

**Univerzita Hradec Králové
Přírodovědecká fakulta
Katedra fyziky**

Studie dosahu interakčních sil v dopravních systémech

Diplomová práce

Autor: Michaela Krbálková
Studijní program: N1701 Fyzika
Studijní obor: Fyzikální měření a modelování
Vedoucí práce: doc. Ing. Tomáš Hobza, Ph.D.

Prohlášení

Prohlašuji, že jsem diplomovou práci vypracovala (pod vedením vedoucího práce) samostatně a s použitím uvedené literatury.

V Hradci Králové dne

Michaela Krbálková

Poděkování

Děkuji doc. Ing. Tomáši Hobzovi, Ph.D. za vedení mé diplomové práce a za přínosné konzultace vedené v přátelské atmosféře. Dále děkuji doc. Mgr. Milanu Krbálkovi, Ph.D. za podnětné rady a příspěvky k tomuto zajímavému tématu.

Anotace

KRBÁLKOVÁ, Michaela. *Studie dosahu interakčních sil v dopravních systémech*. Hradec Králové, 2018. Diplomová práce. Univerzita Hradec Králové, Přírodovědecká fakulta, Katedra fyziky. 69 s.

Cílem práce je matematickými prostředky stanovit dosah interakčních sil mezi vozidly pohybujícími se na reálných komunikacích. Užití metody budou založeny na tzv. poruchovém principu, kdy porušení předpokladů pro odvození rozdělení součtu náhodných veličin užitím konvoluce nenarušuje příslušnost asociovaných hustot pravděpodobnostido zvolené třídy distribucí. Analýza statistických odchylek mezi analytickými a empirickými distribucemi bude pak vhodným nástrojem pro detekci počtu za sebou jedoucích vozidel, která spolu vzájemně interagují. Studie bude navíc podpořena analýzou korelačních koeficientů vyčíslených pro vhodné mikroskopické dopravní veličiny.

Klíčová slova

fyzika dopravy; headway distribuce; rozdělení GIG; poruchová konvoluční funkce

Annotation

KRBÁLKOVÁ, Michaela. *Study of an interaction range in vehicular ensembles*. Hradec Králové, 2018. Diploma Thesis. University of Hradec Králové, Faculty of Science, Department of Physics. 69 pp.

The aim of the project is a mathematical detection of a range in interaction forces among vehicles moving on expressways. Applied methods will be based on the so-called perturbation principle, when violation of conditions for applying the convolution (when deriving a distribution for a sum of two random variables) does not cause any robust changes in a resulting distribution family. Quantitative analysis of statistical discrepancies between analytical and empirical distributions is a suitable instrument for detection of how many succeeding vehicles effectively interact in traffic systems. Moreover, such a study will be supported by an analysis of correlation coefficients enumerated for associated microscopic traffic-quantities.

Keywords

physics of traffic; headway distribution; distribution family GIG; perturbative convolution function

Obsah

1	Konvoluce, hustota a integrální transformace	11
1.1	Lebesgueův integrál	11
1.2	Konvoluce, hustota a jejich vlastnosti	13
1.3	Laplaceova a Fourierova transformace	15
1.4	Konvoluce hustot	20
2	Základní poznatky z teorie pravděpodobnosti	23
2.1	Pravděpodobnost a náhodná veličina	23
2.2	Absolutně spojitá rozdělení	26
2.3	Součet náhodných veličin	30
2.4	Metody odhadu parametrů	31
3	Základní pojmy teorie dopravního proudění	33
3.1	Mikroskopické znaky dopravy	33
3.2	Makroskopický popis dopravy	34
3.3	Fundamentální diagram	35
3.4	Detekce dopravních dat	35
4	Zobecněné inverzní Gaussovo rozdělení GIG a odhadování jeho parametrů	37
4.1	Dvouparametrická třída GIG	37
4.2	Metodika odhadování parametrů distribuce GIG	38
4.3	Testování generátoru pseudonáhodných čísel z dvouparametrické třídy GIG	38
5	Vyhodnocení empirických dat – model pro rozdělení světlostí	41
5.1	Charakteristika zkoumaného datového souboru	41
5.2	Metodika odhadování parametrů a statistické testování	42
6	Poruchová konvoluční funkce – nástroj pro detekci statistické závislosti v dopravních datech	51
6.1	Zavedení pojmu multisvětlost	51
6.2	Statistické odhady pro rozdělení multisvětlostí	51
6.3	Koncept poruchové konvoluční funkce a analýza jejího průběhu	53
6.4	Empirické chování normy poruchové funkce	55
	Přílohy	63

Úvod

Pro studium problematiky dopravního toku jsou zásadním pojmem rozestupy (světlosti) vozidel, viz [1, 2, 3]. Statistická rozdělení těchto fundamentálních dopravních veličin popisují rozložení vozidel na dálnicích [4, 5] či křižovatkách [6, 7]. V matematických predikcích kapacity dálnice či křižovatky představuje rozdělení světlostí klíčový faktor. Disciplína modelování světlostí vozidel se v současné době velice intenzivně rozvíjí, díky čemuž byla odhalena řada pozoruhodných rysů dopravní mikrostruktury. Historický přehled a dosažené úspěchy jsou velmi přehledným způsobem shrnuty v práci [1]. Světlosti vozidel jsou však také velice úzce spojeny se základními makroskopickými dopravními veličinami, a to s hustotou a dopravním tokem.

Pojmem modelování světlostí rozumíme vědeckou disciplínu zabývající se analytickými odhady empirických distribucí světlostí. Tato disciplína prošla za posledních 80 let intenzivním vývojem. První pokus o modelování statistiky světlostí pro reálný dopravní tok byl proveden ve 30-tých letech minulého století, viz [8], kde bylo využito normální a exponenciální rozdělení. Na zmíněnou práci navázalo mnoho dalších autorů, [9, 10, 11, 12, 13, 14, 15, 4, 5, 16, 17, 7]. V počátcích se jednalo pouze o ad hoc modely, ale později se vědci pokoušeli modifikováním známých fyzikálních či matematických modelů predikovat statistické vlastnosti dopravní mikrostruktury co nejhodněji. Jako model se při zmíněných predikcích využívaly jednoparametrické nebo multiparametrické třídy známých distribucí. Typicky se jednalo o distribuční třídy exponenciálních, posunutých exponenciálních, Erlangových, Gamma, log-normálních či GIG rozdělení. Příslušné metody odhadování parametrů však měly a mají několik významných nedostatků. Především žádná z těchto metod nikdy nebyla verifikována standardním statistickým testem, což je vzhledem k historii této disciplíny vlastně logické, protože prvotní odhady byly extrémně nepřesné. Závažnějším nedostatkem uvedených odhadů je ale skutečnost, že většina navrhaných distribucí světlostí nesplňuje základní matematická kritéria, která zkoumaný systém musí podle teorie jednodimenzionálních částicových systému naplňovat.

Aktuálně se jako vhodný kandidát pro distribuce světlostí jeví dvouparametrická třída GIG distribucí (zobecněná inverzní Gaussova distribuce), viz [18, 19], z čehož budeme vycházet i my. Naším cílem pro první část práce je najít vhodný model pro odhad empirických distribucí časové světlosti, ověřit ho statistickým testem a analyzovat změny odhadovaných parametrů vyvolané změnami makroskopických dopravních veličin. Tento model budeme dále ve druhé části práce používat při odhadech statistiky tzv. multisvětlostí (tj. vzdáleností několika sousedních vozidel). Odhady multisvětlostí založené na rozdělení GIG pak využijeme k zodpovězení základní otázky této diplomové práce. Záměrem práce je totiž rozhodnout, zda může být nástroj tzv. poruchové konvoluční funkce využit při detekci dosahu interakčních sil v dopravních systémech. Metoda je založena na základním matematickém principu, podle kterého platí, že jsou-li dvě náhodné veličiny nezávislé, pak lze hustotu pravděpodobnosti jejich součtu vypočítat jako konvoluci hustot příslušných těchto veličinám. Rozdíl mezi hustotou pravděpodobnosti součtu a uvedenou konvolucí v sobě tedy odráží míru závislosti obou veličin. Tento princip zamýšlíme (v dlouhodobějším horizontu)

využít při detekci počtu tzv. aktivně sledovaných vozidel. Jinými slovy, naší snahou je matematickými prostředky odhadnout, kolik předcházejících vozidel zohledňuje řidič při řízení svého vozidla a jak se tento počet (nazývaný v této práci *interakčním dosahem* - viz také [20]) mění s makroskopickými parametry dopravního proudění.

Práce je uspořádána takto: v úvodní kapitole sestavíme přehled důležitých matematických pojmů (hustota, konvoluce, integrální transformace), z nichž budeme vycházet v následujících částech práce. V kapitole druhé tento souhrn doplníme o poznatky a pojmy z teorie pravděpodobnosti (náhodná veličina, hustota pravděpodobnosti, distribuční funkce) a budeme se věnovat vlastnostem absolutně spojitých rozdělení, zejména dvou- a tříparametrického rozdělení GIG. V závěru druhé kapitoly krátce pojednáme o statistických metodách odhadu parametrů distribuce. Třetí kapitola nám představí mikroskopické a makroskopické veličiny teorie dopravního proudění, jakými jsou světlost, dopravní tok či hustota toku. Stručně bude okomentována metoda detekce dopravních dat, jež budeme v této diplomové práci analyzovat, a charakteristika těchto dat. Sestavením modelu dvouparametrické GIG distribuce se zabývá kapitola čtvrtá. Následně pomocí tohoto modelu prověříme generátor pseudonáhodných čísel z distribuce GIG. V páté kapitole přistoupíme ke zpracování reálných souborů světlostí naměřených na Pražském okruhu. Rozebereme členění dat, metodiku odhadu parametrů a vývoj parametrů odhadované GIG distribuce v závislosti na změnách hodnot dopravních makroveličin. Model pak prověříme standardními statistickými testy. V předposlední kapitole zavedeme pojem multisvětlosti a budeme testovat hypotézu, že multisvětlosti libovolného řádu jsou distribuovány podle zobecněného inverzního Gaussova rozdělení se třemi parametry. Opět budeme také analyzovat závislosti odhadovaných parametrů na dopravní hustotě. V rámci kapitoly budeme také definovat tzv. poruchovou konvoluční funkci a zkoumat její průběh v závislosti na jejím řádu a na hustotním pásmu, z něhož je funkce extrahována. Dále představíme dvě funkcionální normy (kolmogorovskou a L_1 -normu), s jejichž pomocí se budeme snažit prokázat, že odchylky poruchové funkce od nulové funkce klesají s jejím řádem. V závěrečné kapitole budeme diskutovat dosažené výsledky a výhled pro navazující výzkum.

Kapitola 1

Konvoluce, hustota a integrální transformace

Úvodní kapitola nás seznámí se dvěma zcela zásadními pojmy, jimiž se budeme v této práci zabývat - konvolucí a hustotou. Zadefinujeme zde Laplaceovu a Fourierovu transformaci a uvedeme některé důležité vlastnosti a vztahy těchto integrálních transformací. Ty aplikujeme v následující části kapitoly, kde se zaměříme na odvození konvoluce dvou hustot pro několik významných rozdělení.

1.1 Lebesgueův integrál

Začátek první kapitoly věnujeme stručnému připomenutí významných pojmů a vět z teorie Lebesgueova integrálu. Budeme se na ně často odkazovat v následujícím textu. Jejich znění včetně důkazů a doplňujících pojmů lze nalézt ve skriptu [21].

Mějme měřitelnou množinu $A \subset \mathbf{R}$. Funkce $f(x)$ je lebesgueovsky integrabilní na množině A , pokud $(\mathcal{L}) \int_A f(x) dx$ existuje. Označuje se zápisem $f(x) \in \mathcal{L}^*(A)$. Pokud $(\mathcal{L}) \int_A f(x) dx$ existuje a je konečný, značí se $f(x) \in \mathcal{L}(A)$. Pokud $(\mathcal{L}) \int_A |f(x)| dx$ existuje a je konečný, označuje se $f(x) \in \mathcal{L}_1(A)$. Všechny integrály použité v této práci mají význam Lebesgueova integrálu, dále tedy nebude znak (\mathcal{L}) před integrály uváděn. Jsou-li dvě funkce ekvivalentní (tj. liší se na množině míry 0), pak mají totožný Lebesgueův integrál ve smyslu existence, neexistence a hodnoty.

