší počet dopravních nehod v sledovaném období byl zaznamenán v roce 2008, v dalším průběhu počet nehod v silničním provozu po roce 2008 rapidně klesl. Tento výrazný pokles byl především ovlivněn změnou hranice hmotné škody, ke které došlo v roce 2009. Od roku 2012 do 2017 dopravní nehodovost měla poměrně rostoucí trend. To mohlo být způsobeno rozvojem dopravní infrastruktury a dlouhodobým nárůstem celkového počtu vozidel na území ČR. Dále byla zkoumána závislost nehodovosti na dopravních investicích. Pro hledání této závislosti byly vybrány hodnoty reprezentující podíl dopravních nehod na počet obyvatel za období 2008-2017 a podíl dopravních investic na HDP. Údaje o dopravních investicích byly získány z webových stránek SFDI, údaje o HDP byly poskytnuté z ČSÚ. Zkoumání závislosti nehodovosti na dopravních investicích je takový, že s růstem investic do dopravní infrastruktury roste dopravní nehodovost a naopak. Takový výsledek není v souladu s předpokladem. Možným důvodem jsou vyšší intenzity dopravy na nových a kvalitnějších komunikacích. Tento vyšší provoz může vést i k vyššímu počtu dopravních nehod a naopak.

V další části práce jsem se zabývala problematikou nehodovosti v hl. městě Praze. Rovněž bylo provedeno celkové zhodnocení situace dopravní nehodovosti v městě. Kromě toho byl stručně popsán Plán udržitelné mobility Prahy a okolí. Rovněž byly popsány hlavní cíle a opatření plánu, zaměřené na zvýšení dopravní bezpečnosti a snížení nehodovosti. V této kapitole byly popsány i vstupní data nehodovosti v Praze, potřebné pro modelování v praktické části. Poskytovatelem těchto dat je Policie ČR.

Pro sledování vztahů mezi veličinami je zapotřebí model. Proto byl v následující kapitole proveden popis vybraných matematických modelů. Zde byla zmíněna teorie modelů popisujících diskrétní data a bylo popsáno sedm statistických rozdělení (kategorické, alternativní, binomické, negativně binomické, geometrické, Poissonovo, rovnoměrné) která byla potřebná pro modelování.

Konec práce je zaměřen na praktickou část, která se věnovala modelování a odhadu diskrétních dat. Ve skutečnosti tato diskrétní data často pochází z různých dotazníků. A tak je tomu také v záznamech o dopravních nehodách, pořízených policií. Pro modelování diskrétních dat, které je základem pro statistickou analýzu, se standardně používá model s kategorickým rozdělením, ale daný model vyžaduje příliš velkou pravděpodobnostní tabulku, kterou často není možno implementovat v paměti počítače. Navíc, pro odhad parametrů modelu je potřeba mít obrovské množství dat. Proto cílem této práce bylo pomoci počítačového programu Scilab zkusit najít vhodné statistické rozdělení, které by mohlo nahradit kategorický model. Na začátku byly provedeny pokusy s modelováním využívající simulovaná data. V prvních pěti testech byla zkoumána data s počtem hodnot 500. Testy se lišily tvarem

pravděpodobnostní funkce. Na základě výsledku většiny testů bylo vybráno binomické rozdělení. Ve většině případů se také dobře projevíly geometrické a Poissonovo rozdělení, ale dané modely hůře odpovídaly modelovaným datům. Dále byly provedeny 3 testy, kde počet hodnot zkoumaných simulovaných dat byl nastaven na hodnotu 1500. Výsledky testu ukázaly, že binomické rozdělení je nejvhodnější pro modelování vybraných simulovaných dat. Dále byly provedeny pokusy s využitím reálných dat. Reálná nehodová diskrétní data byla získána prostřednictvím Policie ČR. Pro modelování byly vybrány veličiny: dělení komunikace, počet zúčastněných vozidel. Ve výsledku pro modelování veličin podle dělení komunikace bylo vybráno geometrické rozdělení jako nejlepší možnost pro modelování zvolených dat. Na druhém místě bylo binomické rozdělení. V dalším pokusu byla modelována nehodová data podle počtu zúčastněných vozidel. Výsledky testu ukázaly, že binomické rozdělení nejlépe vyhovuje danému modelování reálných dat. Binomické rozdělení je preferující pro dané modelování, nicméně geometrické rozdělení by mohlo nahradit binomický model. Střední hodnota rovnoměrného rozdělení u všech pokusů vždy byla odlišná od střední hodnoty simulovaných dat. Proto dané rozdělení bylo považované za nevhodné. Alternativní model má jen dvě hodnoty 0 a 1 a testované veličiny jsou vícehodnotové. Proto při výběru vhodného modelu alternativní rozdělení bylo vynecháno. Z výsledků plyne, že model binomický a v některých případech i geometrický mohou nahradit model s kategoriálním rozdělením.