Věta

Nechť $f(x)$ je měřitelná funkce na množině A . Pak platí $f(x) \in \mathcal{L}(A) \Leftrightarrow f(x) \in \mathcal{L}_1(A)$. Pro měřitelné funkce tedy platí $\mathcal{L}(A) = \mathcal{L}_1(A)$.

Definice

Nechť $v(x)$ je výroková formule, jež má smysl pro každé $x \in E$. Nechť $A \subset E$. Řekneme, že $v(x)$ platí *skoro všude* v množině A , resp. pro *skoro všechna* $x \in A$, jestliže míra $\mu(X)$ množiny všech $x \in A$, kde formule $v(x)$ neplatí, je rovna 0.

Věta - srovnávací kritérium

Nechť $f(x)$ je měřitelná funkce na množině A , $g(x) \in \mathcal{L}(A)$ je lebesgueovsky integrabilní funkce a $|f(x)| \leq g(x)$ skoro všude v A . Pak také $f(x) \in \mathcal{L}(A)$.

Věta - Lebesgueova o integrabilní majorantě

Nechť funkce $f_n(x)$ jsou měřitelné pro $\forall n \in \mathbf{N}$ a $\lim_{n \rightarrow +\infty} f_n(x) = f(x)$ pro skoro všechna $x \in A$. Nechť existuje $g(x) \in \mathcal{L}(A)$ tak, že $|f_n(x)| \leq g(x)$ pro $\forall n \in \mathbf{N}$ skoro všude v A . Pak platí tvrzení $f(x) \in \mathcal{L}(A)$ a rovnost

$$\lim_{n \rightarrow +\infty} \int_A f_n(x) dx = \int_A f(x) dx.$$

Věta - o limitě integrálu s parametrem

Nechť je dána množina $P \subset \mathbf{R}$ všech přípustných parametrů a její hromadný bod $\beta \in \text{der}(P)$, kde $\text{der}(P)$ označuje derivaci množiny P (tj. množinu všech hromadných bodů množiny P). Nechť funkce $f(x, \alpha)$ definovaná na $A \times P$, kde $A \subset \mathbf{R}$ je libovolná měřitelná množina, splňuje následující předpoklady:

- Pro skoro všechna $x \in A$ existuje

$$\lim_{\alpha \rightarrow \beta, \alpha \in P} f(x, \alpha) =: \varphi(x).$$

- Pro každé $\alpha \in P \setminus \{\beta\}$ je funkce $x \mapsto f(x, \alpha)$ měřitelná.
- Existuje funkce $g(x) \in \mathcal{L}(A)$ taková, že pro všechna $\alpha \in P \setminus \{\beta\}$ platí $|f(x, \alpha)| \leq g(x)$ skoro všude v A .

Pak platí rovnost

$$\lim_{\alpha \rightarrow \beta, \alpha \in P} \int_A f(x, \alpha) d\mu(x) = \int_A \varphi(x) dx.$$

Věta - o derivaci integrálu podle parametru

Nechť $I \subset \mathbf{R}$ je otevřený interval a funkce $f(x, \alpha)$ definovaná na $A \times I$ splňuje následující předpoklady:

- Integrál $F(\alpha) := \int_A f(x, \alpha) dx$ konverguje alespoň pro jedno $\alpha \in I$.
- Pro každé $\alpha \in I$ je funkce $x \mapsto f(x, \alpha)$ měřitelná.
- Existuje funkce $g(x) \in \mathcal{L}(A)$ taková, že pro skoro všechna $x \in A$ a všechna $\alpha \in I$ platí $\left| \frac{\partial f(x, \alpha)}{\partial \alpha} \right| \leq g(x)$.

Pak pro všechna $\alpha \in I$ integrál $F(\alpha)$ konverguje a platí

$$\frac{dF}{d\alpha} = \int_A \frac{\partial f(x, \alpha)}{\partial \alpha} dx.$$

Věta - Fubiniova pro Lebesgueův integrál

Nechť $A_1 \subset \mathbf{R}$ a $A_2 \subset \mathbf{R}$ jsou libovolné měřitelné množiny. Označme $A = A_1 \times A_2$. Nechť $f(x, y) \in \mathcal{L}^*(A)$. Potom platí:

- Pro skoro všechna $x \in A_1$ je funkce $g(y) := f(x, y) \in \mathcal{L}^*(A_2)$.

- $H(x) := \int_{A_2} g(y) dy \in \mathcal{L}^*(A_1)$.

Navíc také

$$\int_A f(x, y) d(x, y) = \int_{A_1} \left(\int_{A_2} f(x, y) dy \right) dx.$$

Dále zavedeme *Heavisideovu funkci* $\Theta(\vec{x})$ na \mathbf{R}^r předpisem:

- $\Theta(\vec{x}) = 1$ právě když $x_i > 0$ pro všechna $i \in \{1, \dots, r\}$;
- $\Theta(\vec{x}) = 0$ jinak, tzn. $x_i \leq 0$ alespoň pro jedno $i \in \{1, \dots, r\}$.

1.2 Konvoluce, hustota a jejich vlastnosti

Nyní zadefinujeme konvoluci a hustotu. Důvod pro jejich studium bude zřejmý v další části práce.

Mějme funkce $f(x), g(x) : \mathbf{R} \mapsto \mathbf{R}$. Funkci

$$(f \star g)(x) := \int_{\mathbf{R}} f(s) g(x - s) ds \quad (1.1)$$

nazveme *konvolucí* funkcí $f(x), g(x)$, pokud pravá strana existuje a je vlastní pro skoro všechna $x \in \mathbf{R}$. Pro konvoluci platí, že pokud $f(x), g(x) \in \mathcal{L}_1(\mathbf{R})$, pak funkce $(f \star g)(x)$ existuje a platí $(f \star g)(x) \in \mathcal{L}_1(\mathbf{R})$.

Konvoluce ale může existovat i pro funkce mimo $f(x) \in \mathcal{L}_1(\mathbf{R})$. Určíme konvoluce jednotkových, resp. Heavisideových funkcí:

$$1 \star 1 = \int_{\mathbf{R}} 1(s) 1(x - s) ds = \int_{\mathbf{R}} 1 ds = +\infty;$$

$$\Theta(x) \star \Theta(x) = \int_{\mathbf{R}} \Theta(s) \Theta(x - s) ds = \Theta(x) \int_0^x 1 ds = \Theta(x) x.$$

Konvoluce dvou funkcí $\Theta(x)$ existuje, i když Heavisideova funkce není ze třídy $\mathcal{L}_1(\mathbf{R})$. Konvoluce jednotkových funkcí neexistuje.

Věta

Nechť $f(x), g(x) : \mathbf{R} \mapsto \mathbf{R}$. Pokud existují funkce $F(x), G(x) : \mathbf{R} \mapsto \mathbf{R}$ tak, že pro všechna $x \in \mathbf{R}$ platí $f(x) = \Theta(x) F(x)$ a zároveň $g(x) = \Theta(x) G(x)$, pak $(f \star g)(x) = \Theta(x) \int_0^x F(s) G(x - s) ds$, pokud $F(s) G(x - s) \in \mathcal{L}(\langle 0; x \rangle)$ pro všechna $x \in \mathbf{R}^+$.

Důkaz: Nejprve potvrdíme existenci konvoluce. Protože $F(s) G(x - s) \in \mathcal{L}(\langle 0; x \rangle)$, součin $F(s) G(x - s)$ je na intervalu $\langle 0; x \rangle$ Lebesgueovsky integrovatelný. Integrál $\int_0^x F(s) G(x - s) ds$ tedy existuje a je vlastní pro $x \in \mathbf{R}^+$.

Nyní určíme tvar. Podle definice konvoluce platí $(f \star g)(x) = \int_{\mathbf{R}} \Theta(s) F(s) \Theta(x - s) G(x - s) ds$. Pokud by $s < 0$ a nebo $x - s < 0$, pak by integrand byl roven 0, takže

$$(f \star g)(x) = \Theta(x) \int_0^x F(s) G(x - s) ds.$$

Věta

Ve třídě $\mathcal{L}_1(\mathbf{R})$ je konvoluce komutativní a lineární.

Důkaz: Chceme ukázat, že $(f \star g)(x) = (g \star f)(x)$ pro všechny $f(x), g(x) \in \mathcal{L}_1(\mathbf{R})$. Podle definice je

$$(f \star g)(x) = \int_{-\infty}^{+\infty} f(s) g(x-s) ds.$$

Zavedeme substituci $x-s = t$; $ds = -dt$, pak

$$\int_{-\infty}^{+\infty} f(s) g(x-s) ds = \int_{+\infty}^{-\infty} f(x-t) g(t) (-dt) = \int_{-\infty}^{+\infty} g(t) f(x-t) dt = (g \star f)(x).$$

Konvoluce je tedy komutativní.

Dále chceme ukázat linearitu, tzn. že $((\alpha f + h) \star g)(x) = \alpha(f \star g)(x) + (h \star g)(x)$ pro všechny $f(x), g(x), h(x) \in \mathcal{L}_1(\mathbf{R})$ a $\alpha \in \mathbf{R}$:

$$\begin{aligned} ((\alpha f + h) \star g)(x) &= \int_{\mathbf{R}} (\alpha f + h)(s) g(x-s) ds = \int_{\mathbf{R}} (\alpha f(s) + h(s)) g(x-s) ds = \\ &= \int_{\mathbf{R}} \alpha f(s) g(x-s) ds + \int_{\mathbf{R}} h(s) g(x-s) ds = \alpha \int_{\mathbf{R}} f(s) g(x-s) ds + \int_{\mathbf{R}} h(s) g(x-s) ds = \\ &= \alpha(f \star g)(x) + (h \star g)(x). \end{aligned}$$

Tím je důkaz ukončen.

V následující části zadefinujeme pojmy, jež jsou pro naši práci stěžejní: hustota a konvoluce. Řekneme, že funkce $f(x) : \mathbf{R} \mapsto \mathbf{R}$ je *hustotou*, pokud jsou splněny následující předpoklady:

- 1) $\forall x \in \mathbf{R} : f(x) \geq 0$,
- 2) $f(x) \in \mathcal{L}_1(\mathbf{R})$,
- 3) $\int_{\mathbf{R}} f(x) dx = 1$.

Věta

Jsou-li $f(x), g(x)$ hustoty, pak konvoluce $(f \star g)(x)$ existuje a je rovněž hustotou.

Důkaz: Z definice hustoty víme, že $f(x), g(x) \in \mathcal{L}_1(\mathbf{R}) \Rightarrow (f \star g)(x)$ existuje. Ověříme splnění všech předpokladů:

ad 1) Protože pro $\forall x \in \mathbf{R} : f(x) \geq 0 \wedge g(x) \geq 0$, pak jistě $(f \star g)(x) = \int_{\mathbf{R}} f(s) g(x-s) ds \geq 0$.

ad 2) Obě funkce $f(x), g(x) \in \mathcal{L}_1(\mathbf{R})$, podle výše uvedené věty platí $(f \star g)(x) \in \mathcal{L}_1(\mathbf{R})$.

ad 3) Dosadíme definiční vztah konvoluce

$$\int_{\mathbf{R}} (f \star g)(x) dx = \int_{\mathbf{R}} \left[\int_{\mathbf{R}} (f(s) g(x-s) ds) \right] dx.$$

Aplikujeme Fubiniovu větu a zavedeme substituci $x-s = t$; $dx = dt$, pak

$$\int_{\mathbf{R}} \left[\int_{\mathbf{R}} (f(s) g(x-s) ds) \right] dx = \int_{\mathbf{R}} f(s) \left[\int_{\mathbf{R}} g(x-s) dx \right] ds = \int_{\mathbf{R}} f(s) \left[\int_{\mathbf{R}} g(t) dt \right] ds.$$

Protože $\int_{\mathbf{R}} f(x) dx = 1 \wedge \int_{\mathbf{R}} g(x) dx = 1$, dostáváme rovnost

$$\int_{\mathbf{R}} f(s) \left[\int_{\mathbf{R}} g(t) dt \right] ds = \int_{\mathbf{R}} f(s) ds = 1.$$

Věta

Nechť $f(x), g(x)$ jsou hustoty. Označme $\mu = \int_{\mathbf{R}} x f(x) dx \in \mathbf{R}$ a $\nu = \int_{\mathbf{R}} x g(x) dx \in \mathbf{R}$. Dále označme $h(z) = (f \star g)(z)$. Pak platí

$$\int_{\mathbf{R}} z h(z) dz = \mu + \nu,$$

pokud integrál na levé straně předchozí rovnosti existuje.

Důkaz: Z definice konvoluce je $h(z) = (f \star g)(z) = \int_{\mathbf{R}} f(s) g(z-s) ds$. Ta podle předchozí věty existuje a je to také hustota. Použijeme Fubiniovu větu

$$\int_{\mathbf{R}} z h(z) dz = \int_{\mathbf{R}} z \left[\int_{\mathbf{R}} f(s) g(z-s) ds \right] dz = \int_{\mathbf{R}} f(s) \left[\int_{\mathbf{R}} z g(z-s) dz \right] ds.$$

Zavedeme substituci $z-s = y, dz = dy$, pak

$$\begin{aligned} & \int_{\mathbf{R}} f(s) \left[\int_{\mathbf{R}} z g(z-s) dz \right] ds = \\ &= \int_{\mathbf{R}} f(s) \left[\int_{\mathbf{R}} (y+s) g(y) dy \right] ds = \int_{\mathbf{R}} f(s) \left[\int_{\mathbf{R}} y g(y) dy + s \int_{\mathbf{R}} g(y) dy \right] ds. \end{aligned}$$

Podle předpokladu $\int_{\mathbf{R}} y g(y) dy = \nu$ a $g(y)$ je hustota, tedy $\int_{\mathbf{R}} g(y) dy = 1$. Pak platí

$$\int_{\mathbf{R}} f(s) \left[\int_{\mathbf{R}} y g(y) dy + s \int_{\mathbf{R}} g(y) dy \right] ds = \int_{\mathbf{R}} f(s) (\nu + s) ds = \nu \int_{\mathbf{R}} f(s) ds + \int_{\mathbf{R}} s f(s) ds.$$

Z předpokladů také víme, že $\int_{\mathbf{R}} s f(s) ds = \mu$ a $f(s)$ je hustota, tedy $\int_{\mathbf{R}} f(s) ds = 1$. Dostáváme

$$\nu \int_{\mathbf{R}} f(s) ds + \int_{\mathbf{R}} s f(s) ds = \nu + \mu = \int_{\mathbf{R}} z h(z) dz.$$

1.3 Laplaceova a Fourierova transformace

V této sekci si představíme dvě integrální transformace, jejichž vlastnosti využijeme při odvození konvoluce dvou hustot. Zavedeme Laplaceův prostor $\mathcal{P}(\mathbf{R})$, což je třída všech funkcí $g(x) : \mathbf{R} \mapsto \mathbf{R}$ jež jsou rovny 0 pro všechna $x < 0$, jsou exponenciálního růstu a jsou po částech spojitě na $\langle 0; +\infty \rangle$. Nyní lze zavést Laplaceovu transformaci.

Mějme Laplaceův prostor $\mathcal{P}(\mathbf{R})$. Laplaceův obraz funkce $f(x) \in \mathcal{L}_1(\mathbf{R}) \cap \mathcal{P}(\mathbf{R})$ je pro $s \in \mathbf{R}$ definován vztahem

$$\mathcal{L}[f(x)](s) := \int_{\mathbf{R}} f(x) e^{-sx} dx = F(s).$$

Poznámka

Funkce $g(x)$ je exponenciálního růstu, pokud existují $\alpha \in \mathbf{R}, x_0 \in \mathbf{R}, K > 0$ tak, že pro všechna $x > x_0$ platí, že $|g(x)| \leq K e^{\alpha x}$.

Věta

Nechť $f, g \in \mathcal{L}_1(\mathbf{R}) \cap \mathcal{P}(\mathbf{R})$, kde $\mathcal{P}(\mathbf{R})$ je Laplaceův prostor. Necht' $\mathcal{L}[f]$, resp. $\mathcal{L}[g]$ je Laplaceův obraz funkce $f(x)$, resp. $g(x)$. Pak platí

$$\mathcal{L}[f \star g] = \mathcal{L}[f] \cdot \mathcal{L}[g]. \quad (1.2)$$

Důkaz: Dosadíme vztah pro konvoluci a aplikujeme Fubiniovu větu, pak

$$\mathcal{L}[f \star g](s) = \int_{\mathbf{R}} \left[\int_{\mathbf{R}} f(a) g(x-a) da \right] e^{-sx} dx = \int_{\mathbf{R}} f(a) \left[\int_{\mathbf{R}} g(x-a) e^{-sx} dx \right] da.$$

Zavedeme substituci $x - a = y$, $dx = dy$, takže

$$\begin{aligned} \int_{\mathbf{R}} f(a) \left[\int_{\mathbf{R}} g(x-a) e^{-sx} dx \right] da &= \int_{\mathbf{R}} f(a) \left[\int_{\mathbf{R}} g(y) e^{-s(y+a)} dy \right] da = \\ &= \int_{\mathbf{R}} f(a) e^{-sa} \left[\int_{\mathbf{R}} g(y) e^{-sy} dy \right] da = \int_{\mathbf{R}} f(a) e^{-sa} da \cdot \int_{\mathbf{R}} g(y) e^{-sy} dy. \end{aligned}$$

Obdrželi jsme zřejmě součin Laplaceových obrazů funkcí $f(x)$ a $g(x)$

$$\int_{\mathbf{R}} f(a) e^{-sa} da \cdot \int_{\mathbf{R}} g(y) e^{-sy} dy = \mathcal{L}[f](s) \cdot \mathcal{L}[g](s) = \mathcal{L}[f \star g](s).$$

Fourierův obraz funkce $f(x) \in \mathcal{L}_1(\mathbf{R})$ je pro $t \in \mathbf{R}$ definován vztahem

$$\mathfrak{F}[f(x)](t) := \int_{-\infty}^{+\infty} f(x) e^{itx} dx = F(t).$$

Věta

Necht' $f(x), g(x) \in \mathcal{L}_1(\mathbf{R})$. Necht' $\mathfrak{F}[f]$, resp. $\mathfrak{F}[g]$ je Fourierův obraz funkce $f(x)$, resp. $g(x)$. Pak platí

$$\mathfrak{F}[f \star g] = \mathfrak{F}[f] \cdot \mathfrak{F}[g]. \quad (1.3)$$

Důkaz: Dosadíme vztah pro konvoluci a následně definiční vztah pro Fourierův obraz

$$\mathfrak{F}[f \star g] = \mathfrak{F} \left[\int_{\mathbf{R}} f(s) g(x-s) ds \right] = \int_{\mathbf{R}} \left(\int_{\mathbf{R}} f(s) g(x-s) ds \right) e^{itx} dx.$$

Aplikací Fubiniovy věty obdržíme

$$\int_{\mathbf{R}} \left(\int_{\mathbf{R}} f(s) g(x-s) ds \right) e^{itx} dx = \int_{\mathbf{R}} f(s) \left(\int_{\mathbf{R}} g(x-s) e^{itx} dx \right) ds.$$

Po zavedení substituce $x - s = a$, $dx = da$ lze upravit na tvar

$$\begin{aligned} \int_{\mathbf{R}} f(s) \left(\int_{\mathbf{R}} g(x-s) e^{itx} dx \right) ds &= \int_{\mathbf{R}} f(s) \left(\int_{\mathbf{R}} g(a) e^{it(s+a)} da \right) ds = \\ &= \int_{\mathbf{R}} f(s) e^{its} \left(\int_{\mathbf{R}} g(a) e^{ita} da \right) ds = \int_{\mathbf{R}} f(s) e^{its} ds \cdot \int_{\mathbf{R}} g(a) e^{ita} da. \end{aligned}$$

Porovnáním s definicí Fourierova obrazu tedy dostáváme rovnost

$$\int_{\mathbf{R}} f(s) e^{its} ds \cdot \int_{\mathbf{R}} g(a) e^{ita} da = \mathfrak{F}[f](t) \cdot \mathfrak{F}[g](t) = \mathfrak{F}[f \star g](t).$$

Souvislost Laplaceova obrazu funkce $x^n f(x)$ a Laplaceova obrazu funkce $f(x)$

Uvažujme funkce $f(x) \in \mathcal{L}_1(\mathbf{R}) \cap \mathcal{P}(\mathbf{R})$ a $x^n f(x) \in \mathcal{L}_1(\mathbf{R}) \cap \mathcal{P}(\mathbf{R})$. Z definice Laplaceova obrazu dostaneme

$$\mathcal{L}[x^n f(x)](s) = \int_{\mathbf{R}} x^n f(x) e^{-sx} dx.$$

Mějme $\mathcal{L}[f(x)](s) = \int_{\mathbf{R}} f(x) e^{-sx} dx$. Tento integrál postupně derivujeme podle parametru s . Využitím věty o záměně integrálu a derivace, jejíž podmínky jsou splněny volbou definičního oboru, lze psát

- 1. derivace $\frac{d}{ds} \left(\int_{\mathbf{R}} f(x) e^{-sx} dx \right) = \int_{\mathbf{R}} \left[\frac{\partial}{\partial s} (f(x) e^{-sx}) \right] dx = \int_{\mathbf{R}} (-x) f(x) e^{-sx} dx,$
- 2. derivace $\frac{d^2}{ds^2} \left(\int_{\mathbf{R}} f(x) e^{-sx} dx \right) = \int_{\mathbf{R}} \left[\frac{\partial^2}{\partial s^2} (f(x) e^{-sx}) \right] dx = \int_{\mathbf{R}} (-x)^2 f(x) e^{-sx} dx,$
- 3. derivace $\frac{d^3}{ds^3} \left(\int_{\mathbf{R}} f(x) e^{-sx} dx \right) = \int_{\mathbf{R}} \left[\frac{\partial^3}{\partial s^3} (f(x) e^{-sx}) \right] dx = \int_{\mathbf{R}} (-x)^3 f(x) e^{-sx} dx.$

Pro n -tou derivaci podle s pak platí

$$\frac{d^n}{ds^n} \mathcal{L}[f(x)](s) = \frac{d^n}{ds^n} \left(\int_{\mathbf{R}} f(x) e^{-sx} dx \right) = (-1)^n \int_{\mathbf{R}} x^n f(x) e^{-sx} dx = (-1)^n \mathcal{L}[x^n f(x)](s).$$

Dostáváme tedy vztah

$$\frac{d^n}{ds^n} \mathcal{L}[f(x)](s) = (-1)^n \mathcal{L}[x^n f(x)](s). \quad (1.4)$$

Souvislost Fourierova obrazu funkce $x^n f(x)$ a Fourierova obrazu funkce $f(x)$

Uvažujme funkce $f(x) \in \mathcal{L}_1(\mathbf{R})$ a $x^n f(x) \in \mathcal{L}_1(\mathbf{R})$. Z definice Fourierova obrazu

$$\mathfrak{F}[x^n f(x)](t) = \int_{\mathbf{R}} x^n f(x) e^{itx} dx.$$

Mějme $\mathfrak{F}[f(x)](t) = \int_{\mathbf{R}} f(x) e^{itx} dx$. Tento integrál postupně derivujeme podle parametru t . Využitím věty o záměně integrálu a derivace dostaneme