Použité zdroje

- [1] Centrum služeb pro silniční dopravu. Cspsd.cz [Online] Dostupné z: https://www.cspsd.cz/storage/files/historie_nehody.pdf
- [2] Newsland. [Online] Dostupné z: <http://newsland.com/community/1793/content/pervaia-v-istorii-zhertva-dtp/6157637>
- [3] Poznámky z přednášek Technologie silniční dopravy
- [4] Policie České republiky. [Online] Dostupné z: <https://www.policie.cz>
- [5] Metodika aplikace nehody v mapě [Online] Dostupné z: <http://pcr.jdvm.cz>
- [6] Centrum dopravního výzkumu [Online] Dostupné z: <https://www.cdv.cz/publikace/>
- [7] Ministerstvo dopravy ČR [Online] Dostupné z: https://www.sydos.cz/cs/rocenka_pdf/Rocenka_dopravy_2017.pdf
- [8] PAVELKA, Tomáš. *Makroekonomie: základní kurz*. 2. vyd. Praha: Melandrium, 2007. ISBN 978-80-86175-52-2.
- [9] Eurostat: [Online] Dostupné z: <https://ec.europa.eu/eurostat/en/data/database>
- [10] Eurostat Annual Accident report. [Online] Dostupné z https://ec.europa.eu/transport/road_safety/sites/roadsafety/files/pdf/statistics/dacota/asr2018.pdf
- [11] Český statistický úřad. [Online] Dostupné z: <https://www.czso.cz/documents/10180/20534694/32025414a02.docx/fb7cba2c-a8e7-4d8f-bd3c-b3aab5380ac7?version=1.0>.
- [12] Státní fond dopravní infrastruktury. [Online] Dostupné z: <https://www.sfdi.cz/rozpocet/>
- [13] [Online] Dostupné z: https://www.idnes.cz/auto/historie/historie-automobilismu-v-datech.A060127_172200_auto_ojetiny_fdv
- [14] Web skriptům Ivan Nagy " Statistika" [Online] Dostupné z: <https://www.fd.cvut.cz/personal/nagyivan/Statistika/Statistika.pdf>
- [15] Daniel Zelterman Models for Discrete Data, vyd. Oup Oxford, 2006. ISBN: 9780198567011, 0198567014
- [16] Technická správa komunikací. [Online] Dostupné z: <http://www.tsk-praha.cz/wps/portal/root/nabidka-sluzeb/rocenky>
- [17] Polad' Prahu, plán udržitelné mobility [Online] Dostupné z: <https://www.poladprahu.cz/cs/dokumenty-p-plus>