- 1. derivace $\frac{d}{dt} \left(\int_{\mathbf{R}} f(x) e^{itx} dx \right) = \int_{\mathbf{R}} \left[\frac{\partial}{\partial t} (f(x) e^{itx}) \right] dx = \int_{\mathbf{R}} (ix) f(x) e^{itx} dx,$
- 2. derivace $\frac{d^2}{dt^2} \left(\int_{\mathbf{R}} f(x) e^{itx} dx \right) = \int_{\mathbf{R}} \left[\frac{\partial^2}{\partial t^2} (f(x) e^{itx}) \right] dx = \int_{\mathbf{R}} (ix)^2 f(x) e^{itx} dx,$
- 3. derivace $\frac{d^3}{dt^3} \left(\int_{\mathbf{R}} f(x) e^{itx} dx \right) = \int_{\mathbf{R}} \left[\frac{\partial^3}{\partial t^3} (f(x) e^{itx}) \right] dx = \int_{\mathbf{R}} (ix)^3 f(x) e^{itx} dx.$

Pro n -tou derivaci podle t pak platí

$$\frac{d^n}{dt^n} \mathfrak{F}[f(x)](t) = \frac{d^n}{dt^n} \left(\int_{\mathbf{R}} f(x) e^{itx} dx \right) = i^n \int_{\mathbf{R}} x^n f(x) e^{itx} dx = i^n \mathfrak{F}[x^n f(x)](t).$$

Dostáváme tedy vztah

$$\frac{d^n}{dt^n} \mathfrak{F}[f(x)](t) = i^n \mathfrak{F}[x^n f(x)](t). \quad (1.5)$$

Souvislost Laplaceova obrazu funkce $f(x - a)$ a Laplaceova obrazu funkce $f(x)$

Podle definice platí rovnost

$$\mathcal{L}[f(x - a)](s) = \int_{\mathbf{R}} f(x - a) e^{-sx} dx.$$

Zavedeme substituci $y = x - a$, $dy = dx$, takže po úpravách dostaneme

$$\int_{\mathbf{R}} f(x - a) e^{-sx} dx = \int_{\mathbf{R}} f(y) e^{-s(y+a)} dy = e^{-sa} \int_{\mathbf{R}} f(y) e^{-sy} dy = e^{-sa} \mathcal{L}[f(y)].$$

Výsledný vztah je

$$\mathcal{L}[f(x - a)](s) = e^{-sa} \mathcal{L}[f(x)](s). \quad (1.6)$$

Souvislost Fourierova obrazu funkce $f(x - a)$ a Fourierova obrazu funkce $f(x)$

Podle definice platí

$$\mathfrak{F}[f(x - a)](t) = \int_{\mathbf{R}} f(x - a) e^{itx} dx.$$

Zavedeme substituci $y = x - a$, $dy = dx$ a po jednoduché úpravě máme

$$\int_{\mathbf{R}} f(x - a) e^{itx} dx = \int_{\mathbf{R}} f(y) e^{it(y+a)} dy = e^{ita} \int_{\mathbf{R}} f(y) e^{ity} dy = e^{ita} \mathfrak{F}[f(y)].$$

Výsledný vztah je

$$\mathfrak{F}[f(x - a)](t) = e^{ita} \mathfrak{F}[f(x)](t). \quad (1.7)$$

Odvození Laplaceova obrazu exponenciální hustoty

Mějme exponenciální hustotu ve tvaru $p(x) = \Theta(x) A e^{-\lambda x}$, kde $\lambda > 0$. Nejprve určíme hodnotu konstanty A normováním, tedy ze vztahu $\int_{\mathbf{R}} p(x) dx = 1$,

$$\int_{\mathbf{R}} \Theta(x) A e^{-\lambda x} dx = A \int_0^{+\infty} e^{-\lambda x} dx = A \left[\frac{e^{-\lambda x}}{-\lambda} \right]_0^{+\infty} = A \frac{1}{\lambda} = 1.$$

Konstanta A má tedy hodnotu $A = \lambda$ a vztah pro exponenciální hustotu je

$$p(x) = \Theta(x) \lambda e^{-\lambda x}, \quad \lambda > 0.$$

Nyní můžeme hledat Laplaceův obraz dosazením $p(x)$ do definičního vztahu

$$\begin{aligned} \mathfrak{L}[p(x)] &= \int_{\mathbf{R}} \Theta(x) \lambda e^{-\lambda x} e^{-sx} dx = \lambda \int_{\mathbf{R}} \Theta(x) e^{-(\lambda+s)x} dx = \lambda \int_0^{+\infty} e^{-(\lambda+s)x} dx = \\ &= \lambda \left[\frac{e^{-(\lambda+s)x}}{-(\lambda+s)} \right]_0^{+\infty} = \lambda \frac{1}{\lambda+s}. \end{aligned}$$

Nalezli jsme tedy Laplaceův obraz ve tvaru

$$\mathfrak{L}[\Theta(x) \lambda e^{-\lambda x}] = \frac{\lambda}{\lambda+s}. \quad (1.8)$$

Odvození Laplaceova obrazu Gamma hustoty

Mějme Gamma hustotu ve tvaru $p(x) = \Theta(x) \frac{\lambda^{\alpha+1}}{\Gamma(\alpha+1)} x^{\alpha} e^{-\lambda x}$, kde $\lambda > 0$ a $\alpha \in \mathbf{R}^+$. Ve jmenovateli vystupuje Γ funkce, která je definována vztahem $\Gamma(\alpha + 1) = \int_0^{+\infty} x^{\alpha} e^{-x} dx$, jde o tzv. Eulerův integrál druhého druhu. Při výpočtu Laplaceova obrazu Gamma hustoty vyjdeme tedy z definice Laplaceovy transformace

$$\begin{aligned} \mathfrak{L}[p(x)] &= \mathfrak{L}\left[\Theta(x) \frac{\lambda^{\alpha+1}}{\Gamma(\alpha+1)} x^{\alpha} e^{-\lambda x}\right] = \int_{-\infty}^{+\infty} \Theta(x) \frac{\lambda^{\alpha+1}}{\Gamma(\alpha+1)} x^{\alpha} e^{-\lambda x} e^{-sx} dx = \\ &= \frac{\lambda^{\alpha+1}}{\Gamma(\alpha+1)} \int_0^{+\infty} x^{\alpha} e^{-(\lambda+s)x} dx. \end{aligned}$$

Zavedením substituce $y = (\lambda + s)x$, $dy = (\lambda + s) dx$ dostaneme

$$\frac{\lambda^{\alpha+1}}{\Gamma(\alpha+1)} \int_0^{+\infty} x^{\alpha} e^{-(\lambda+s)x} dx = \frac{\lambda^{\alpha+1}}{\Gamma(\alpha+1)} \int_0^{+\infty} \left(\frac{y}{\lambda+s}\right)^{\alpha} e^{-y} \frac{dy}{\lambda+s} =$$

$$= \frac{\lambda^{\alpha+1}}{\Gamma(\alpha+1)} \frac{1}{(\lambda+s)^{\alpha+1}} \int_0^{+\infty} y^\alpha e^{-y} dy.$$

Jak již bylo zmíněno, integrál v tomto tvaru definuje Γ funkci, rovnost tedy můžeme dále upravit

$$\frac{\lambda^{\alpha+1}}{\Gamma(\alpha+1)} \frac{1}{(\lambda+s)^{\alpha+1}} \int_0^{+\infty} y^\alpha e^{-y} dy = \frac{\lambda^{\alpha+1}}{\Gamma(\alpha+1)} \frac{1}{(\lambda+s)^{\alpha+1}} \Gamma(\alpha+1) = \left(\frac{\lambda}{\lambda+s} \right)^{\alpha+1}.$$

Výsledný vztah má tedy tvar

$$\mathfrak{L} \left[\Theta(x) \frac{\lambda^{\alpha+1}}{\Gamma(\alpha+1)} x^\alpha e^{-\lambda x} \right] = \left(\frac{\lambda}{\lambda+s} \right)^{\alpha+1}, \quad \alpha \in \mathbf{R}^+, \lambda > 0. \quad (1.9)$$

Odvození Laplaceova obrazu Erlangovy hustoty

Mějme Erlangovu hustotu ve tvaru $p(x) = \Theta(x) \frac{\lambda^{\alpha+1}}{\alpha!} x^\alpha e^{-\lambda x}$, kde $\lambda > 0, \alpha \in \mathbf{N}$. Zřejmě jde o speciální případ Gamma hustoty, neboť pro $\alpha \in \mathbf{N}$ platí rovnost $\Gamma(\alpha+1) = \alpha!$. Postup při odvození Laplaceova obrazu by byl identický s postupem u Gamma hustoty. Odkazujeme se tedy pouze na vztah (1.9), odkud v kombinaci s výše zmíněnou vlastností Gamma funkce dostáváme Laplaceův obraz ve tvaru

$$\mathfrak{L} \left[\Theta(x) \frac{\lambda^{\alpha+1}}{\alpha!} x^\alpha e^{-\lambda x} \right] = \left(\frac{\lambda}{\lambda+s} \right)^{\alpha+1}, \quad \alpha \in \mathbf{N}, \lambda > 0. \quad (1.10)$$

Odvození Fourierova obrazu Gaussovy hustoty

Nyní odvodíme pomocný vztah, který dále aplikujeme při výpočtu konvoluce dvou Gaussových hustot. Mějme Gaussovu hustotu ve tvaru $p(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$, pro jejíž parametry platí $-\infty < \mu < \infty$ a $\sigma > 0$. Uvažujme nejprve funkci $d(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}}$, jejíž Fourierův obraz má tvar

$$\mathfrak{F}[d(x)] = \int_{\mathbf{R}} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}} e^{itx} dx.$$

Pomocí rovnosti $e^{itx} = \cos(tx) + i \sin(tx)$ lze rozdělit na dva integrály

$$\begin{aligned} \int_{\mathbf{R}} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}} e^{itx} dx &= \frac{1}{\sigma\sqrt{2\pi}} \int_{\mathbf{R}} e^{-\frac{x^2}{2\sigma^2}} (\cos(tx) + i \sin(tx)) dx = \\ &= \frac{1}{\sigma\sqrt{2\pi}} \left[\int_{\mathbf{R}} e^{-\frac{x^2}{2\sigma^2}} \cos(tx) dx + i \int_{\mathbf{R}} e^{-\frac{x^2}{2\sigma^2}} \sin(tx) dx \right]. \end{aligned}$$

Funkce $e^{-\frac{x^2}{2\sigma^2}} \sin(tx)$ je lichá a její integrál existuje, takže po integrování přes celý obor \mathbf{R} bude hodnota integrálu rovna 0. Funkce $e^{-\frac{x^2}{2\sigma^2}} \cos(tx)$ je naopak sudá, integrál také existuje a jeho hodnota na intervalech $(-\infty; 0)$ a $(0; +\infty)$ je stejná. Pak lze psát

$$\frac{1}{\sigma\sqrt{2\pi}} \left[\int_{\mathbf{R}} e^{-\frac{x^2}{2\sigma^2}} \cos(tx) dx + i \int_{\mathbf{R}} e^{-\frac{x^2}{2\sigma^2}} \sin(tx) dx \right] = \frac{2}{\sigma\sqrt{2\pi}} \int_0^{+\infty} e^{-\frac{x^2}{2\sigma^2}} \cos(tx) dx.$$

Aplikujeme vztah $\int_0^{+\infty} e^{-ax^2} \cos(bx) dx = \frac{1}{2} \sqrt{\frac{\pi}{a}} e^{-\frac{b^2}{4a}}$, viz [22], str. 238. Označme $a = \frac{1}{2\sigma^2}$ a $b = t$, dostáváme tak

$$\mathfrak{F}[d(x)] = \frac{2}{\sigma\sqrt{2\pi}} \int_0^{+\infty} e^{-\frac{x^2}{2\sigma^2}} \cos(tx) dx = \frac{2}{\sigma\sqrt{2\pi}} \frac{1}{2} \sqrt{\frac{\pi}{\frac{1}{2\sigma^2}}} e^{-\frac{t^2}{2\sigma^2}} = \frac{\sigma\sqrt{2\pi}}{\sigma\sqrt{2\pi}} e^{-\frac{2\sigma^2 t^2}{4}} = e^{-\frac{(\sigma t)^2}{2}}.$$

Je zřejmé, že $p(x) = d(x - \mu)$. Využijeme vlastnosti (1.7), takže

$$\mathfrak{F}[p(x)] = \mathfrak{F}[d(x - \mu)] = e^{it\mu} \mathfrak{F}[d(x)].$$

Obdržíme tak hledaný vztah pro Fourierův obraz Gaussova rozdělení

$$\mathfrak{F}\left[\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\right] = e^{it\mu} e^{-\frac{(\sigma t)^2}{2}}. \quad (1.11)$$

1.4 Konvoluce hustot

Konvoluce dvou exponenciálních hustot

Mějme exponenciální hustotu $p(x) = \Theta(x)\lambda e^{-\lambda x}$, $\lambda > 0$, $x \in \mathbf{R}$. Protože $p(x)$ je hustota, $(p \star p)$ existuje a patří do $\mathcal{L}_1(\mathbf{R})$. Konvoluci odvodíme z definice konvoluce takto: do definičního vztahu konvoluce dosadíme hustotu $p(x)$

$$\begin{aligned} (p \star p)(x) &= \int_{\mathbf{R}} p(s) p(x-s) ds = \int_{\mathbf{R}} \Theta(s)\lambda e^{-\lambda s} \Theta(x-s)\lambda e^{-\lambda(x-s)} ds = \\ &= \lambda^2 \int_{\mathbf{R}} \Theta(s)\Theta(x-s) e^{-\lambda x} ds = \lambda^2 \int_0^{+\infty} \Theta(x-s) e^{-\lambda x} ds. \end{aligned}$$

Abychom mohli integrovat, odstraníme z argumentu integrálu funkci $\Theta(x-s)$. Zjevně musí být $x-s > 0$, takže $s \in (0, x)$. Pak musí x být kladné, což zaručíme vynásobením integrálu funkcí $\Theta(x)$, tedy

$$\begin{aligned} \lambda^2 \int_0^{+\infty} \Theta(x-s) e^{-\lambda x} ds &= \lambda^2 \Theta(x) \int_0^x e^{-\lambda x} ds = \lambda^2 \Theta(x) e^{-\lambda x} \int_0^x 1 ds = \\ &= \lambda^2 \Theta(x) e^{-\lambda x} x = \lambda x \Theta(x) \lambda e^{-\lambda x}. \end{aligned}$$

Zpětným dosazením $p(x)$ docházíme k výslednému vztahu

$$(p \star p)(x) = \lambda x p(x). \quad (1.12)$$

Konvoluce dvou Gaussových hustot

Mějme dvě Gaussovy hustoty $p(x) = \frac{1}{\sigma_1\sqrt{2\pi}} e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}$, $q(x) = \frac{1}{\sigma_2\sqrt{2\pi}} e^{-\frac{(x-\mu_2)^2}{2\sigma_2^2}}$ s různými parametry $\mu_1 \neq \mu_2$ a $\sigma_1 \neq \sigma_2$. Zkombinujeme vztah pro Fourierův obraz konvoluce (1.3) a Fourierův obraz Gaussova rozdělení (1.11)

$$\mathfrak{F}[p \star q] = \mathfrak{F}[p] \cdot \mathfrak{F}[q] = e^{it\mu_1} e^{-\frac{(\sigma_1 t)^2}{2}} e^{it\mu_2} e^{-\frac{(\sigma_2 t)^2}{2}} = e^{it(\mu_1+\mu_2)} e^{-\frac{(\sigma_1^2+\sigma_2^2)t^2}{2}} = e^{it(\mu_1+\mu_2)} e^{-\frac{(\sqrt{\sigma_1^2+\sigma_2^2}t)^2}{2}}.$$

Označme nyní $\mu = \mu_1 + \mu_2$ a $\sigma = \sqrt{\sigma_1^2 + \sigma_2^2}$, potom

$$\mathfrak{F}[p \star q] = e^{it\mu} e^{-\frac{(\sigma t)^2}{2}} = \mathfrak{F}[r], \quad r(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

Pokud se rovnají Fourierovy obrazy funkcí, musejí se rovnat i funkce samotné, takže pro konvoluci dostáváme

$$(p \star q)(x) = \frac{1}{\sqrt{2\pi(\sigma_1^2 + \sigma_2^2)}} e^{-\frac{(x-(\mu_1+\mu_2))^2}{2(\sigma_1^2+\sigma_2^2)}}. \quad (1.13)$$

Konvoluce dvou Gamma hustot

Mějme dvě Gamma hustoty $p(x) = \Theta(x) \frac{\lambda^{\alpha+1}}{\Gamma(\alpha+1)} x^\alpha e^{-\lambda x}$, $q(x) = \Theta(x) \frac{\lambda^{\beta+1}}{\Gamma(\beta+1)} x^\beta e^{-\lambda x}$, kde $\lambda > 0$, $\alpha \in \mathbf{R}^+$, $\beta \in \mathbf{R}^+$. K odvození konvoluce těchto hustot využijeme Laplaceův obraz konvoluce, do něhož dosadíme vztah (1.9), tedy

$$\mathfrak{L}[p \star q] = \mathfrak{L}[p] \cdot \mathfrak{L}[q] = \left(\frac{\lambda}{\lambda+s}\right)^{\alpha+1} \left(\frac{\lambda}{\lambda+s}\right)^{\beta+1} = \left(\frac{\lambda}{\lambda+s}\right)^{\alpha+\beta+2}.$$

Označme $\gamma = \alpha + \beta + 1$. Protože $\alpha \in \mathbf{R}^+$ a $\beta \in \mathbf{R}^+$, pak jistě i $\gamma \in \mathbf{R}^+$. Na pravé straně předchozí rovnosti dostaneme výraz $\left(\frac{\lambda}{\lambda+s}\right)^{\gamma+1}$, což je podle vztahu (1.9) Laplaceův obraz hustoty $q(x)$, která má tvar

$$q(x) = \Theta(x) \frac{\lambda^{\gamma+1}}{\Gamma(\gamma+1)} x^\gamma e^{-\lambda x}.$$

Dokázali jsme, že platí rovnost

$$\mathfrak{L}[p \star q] = \mathfrak{L}\left[\Theta(x) \frac{\lambda^{\alpha+\beta+2}}{\Gamma(\alpha+\beta+2)} x^{\alpha+\beta+1} e^{-\lambda x}\right].$$

Pokud se však rovnají Laplaceovy obrazy funkcí, musejí se rovnat i funkce samotné, viz Lerchův teorém [23], takže docházíme k rovnici

$$(p \star q) = \Theta(x) \frac{\lambda^{\alpha+\beta+2}}{\Gamma(\alpha+\beta+2)} x^{\alpha+\beta+1} e^{-\lambda x}, \quad \alpha \in \mathbf{R}^+, \beta \in \mathbf{R}^+. \quad (1.14)$$

Konvoluce dvou Erlangových hustot

Mějme pro $\lambda > 0, x \in \mathbf{R}$ dvě Erlangovy hustoty $p(x) = \Theta(x) \frac{\lambda^{\alpha+1}}{\alpha!} x^\alpha e^{-\lambda x}$, $q(x) = \Theta(x) \frac{\lambda^{\beta+1}}{\beta!} x^\beta e^{-\lambda x}$, kde $\alpha, \beta \in \mathbf{N}_0$. Protože $p(x), q(x)$ jsou hustoty, $(p \star q)$ existuje a patří do $\mathcal{L}_1(\mathbf{R})$. K výpočtu konvoluce opět využijeme fakt, že Erlangova hustota je speciální případ Gamma hustoty, protože pro $\alpha, \beta \in \mathbf{N}_0$ platí rovnost $\Gamma(\alpha+1) = \alpha!$, resp. $\Gamma(\beta+1) = \beta!$. Postup při odvození konvoluce by byl identický s postupem u Gamma hustoty. Odkazujeme se tedy pouze na vztah (1.14), odkud v kombinaci s výše zmíněnou vlastností Γ - funkce dostáváme

$$(p \star q) = \Theta(x) \frac{\lambda^{\alpha+\beta+2}}{(\alpha+\beta+1)!} x^{\alpha+\beta+1} e^{-\lambda x}, \quad \alpha \in \mathbf{N}_0, \beta \in \mathbf{N}_0. \quad (1.15)$$

Konvoluci pro dvě Erlangovy hustoty $p(x)$ o stejném parametru $\alpha \in \mathbf{N}_0$ lze po krátké úpravě vyjádřit rovností

$$(p \star p)(x) = \frac{\alpha!}{(2\alpha+1)!} (\lambda x)^{\alpha+1} p(x). \quad (1.16)$$

Kapitola 2

Základní poznatky z teorie pravděpodobnosti

V úvodu druhé kapitoly zavedeme několik významných pojmů z teorie pravděpodobnosti, přičemž budeme vycházet z teorie míry. Tato problematika je podrobně zpracována v práci [21], z níž budeme v následujícím textu čerpat. V další sekci se budeme věnovat absolutně spojitě náhodné veličině a absolutně spojitěmu rozdělení, kde navážeme na znalosti z kapitoly první a nastíníme statistické metody pro prokládání empirických dat. Závěr tvoří souhrn dopravních veličin a stručná specifikace použitých dopravních dat.

2.1 Pravděpodobnost a náhodná veličina

Pravděpodobnost

Uvažujme nějaký náhodný pokus. Množinu všech možných výsledků takového pokusu nazveme *základní pravděpodobnostní prostor* Ω . Na tomto prostoru můžeme zavést množinovou σ -algebru $\mathcal{X} \subset 2^\Omega$. Každou normalizovanou σ -aditivní (tedy spočetně aditivní) míru

$$P(A) : \mathcal{X} \mapsto \langle 0, 1 \rangle$$

pak nazveme *pravděpodobnostní mírou* (*pravděpodobností*) na \mathcal{X} . Míra se nazývá normalizovaná, pokud pro ni platí podmínka $P[\Omega] = 1$.

Každá pravděpodobnostní míra musí splňovat 3 axiomy:

- *axiom nezápornosti*: $\forall A \in \mathcal{X} : P(A) \geq 0$,
- *axiom normality*: $P(\Omega) = 1$,
- *axiom σ -aditivity*: Pro všechny disjunktní posloupnosti jevů platí $P(\uplus_{\ell=1}^{\infty} A_\ell) = \sum_{\ell=1}^{\infty} P(A_\ell)$, kde symbol \uplus představuje disjunktní sjednocení.

Z těchto axiomů vyplývají následující důležité vlastnosti:

- $P(\emptyset) = 0$,
- $0 \leq P(A) \leq 1$,
- *monotonie*: $A \subset B \implies P(A) \leq P(B)$,
- *aditivita*: $P(A \uplus B) = P(A) + P(B)$,

Řekneme, že dva jevy $A, B \subset \Omega$ jsou *nezávislé*, pokud platí

$$P(A \cap B) = P(A) \cdot P(B). \quad (2.1)$$

Díky pravděpodobnostní míře lze zavést tzv. *pravděpodobnostní prostor*, který je tvořen trojicí $\{\Omega, \mathcal{X}, P\}$, tedy základním pravděpodobnostním prostorem Ω , σ -algebrou $\mathcal{X} \subset 2^\Omega$ P -měřitelných množin a příslušnou mírou $P(A)$. V takovém pravděpodobnostním prostoru nazveme *náhodnou veličinou* každé zobrazení $X : \Omega \mapsto \mathbf{R}$ takové, že pro $\forall c \in \mathbf{R}$ platí

$$X^{-1}((-\infty, c)) = [X \leq c] = \{\omega \in \Omega : X(\omega) \leq c\} \in \mathcal{X}.$$

Tento vztah klade na vzory všech intervalů $(-\infty, c)$ podmínku, aby byly P -měřitelnými množinami. Z hlediska obecné teorie míry tedy definice náhodné veličiny odpovídá definici měřitelné funkce.

Pravděpodobnost, že náhodná veličina X nabude hodnoty menší než x , vyjádříme jako $P\{\omega \in \Omega : X(\omega) < x\}$ a značíme symbolem $P[X < x]$. Pravděpodobnost, že náhodná veličina X nabude hodnoty z množiny A se tedy rovná $P[X \in A] = P\{\omega \in \Omega : X(\omega) \in A\}$. Obdobně lze zavést $P[X \geq x]$, $P[X = 3]$, a tak dále.