Seznam obrázků

Obrázek 1: Silniční nehoda XIV století [2]	11
Obrázek 2: Podíl jednotlivých činitelů na vzniku dopravních nehod za období 2011-2015 [6]	15
Obrázek 3: Intenzity automobilové dopravy v Praze rok 2018 [16]	27
Obrázek 4: Problémová mapa automobilové dopravy [17]	30
Obrázek 6: Histogram posloupnosti simulovaných dat, vlastní výpočty, Scilab.....	46
Obrázek 5: Histogram binomického modelu, vlastní výpočty, Scilab	46
Obrázek 7: Histogram geometrického modelu, vlastní výpočty, Scilab	46
Obrázek 10: Histogram kategorického modelu, vlastní výpočty, Scilab	47
Obrázek 8: Histogram negativně binomického modelu, vlastní výpočty, Scilab	47
Obrázek 9: Histogram negativně binomického modelu, vlastní výpočty, Scilab	47
Obrázek 11: Histogram modelu s rovnoměrným rozdělením u, vlastní výpočty, Scilab	47
Obrázek 12: Histogram geometrického modelu, vlastní výpočty, Scilab	48
Obrázek 13: Histogram klesajících simulovaných dat, vlastní výpočty, Scilab	48
Obrázek 14: Histogram binomického modelu	48
Obrázek 17: Histogram Poissonova modelu, vlastní výpočty, Scilab	49
Obrázek 18: Histogram geometrického modelu, vlastní výpočty, Scilab	49
Obrázek 15: Histogram binomického modelu, vlastní výpočty, Scilab	49
Obrázek 16: Histogram simulovaných dat, vlastní výpočty, Scilab	49
Obrázek 19: Histogram binomického modelu, vlastní výpočty, Scilab	50
Obrázek 22: Histogram geometrického modelu, vlastní výpočty, Scilab	50
Obrázek 21: Histogram Poissonova modelu, vlastní výpočty, Scilab	50
Obrázek 20: Histogram simulovaných dat, vlastní výpočty, Scilab	50
Obrázek 23: Histogram geometrického modelu, vlastní výpočty, Scilab	51
Obrázek 25: Histogram simulovaných dat, vlastní výpočty, Scilab	51
Obrázek 26: Histogram binomického modelu, vlastní výpočty, Scilab	51
Obrázek 24: Histogram Poissonova modelu, vlastní výpočty, Scilab	51
Obrázek 27: Histogram binomického modelu, vlastní výpočty, Scilab	52
Obrázek 29: Histogram Poissonova rozdělení, vlastní výpočty, Scilab	52
Obrázek 28: Histogram simulovaných dat, vlastní výpočty, Scilab	52
Obrázek 30: Histogram geometrického modelu, vlastní výpočty, Scilab	53
Obrázek 31: Histogram binomického modelu, vlastní výpočty, Scilab	54
Obrázek 32: Histogram simulovaných dat, vlastní výpočty, Scilab	54
Obrázek 33: Histogram modelu s kategorickým rozdělením, vlastní výpočty, Scilab	54
Obrázek 34: Histogram modelu s rovnoměrným rozdělením, vlastní výpočty, Scilab	54
Obrázek 36: Histogram geometrického modelu, vlastní výpočty, Scilab	55
Obrázek 35: Histogram Poissonova modelu, vlastní výpočty, Scilab	55

Obrázek 38: Histogram binomického modelu, vlastní výpočty, Scilab	56
Obrázek 37: Histogram simulovaných dat, vlastní výpočty, Scilab	56
Obrázek 41: Histogram reálných dat, vlastní výpočty, zdroj: Policie ČR.....	58
Obrázek 42: Histogram binomického modelu, vlastní výpočty, zdroj: Policie ČR	58
Obrázek 43: Histogram geometrického modelu, vlastní výpočty, zdroj: Policie ČR	58
Obrázek 44: Histogram Poissonova rozdělení, vlastní výpočty, zdroj: Policie ČR.....	59
Obrázek 45: Histogram negativně binomického rozdělení, vlastní výpočty, zdroj: Policie ČR	59
Obrázek 47: Histogram modelu s kategoričným rozdělením, vlastní výpočty, zdroj: Policie ČR	59
Obrázek 46: Histogram modelu s rovnoměrným rozdělením, vlastní výpočty, zdroj: Policie ČR	59
Obrázek 49: Histogram reálných dat, zdroj: Policie ČR, vlastní výpočty	60
Obrázek 48: Histogram binomického modelu, zdroj: Policie ČR, vlastní výpočty	60
Obrázek 52: Histogram geometrického modelu, vlastní výpočty, zdroj: Policie ČR	61
Obrázek 51: Histogram Poissonova modelu, vlastní výpočty, zdroj: Policie ČR.....	61
Obrázek 50: Histogram negativně binomického modelu, vlastní výpočty, zdroj: Policie ČR..	61

Seznam tabulek

Tabulka 1: Dopravní nehodovost a obyvatelstvo v sledovaných zemích EU v roce 2015 [9]	15
Tabulka 2: nehodovost, počet obyvatelů, HDP v PPS v roce 2015[zdroj: Eurostat [9].....	16
Tabulka 3: Dopravní infrastruktura ve vybraných zemích EU v roce 2015 [9].....	17
Tabulka 4: Podíl nehod s usmrcením na počet obyvatelů a dopravní infrastruktura v roce 2015 [zdroje: Ministerstvo dopravy, Eurostat, vlastní výpočty]	18
Tabulka 5: Počet nehod v silničním provozu, při kterých došlo ke zranění nebo usmrcení osob za období 2008-2017 [7]	20
Tabulka 6: Investiční výdaje do dopravní infrastruktury v sledovaném období [12]	23
Tabulka 7: Hrubý domácí produkt podle krajů za období 2008-2017 [11].....	24
Tabulka 8: Vývoj dopravní nehodovosti v Praze [16].....	28