Pro pravděpodobnostní prostor $\{\Omega, \mathcal{X}, P\}$ a náhodnou veličinu $X : \Omega \mapsto \mathbf{R}$ definujeme *distribuční funkci* náhodné veličiny X předpisem

$$F_X(x) := P[X \leq x],$$

což je reálná funkce. Protože jsme pravděpodobnost $P(A) : \mathcal{X} \mapsto \langle 0, 1 \rangle$ definovali jako míru, reprezentuje distribuční funkce $F_X(x)$ vlastně vytvořující funkci míry. Pak musí splňovat následující předpoklady:

- je neklesající na \mathbf{R} ,
- $\text{Ran}(F_X) \subset \langle 0, 1 \rangle$,
- $\lim_{x \rightarrow -\infty} F_X(x) = 0$,
- $\lim_{x \rightarrow +\infty} F_X(x) = 1$,
- $F_X(x)$ je spojitá zprava na \mathbf{R} , tj. pro každé $c \in \mathbf{R}$ platí $\lim_{x \rightarrow c^+} F_X(x) = F_X(c)$,
- $F_X(x)$ má nejvýše spočetně mnoho bodů nespojitosti.

Absolutně spojitá náhodná veličina

Mějme pravděpodobnostní prostor $\{\Omega, \mathcal{X}, P\}$ a náhodnou veličinu $X : \Omega \mapsto \mathbf{R}$. Pokud existuje nezáporná funkce $f_X(t) : \mathbf{R} \mapsto \mathbf{R}$ taková, že pro distribuční funkci $F_X(x)$ náhodné veličiny X platí pro všechna $x \in \mathbf{R}$

$$F_X(x) = \int_{-\infty}^x f_X(t) dt, \quad (2.2)$$

řekneme, že náhodná veličina X má *absolutně spojitě rozdělení*.

Pokud tedy pro náhodnou veličinu X existuje funkce $f_X(x)$ z předchozího vztahu, nazýváme tuto funkci *hustota pravděpodobnosti* náhodné veličiny X .

Zvolme si nyní pevný pravděpodobnostní prostor $\{\Omega, \mathcal{X}, P\}$, ze kterého budeme v následujících úvahách vycházet. Mějme dále náhodnou veličinu X s absolutně spojitým rozdělením, jejíž

distribuční funkce je $F_X(x)$ a hustota pravděpodobnosti $f_X(x)$. Z vlastnosti integrálu a derivace pak plyne, že ve všech bodech, kde existuje derivace funkce $F_X(x)$, platí

$$f_X(x) = \frac{dF_X}{dx}(x).$$

Z vlastností distribuční funkce, konkrétně z rovnosti $\lim_{x \rightarrow +\infty} F_X(x) = 1$ a definice hustoty pravděpodobnosti $F_X(x) = \int_{-\infty}^x f_X(t) dt$, plyne pro hustotu pravděpodobnosti tzv. *normalizační podmínka* ve tvaru

$$\int_{\mathbf{R}} f_X(x) dx = 1.$$

Integrál $\int_A f_X(x) dx$ pak reprezentuje pravděpodobnost, že náhodná veličina X nabyde hodnoty z množiny $A \subset \mathbf{R}$, $P[X \in A] = \int_A f_X(x) dx$. Na každou nezápornou funkci $f(x) : \mathbf{R} \mapsto \mathbf{R}$, která splňuje normalizační podmínku, pak můžeme nahlížet jako na hustotu pravděpodobnosti určité jednorozměrné náhodné veličiny. Takovými funkcemi jsme se ovšem zabývali v první kapitole při zavedení pojmu *hustota*. Všechny vztahy odvozené pro hustotu tedy zřejmě musejí platit i pro hustotu pravděpodobnosti jednorozměrné náhodné veličiny (a naopak), neboť jejich vlastnosti jsou identické.

Uvažujme opět náhodnou veličinu X , která má absolutně spojitě rozdělení s hustotou pravděpodobnosti $f_X(x)$. Potom pro každou množinu $A = (a, b)$, kde $a, b \in \mathbf{R}$ a $a \leq b$ (obecně pro každou měřitelnou množinu), platí

$$P[X \in A] = \int_a^b f_X(x) dx.$$

Pokud konverguje integrál

$$\int_{\mathbf{R}} x f_X(x) dx, \quad (2.3)$$

pak jeho hodnota vyjadřuje *střední hodnotu* náhodné veličiny X . Označuje se symbolem $E(X)$. Máme-li dánu náhodnou veličinu X , příslušnou hustotu pravděpodobnosti $f_X(x)$ a její střední hodnotu $E(X)$ a integrál

$$\int_{\mathbf{R}} (x - E(X))^2 f_X(x) dx,$$

konverguje, pak jeho hodnota vyjadřuje *rozptyl* náhodné veličiny X . Značí se symbolem $\text{VAR}(X)$. Konverguje-li také integrál $E(X^2) = \int_{\mathbf{R}} x^2 f_X(x) dx$, platí mezi rozptylem $\text{VAR}(X)$ a střední hodnotou $E(X)$ náhodné veličiny X významný vztah

$$\text{VAR}(X) = E(X^2) - (E(X))^2. \quad (2.4)$$

Směrodatnou odchylku náhodné veličiny X získáme z daného rozptylu $\text{VAR}(X)$ podle vztahu

$$\text{SD}(X) := \sqrt{\text{VAR}(X)}. \quad (2.5)$$

Výčet pojmů zakončíme definicí ℓ -tého momentu μ_ℓ náhodné veličiny X , který je roven hodnotě integrálu

$$\mu_\ell = \int_{\mathbf{R}} x^\ell f_X(x) dx, \quad (2.6)$$

pokud integrál konverguje. Střední hodnotu a rozptyl lze pak vyjádřit pomocí ℓ -tého momentu rovnostmi

$$E(X) = \mu_1, \quad \text{VAR}(X) = \mu_2 - \mu_1^2.$$

Poznámka:

Pomocí Laplaceovy transformace lze první a druhý moment $E(X)$ a $E(X^2)$ vypočítat podle vztahů

$$\mu_1 = E(X) = -\frac{d}{ds} \mathcal{L}[f_X(x)](0), \quad \mu_2 = E(X^2) = \frac{d^2}{ds^2} \mathcal{L}[f_X(x)](0).$$

Rovnosti platí zcela obecně pro jakoukoliv náhodnou veličinu s absolutně spojitým rozdělením, jejíž hustota pravděpodobnosti je omezenou funkcí.

2.2 Absolutně spojitá rozdělení

Exponenciální rozdělení

Náhodná veličina X má *exponenciální rozdělení* s parametry $\mu \in \mathbf{R}, \lambda > 0$, pokud pro její hustotu pravděpodobnosti platí vztah

$$f_X(x) = \Theta(x - \mu) \lambda e^{-\lambda(x-\mu)}.$$

Značí se $X \sim \text{Exp}(\mu, \lambda)$. Střední hodnota a rozptyl takové náhodné veličiny pak mají hodnotu

$$E(X) = \mu + \frac{1}{\lambda}, \quad \text{VAR}(X) = \frac{1}{\lambda^2}.$$

Snadno se přesvědčíme se, že hustota pravděpodobnosti je normalizována správně.

Při volbě parametrů $\mu = 0$ a $\lambda = 1$ platí pro střední hodnotu, rozptyl a hustotu pravděpodobnosti náhodné veličiny X rovnice

$$E(X) = 1, \quad \text{VAR}(X) = 1, \quad f_X(x) = \Theta(x) e^{-x}. \quad (2.7)$$

Gamma rozdělení

Náhodná veličina X má *Gamma rozdělení* s parametry $\alpha, \lambda \in \mathbf{R}$, jestli-že pro její hustotu pravděpodobnosti platí vztah

$$f_X(x) = \Theta(x) \frac{\lambda^{\alpha+1}}{\Gamma(\alpha+1)} x^\alpha e^{-\lambda x}.$$

Značíme $X \sim \text{Gamma}(\alpha, \lambda)$. Střední hodnota a rozptyl takové náhodné veličiny mají hodnotu

$$E(X) = \frac{\alpha+1}{\lambda}, \quad \text{VAR}(X) = \frac{\alpha+1}{\lambda^2}.$$

Opět se přesvědčíme, že hustota pravděpodobnosti je normalizována správně, k čemuž využijeme vztah (1.9) při $s = 0$:

$$G(s) = \mathcal{L} \left[\Theta(x) \frac{\lambda^{\alpha+1}}{\Gamma(\alpha+1)} x^\alpha e^{-\lambda x} \right] = \left(\frac{\lambda}{\lambda+s} \right)^{\alpha+1} \Rightarrow G(0) = 1 \Rightarrow \int_{-\infty}^{+\infty} \Theta(x) \frac{\lambda^{\alpha+1}}{\Gamma(\alpha+1)} x^\alpha e^{-\lambda x} dx = 1.$$

Při volbě parametru $\lambda = \alpha + 1$ platí pro střední hodnotu, rozptyl a hustotu pravděpodobnosti náhodné veličiny X rovnice

$$E(X) = 1, \quad \text{VAR}(X) = \frac{1}{\alpha+1}, \quad f_X(x) = \Theta(x) \frac{(\alpha+1)^{\alpha+1}}{\Gamma(\alpha+1)} x^\alpha e^{-(\alpha+1)x}. \quad (2.8)$$

GIG rozdělení

Náhodná veličina X má *tříparametrické GIG rozdělení* (Generalized Inverse Gaussian distribution), jestliže pro její hustotu pravděpodobnosti platí vztah

$$f_X(x) = \Theta(x) \sqrt{\left(\frac{D}{\beta}\right)^{\alpha+1}} \frac{1}{2 \mathcal{K}_{\alpha+1}[2\sqrt{\beta D}]} x^\alpha e^{-\frac{\beta}{x}} e^{-Dx}, \quad (2.9)$$

kde $\alpha \in \mathbf{R}, \beta > 0$. Symbol $\mathcal{K}_\ell, \ell \in \mathbf{R}$ reprezentuje MacDonaldu funkci řádu ℓ (tj. modifikovanou Besselovu funkci druhého druhu řádu ℓ). Hustota je normalizována správně, výpočet ovšem pro jeho složitost neuvádíme. GIG rozdělení se značí $X \sim GIG(\alpha, \beta, D)$. Laplaceův obraz funkce $f_X(x)$ má podle [26] tvar

$$\mathfrak{L}[f_X(x)] = \sqrt{\left(\frac{D}{D+s}\right)^{\alpha+1}} \frac{\mathcal{K}_{\alpha+2}[2\sqrt{\beta(D+s)}]}{\mathcal{K}_{\alpha+1}[2\sqrt{\beta D}]}.$$

Střední hodnota $E(X)$ je pak dána vztahem

$$E(X) = \sqrt{\frac{\beta}{D}} \frac{\mathcal{K}_{\alpha+2}[2\sqrt{\beta D}]}{\mathcal{K}_{\alpha+1}[2\sqrt{\beta D}]}.$$

Položme $E(X) = 1$ a hledejme závislost D na α, β tak, aby rovnost platila. Podle zdroje [26] dostaneme výraz

$$D(\alpha, \beta) = \alpha + \beta + \frac{3 - g(\beta)}{2}, \quad (2.10)$$

kde $g(\beta)$ reprezentuje korekční funkci aproximativně odhadnutou na $g(\beta) \approx e^{-\sqrt{\beta}}$. Analogicky lze odvodit rozptyl GIG rozdělení, viz [26],

$$\text{VAR}(X) = \frac{\beta \mathcal{K}_{\alpha+3}[2\sqrt{\beta D}]}{D \mathcal{K}_{\alpha+1}[2\sqrt{\beta D}]} - (E(X))^2.$$

Pokud uvažujeme speciální případ, kdy $E(X) = 1$, což bude typická situace v této diplomové práci, vztah pro rozptyl se značně zjednoduší na

$$\text{VAR}(X) = \frac{\alpha + \beta + 2}{D} - 1.$$

Často se využívá pouze *dvouparametrické GIG rozdělení*, kde $\alpha = 0$, tedy

$$f_X(x) = \Theta(x) \sqrt{\frac{D}{\beta}} \frac{1}{2 \mathcal{K}_1[2\sqrt{\beta D}]} e^{-\frac{\beta}{x}} e^{-Dx}. \quad (2.11)$$

V této diplomové práci budou aplikovány oba typy rozdělení, dvou- i tříparametrické.

Konvoluce dvou identických GIG hustot

Nejprve vyslovíme pomocnou větu z [28], bez níž bychom odvození konvoluce dvou GIG hustot nemohli provést.

Věta o hrubém leadingu

Nechť $p(x), r(x) : (a, b) \mapsto \mathbf{R}$ jsou měřitelnými funkcemi na $(a, b) \subset \mathbf{R}$. Nechť $S := \sup_{x \in (a, b)} p(x) < +\infty$. Označme

$$H(t) = \int_a^b r(x) e^{tp(x)} dx.$$

Nechť $\exists t_0 \in \mathbf{R}$ takové, že $r(x) e^{t_0 p(x)} \in \mathcal{L}(a, b)$. Potom

- $r(x) e^{t(p(x)-S)} \in \mathcal{L}(a, b)$
- $\lim_{t \rightarrow +\infty} e^{-St} H(t) = \int_a^b r_0(x) dx$, kde $r_0(x) = \begin{cases} r(x) & p(x) = S, \\ 0 & p(x) \neq S. \end{cases}$

Přímý důsledek této věty je integrální aproximace

$$\int_a^b r(x) e^{tp(x)} dx \doteq \int_a^b r(x) e^{tS} dx,$$

jejíž přesnost se zvyšuje s rostoucím t . Té využijeme v následujícím výpočtu konvoluce dvou GIG hustot. Vyjdeme z tříparametrického GIG rozdělení s hustotou pravděpodobnosti dané vztahem

(2.9). Označme výraz $A = \sqrt{\left(\frac{D}{\beta}\right)^{\alpha+1} \frac{1}{2\mathcal{K}_{\alpha+1}[2\sqrt{\beta D}]}}$. Ten nezávisí na proměnné x a pro daná α, β, D je konstantou. Potom

$$\begin{aligned} (f_X \star f_X)(x) &= \int_{\mathbb{R}} \Theta(s) A s^\alpha e^{-\frac{\beta}{s}} e^{-Ds} \Theta(x-s) A (x-s)^\alpha e^{-\frac{\beta}{x-s}} e^{-D(x-s)} ds = \\ &= \Theta(x) A^2 e^{-Dx} \int_0^x s^\alpha (x-s)^\alpha e^{\beta \frac{-x}{sx-s^2}} ds. \end{aligned}$$

Označíme-li $r(s) = s^\alpha (x-s)^\alpha$, $t = \beta$, $p(s) = \frac{-x}{sx-s^2}$ na intervalu $(a, b) = (0, x)$, získáme integrál z předchozí věty a můžeme použít odpovídající aproximaci

$$\int_0^x r(s) e^{\beta p(s)} ds \doteq \int_0^x r(s) e^{\beta S} ds.$$

Body, v nichž funkce $p(s)$ může mít extrém, najdeme z první derivace $p(s)$, z druhé derivaci ověříme, že jde o maximum:

$$\frac{dp(s)}{ds} = \frac{x(x-2s)}{(sx-s^2)^2} = 0 \Rightarrow x \neq s \neq 0, s_0 = \frac{x}{2},$$

$$\frac{d^2p(s)}{ds^2} = \frac{2s}{(sx-s^2)^2} + \frac{4s-2x}{(sx-s^2)^3} \Rightarrow \frac{d^2p(s)}{ds^2} \left(\frac{x}{2}\right) = -\frac{16}{x^3} < 0.$$

Funkce $p(s)$ je na konečném intervalu $(0, x)$ spojitá, nalezené maximum v bodě $s_0 = \frac{x}{2}$ je tedy zároveň rovno supremu $S = \sup_{s \in (0, x)} p(s) = p\left(\frac{x}{2}\right) = -\frac{4}{x}$. Dostaneme aproximaci

$$(f_X \star f_X)(x) = \Theta(x) A^2 e^{-Dx} \int_0^x r(s) e^{\beta S} ds = A^2 e^{-Dx} e^{-\beta \frac{4}{x}} \Theta(x) \int_0^x s^\alpha (x-s)^\alpha ds.$$

Výraz $\Theta(x) \int_0^x s^\alpha (x-s)^\alpha ds$ představuje konvoluci $\Theta(x)x^\alpha \star \Theta(x)x^\alpha$, na kterou s výhodou aplikujeme Laplaceovu transformaci

$$\begin{aligned} \mathfrak{L}[(\Theta(x)x^\alpha \star \Theta(x)x^\alpha)] &= \mathfrak{L}[\Theta(x)x^\alpha] \cdot \mathfrak{L}[\Theta(x)x^\alpha] = \left(\frac{\Gamma(\alpha+1)}{s^{\alpha+1}}\right)^2 = \frac{\Gamma(\alpha+1)\Gamma(\alpha+1)}{\Gamma(2\alpha+2)} \cdot \frac{\Gamma(2\alpha+2)}{s^{2\alpha+2}} = \\ &= \frac{\Gamma(\alpha+1)\Gamma(\alpha+1)}{\Gamma(2\alpha+2)} \mathfrak{L}[\Theta(x)x^{2\alpha+1}] = \mathfrak{L}\left[\frac{\Gamma(\alpha+1)\Gamma(\alpha+1)}{\Gamma(2\alpha+2)} \Theta(x)x^{2\alpha+1}\right]. \end{aligned}$$

Rovnájí-li se Laplaceovy obrazy funkcí, musejí se podle Lerchova teorému [23] rovnat i funkce samotné, tedy

$$\Theta(x) \int_0^x s^\alpha (x-s)^\alpha ds = (\Theta(x)x^\alpha \star \Theta(x)x^\alpha) = \frac{\Gamma(\alpha+1)\Gamma(\alpha+1)}{\Gamma(2\alpha+2)} \Theta(x)x^{2\alpha+1}.$$

Pro konvoluci dvou hustot GIG rozdělení pak získáváme vztah

$$(f_X \star f_X)(x) = g_X(x) = A^2 \frac{\Gamma(\alpha+1)\Gamma(\alpha+1)}{\Gamma(2\alpha+2)} \Theta(x)x^{2\alpha+1} e^{-\frac{4\beta}{x}} e^{-Dx}.$$

Funkce $g_X(x)$ má zjevně tvar hustoty pravděpodobnosti GIG rozdělení s parametry $2\alpha+1$, 4β , D . Výraz $A^2 \frac{\Gamma(\alpha+1)\Gamma(\alpha+1)}{\Gamma(2\alpha+2)}$ nezávisí na proměnné x a vyjadřuje jistou normalizační konstantu. Jak již zaznělo v první kapitole, konvolucí dvou hustot vznikne opět hustota, která je správně normovaná. V tomto případě jsme však použili aproximaci, takže výše zmíněná konstanta je nevyhovující a je nutné výslednou funkci $g_X(x)$ normalizovat znovu. Napíšeme ji ve tvaru

$$g_X(x) = B \Theta(x)x^{2\alpha+1} e^{-\frac{4\beta}{x}} e^{-Dx}$$

a budeme hledat normalizační konstantu B . Víme, že normalizační konstanta funkce $f_X(x)$ s parametry γ, δ, D je rovna $A = \sqrt{\left(\frac{D}{\delta}\right)^{\gamma+1} \frac{1}{2\mathcal{K}_{\gamma+1}[2\sqrt{\delta D}]}}$, odkud při volbě $\gamma = 2\alpha+1, \delta = 4\beta$ získáme $B = \left(\frac{D}{4\beta}\right)^{\alpha+1} \frac{1}{2\mathcal{K}_{2\alpha+2}[4\sqrt{\beta D}]}$. Výsledný vztah pro konvoluci identických hustot pravděpodobnosti tří-parametrického GIG rozdělení je tedy

$$(f_X \star f_X)(x) = \left(\frac{D}{4\beta}\right)^{\alpha+1} \frac{1}{2\mathcal{K}_{2\alpha+2}[4\sqrt{\beta D}]} \Theta(x)x^{2\alpha+1} e^{-\frac{4\beta}{x}} e^{-Dx}. \quad (2.12)$$

Konvoluce dvou odlišných GIG hustot

Předpokládejme nyní dvě hustoty GIG rozdělení $f_X(x)$ a $g_X(x)$ s parametry (α_1, β_1, D) a (α_2, β_2, D) . Odvození konvoluce $(f_X \star g_X)(x)$ provedeme analogicky, zmíníme zde jen dílčí kroky. Označíme-li konstanty příslušné první a druhé trojici parametrů A , resp. B , lze upravit konvoluci na tvar

$$(f_X \star g_X)(x) = \Theta(x) A B e^{-Dx} \int_0^x s^{\alpha_1} (x-s)^{\alpha_2} e^{\beta_1 \frac{(1-\frac{\beta_2}{\beta_1})s-x}{sx-s^2}} ds.$$

Označíme $r(s) = s^{\alpha_1} (x-s)^{\alpha_2}$, $t = \beta_1$, $p(s) = \frac{(1-\frac{\beta_2}{\beta_1})s-x}{sx-s^2}$. Stacionární bod s_0 určíme z první derivace $p(s)$

$$\frac{dp(s)}{ds} = \frac{1}{s^2} - \frac{\beta_2}{\beta_1(x-s)^2} = 0 \Rightarrow x \neq s \neq 0, s_0 = \frac{x\sqrt{\beta_1}}{\sqrt{\beta_1} + \sqrt{\beta_2}}.$$

Funkce $p(s)$ je na $(0, x)$ spojitě diferencovatelná, má jediný stacionární bod a $\lim_{s \rightarrow 0^+} p(s) = \lim_{s \rightarrow x^-} p(s) = -\infty$. Bod s_0 je tedy zřejmě maximum a zároveň supremum a $p(s_0) = -\frac{(\sqrt{\beta_1} + \sqrt{\beta_2})^2}{x\beta_1}$. Dostaneme aproximaci

$$(f_X \star g_X)(x) = \Theta(x) A B e^{-Dx} e^{-\frac{(\sqrt{\beta_1} + \sqrt{\beta_2})^2}{x}} \int_0^x s^{\alpha_1} (x-s)^{\alpha_2} ds.$$

Výraz $\Theta(x) \int_0^x s^{\alpha_1} (x-s)^{\alpha_2} ds$ upravíme pomocí Laplaceovy transformace

$$\mathcal{L} \left[\int_0^x s^{\alpha_1} (x-s)^{\alpha_2} ds \right] = \mathcal{L} [\Theta(x)x^{\alpha_1} \star \Theta(x)x^{\alpha_2}] = \mathcal{L} \left[\frac{\Gamma(\alpha_1+1)\Gamma(\alpha_2+1)}{\Gamma(\alpha_1+\alpha_2+2)} \Theta(x)x^{\alpha_1+\alpha_2+1} \right].$$

Na závěr musíme analogicky předchozímu odvození modifikovat konstantu $A B \frac{\Gamma(\alpha_1+1)\Gamma(\alpha_2+1)}{\Gamma(\alpha_1+\alpha_2+2)}$ na vhodnou konstantu C tak, aby konvoluce byla správně normována. Pro konvoluci tak dostáváme výsledný vztah

$$(f_X \star g_X)(x) = C \Theta(x) x^{\alpha_1+\alpha_2+1} e^{-\frac{(\sqrt{\beta_1}+\sqrt{\beta_2})^2}{x}} e^{-Dx}. \quad (2.13)$$

Kvalita aproximace byla testována v článku [17].

2.3 Součet náhodných veličin

Mějme dány dvě nezávislé jednorozměrné náhodné veličiny X, Y s absolutně spojitým rozdělením a jejich hustoty pravděpodobnosti $f_X(x)$ a $f_Y(y)$. Pak hustotou pravděpodobnosti $f_Z(z)$ náhodné veličiny $Z = X + Y$ je funkce

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z-x) dx.$$

Podívejme se blíže na odvození této rovnosti. Označme $F_Z(z)$ distribuční funkci náhodné veličiny $Z = X + Y$. Pro ni podle definice distribuční funkce náhodné veličiny s absolutně spojitým rozdělením platí

$$F_Z(z) = P[Z \leq z] = P[X + Y \leq z] = \int \int_{x+y \leq z} f_{X,Y}(x, y) dx dy,$$

kde $f_{X,Y}(x, y)$ představuje sdruženou hustotu pravděpodobnosti náhodných veličin X, Y . Ty jsou nezávislé, takže $f_{X,Y}(x, y) = f_X(x) \cdot f_Y(y)$ a pro distribuční funkci dostaneme

$$\begin{aligned} F_Z(z) &= \int \int_{x+y \leq z} f_X(x) f_Y(y) dx dy = \int_{-\infty}^{\infty} \int_{-\infty}^{z-x} f_X(x) f_Y(y) dy dx = \left| \begin{array}{l} r = x + y \\ dr = dy \end{array} \right| = \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^z f_X(x) f_Y(r-x) dr dx = \int_{-\infty}^z \left(\int_{-\infty}^{\infty} f_X(x) f_Y(r-x) dx \right) dr = \int_{-\infty}^z f_Z(r) dr. \end{aligned}$$

Hledanou hustotou pravděpodobnosti je (po přeznačení) opravdu funkce

$f_Z(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z-x) dx$. Porovnáme-li vztah s definicí konvoluce, zjistíme, že jsou identické. Pro hustotu pravděpodobnosti $f_Z(z)$ náhodné veličiny $Z = X + Y$ tedy platí

$$f_Z(z) = (f_X \star f_Y)(z). \quad (2.14)$$

Protože $f_X(x), f_Y(y)$ mají všechny vlastnosti hustoty a patří tak do $\mathcal{L}_1(\mathbf{R})$, $f_Z(z)$ musí existovat a patří také do $\mathcal{L}_1(\mathbf{R})$. Platí pro ně všechny vztahy odvozené v první kapitole.

Střední hodnota součtu náhodných veličin

Uvažujme dvě náhodné veličiny X, Y s absolutně spojitým rozdělením a jejich hustoty pravděpodobnosti $f_X(x), f_Y(y)$. Označme střední hodnoty $E(X) = \mu, E(Y) = \nu$. Potom hustotou pravděpodobnosti náhodné veličiny $Z = X + Y$ je zřejmě funkce $h_Z(z) = (f_X \star f_Y)(z)$ a platí

$$E(Z) = E(X) + E(Y).$$

Vezmeme-li v úvahu definici střední hodnoty (2.3), je toto tvrzení splněno na základě věty, kterou jsme dokázali v první kapitole.

Rozptyl součtu náhodných veličin

Nechť má náhodná veličina X rozptyl roven $\text{VAR}(X) = \sigma_x^2$ a náhodná veličina Y rozptyl $\text{VAR}(Y) = \sigma_y^2$. Nechť X, Y jsou nezávislé. Pak pro rozptyl náhodné veličiny $Z = X + Y$ platí

$$\text{VAR}(Z) = \sigma_x^2 + \sigma_y^2.$$

2.4 Metody odhadu parametrů

Nyní se budeme krátce zabývat statistickými metodami, pomocí nichž lze najít optimální parametry matematického modelu. Uvažujme náhodný výběr X_1, X_2, \dots, X_n (posloupnost n nezávislých stejně rozdělených náhodných veličin) z rozdělení, jež závisí na parametru $\vec{\theta}$. Cílem odhadu je určení tohoto parametru (obecně vektorového) na základě informace obsažené ve výběrovém souboru. Hledáme *bodový odhad*, kdy neznámý parametr $\vec{\theta}$ charakterizujeme jedinou hodnotou. Nalézt bodový odhad znamená nalézt výběrovou charakteristiku $\vec{T} = \vec{T}(X_1, X_2, \dots, X_n)$, tj. funkci náhodných veličin X_1, X_2, \dots, X_n , která je blízko skutečné hodnotě $\vec{\theta}$. Existují některé výhodné vlastnosti odhadů, např. konzistence či nestrannost. My na odhad \vec{T} klademe dva požadavky:

- Chceme, aby odhad \vec{T} byl *nestranný*, tedy $E(\vec{T}) = \vec{\theta}$ pro každé $\vec{\theta}$.
- Chceme, aby odhad \vec{T} byl *nejlepší*. Máme-li k dispozici více nestranných odhadů, nazveme odhad \vec{T} nejlepším, pokud má nejmenší rozptyl.

Tyto poznatky byly čerpány z [24]. Nyní uvedeme dvě obecné statistické metody k nalezení odhadu.

Metoda maximální věrohodnosti (Maximum Likelihood Estimation)

Předpokládejme náhodný výběr z rozdělení o hustotě pravděpodobnosti $f(x; \vec{\theta})$. Sdružená hustota pro soubor stejně rozdělených nezávislých náhodných veličin je podle [24] dána vztahem

$$f(x_1, x_2, \dots, x_n; \vec{\theta}) = \prod_{i=1}^n f(x_i; \vec{\theta}). \quad (2.15)$$

Nahlížíme-li na sdruženou hustotu jako na funkci $\vec{\theta}$, dostaneme *věrohodnostní funkci*. Obvykle se používá logaritmická věrohodnostní funkce $l(\vec{\theta})$, která převede součin na součet

$$l(\vec{\theta}) = \sum_{i=1}^n \ln f(x_i; \vec{\theta}). \quad (2.16)$$

Pro $f(x) = 0$ se definuje $l(\vec{\theta}) = -\infty$. Hledaný odhad je pak taková hodnota parametru $\hat{\vec{\theta}}$, která na množině všech možných hodnot parametru $\vec{\theta}$ maximalizuje věrohodnostní (resp. logaritmickou věrohodnostní) funkci. Parametry získané touto metodou budeme nadále označovat dolním indexem *MLE*.

Metoda minimální vzdálenosti (Minimum Distance Estimation)

Tato statistická metoda vychází podle [25] z náhodného výběru z rozdělení o distribuční funkci $F(x; \vec{\theta})$, kde $\theta \in \Theta$. Symbolem $F_E(x)$ označíme empirickou distribuční funkci, která popisuje empirická data a je dána vztahem (2.2). Pokud $\hat{\vec{\theta}}$ je odhad parametru $\vec{\theta}$, pak $F(x; \hat{\vec{\theta}})$ je odhad distribuční funkce $F(x; \vec{\theta})$. Nyní zavedeme tzv. *distanční funkcionál* $d[\cdot, \cdot]$, který vrací hodnotu obecné vzdálenosti (metriky) vstupujících funkcí. Následně hledáme v parametrické množině Θ takové $\hat{\vec{\theta}}$, pro které je vzdálenost empirické distribuční funkce a odhadu distribuční funkce na množině Θ minimální, tedy platí

$$d[F(x; \hat{\vec{\theta}}), F_E(x)] = \inf \{ d[F(x; \vec{\theta}), F_E(x)] : \vec{\theta} \in \Theta \}. \quad (2.17)$$

Funkcionální metriku, tj. distanční funkcionál lze volit různě. V této práci budeme pracovat především s tzv. *kolmogorovskou vzdáleností*, kterou lze vyjádřit rovnicí

$$d[F(x; \vec{\theta}), F_E(x)] = \sup_{x \in \mathbf{R}} |F(x; \vec{\theta}) - F_E(x)|. \quad (2.18)$$

Parametry získané touto metodou budeme nadále označovat dolním indexem $_{KD}$ (Kolmogorov Distance).

Metodu minimální vzdálenosti lze analogicky aplikovat také na příslušné hustoty pravděpodobnosti $f(x, \hat{\theta})$ a $f_E(x)$. Jako typické zástupce funkcionálních metrik obvykle aplikovaných na hustoty pravděpodobnosti uvádíme

- L_1 - metriku zavedenou pomocí L_1 - normy $\|h(x)\|_1 := \int_{\mathbf{R}} |h(x)| dx$ předpisem

$$d[f(x), g(x)] := \|f(x) - g(x)\|_1 = \int_{\mathbf{R}} |f(x) - g(x)| dx, \quad (2.19)$$

popřípadě

- L_2 - metriku zavedenou pomocí L_2 - normy $\|h(x)\|_2 := \sqrt{\int_{\mathbf{R}} |h(x)|^2 dx}$ předpisem

$$d[f(x), g(x)] := \|f(x) - g(x)\|_2 = \sqrt{\int_{\mathbf{R}} |f(x) - g(x)|^2 dx}. \quad (2.20)$$

Kapitola 3

Základní pojmy teorie dopravního proudění

V této kapitole budou představeny základní definice popisující mikroskopickou a makroskopickou strukturu dopravy. Níže uvedené vztahy jsou přejaty z [17]. Nastíníme charakteristiky datových souborů, jež budeme zpracovávat.

3.1 Mikroskopické znaky dopravy

Dopravní mikroskopické veličiny charakterizují jednotlivá vozidla. Lze je rozdělit na primární (přímo měřené detektorem) a sekundární (odvozené z primárních).

3.1.1 Primární veličiny

V daném jízdním pruhu mějme N vozidel. Pak množina

$$T^{(in)} = \{\tau_k^{(in)} \in \mathbf{R}_0^+ \mid k = 1, 2, \dots, N\} \quad (3.1)$$

je soubor chronologicky uspořádaných časových údajů, které popisují, kdy přední nárazník k -tého vozidla protnul linii detektoru. Obdobně se zavádí soubor časových údajů, kdy linii detektoru protnul zadní nárazník k -tého vozidla

$$T^{(out)} = \{\tau_k^{(out)} \in \mathbf{R}_0^+ \mid k = 1, 2, \dots, N\}. \quad (3.2)$$

Množina rychlostí jednotlivých vozidel zaznamenaných detektorem tvoří soubor

$$V = \{v_k \in \mathbf{R}^+ \mid k = 1, 2, \dots, N\}. \quad (3.3)$$

Uvedené veličiny jsou přímo měřitelné detektorem, a tudíž nejsou ovlivněny žádnou systematickou chybou. Analogicky lze tyto soubory zavést pro libovolný jízdní pruh.

3.1.2 Sekundární veličiny

V následujícím přehledu uvedeme některé významné sekundární veličiny. Jsou to

- *časový odstup* popisující časový interval mezi detekcí předního nárazníku vedoucího vozidla a detekcí předního nárazníku následujícího vozidla

$$z_k = \tau_k^{(in)} - \tau_{k-1}^{(in)}, \quad (3.4)$$

- *časová světlost* popisující časový interval mezi detekcí zadního nárazníku vedoucího vozidla a detekcí předního nárazníku následujícího vozidla

$$t_k = \tau_k^{(in)} - \tau_{k-1}^{(out)}, \quad (3.5)$$

- *délka vozidla* aproximovaná vztahem

$$d_k = v_k(\tau_k^{(in)} - \tau_k^{(out)}), \quad (3.6)$$

- *prostorový odstup* předních nárazníků dvou následujících vozidel

$$s_k = v_k z_k, \quad (3.7)$$

- *prostorová světlost* mezi zadním nárazníkem vedoucího vozidla a předním nárazníkem následujícího vozidla

$$r_k = v_k t_k. \quad (3.8)$$

Časové světlosti nejsou, jak je zřejmé, zatíženy žádnou systematickou chybou. Při odvození ostatních veličin předpokládáme, že rychlost $v(\tau) = v_k$ je po dobu časového intervalu $\tau \in [\tau_{(k-1)}^{(in)}, \tau_{(k)}^{(in)}]$ konstantní. Tato podmínka je však diskutabilní, především v oblastech s velkými časovými odstupy (tj. v oblastech s malou hustotou provozu). V případě veličin charakterizujících prostorový rozestup musíme tedy uvažovat i výpočetní chybu.

3.2 Makroskopický popis dopravy

Nejprve zavedeme velikost vzorkování m . V této práci budeme nadále uvažovat $m = 50$. Pro celkový počet vzorků platí

$$M = \left\lfloor \frac{N}{m} \right\rfloor,$$

kde $N = \text{card}\{\tau_k^{(in)}\}$ a symbol $\text{card}(A)$ reprezentuje tzv. kardinální číslo množiny A .

Datový vzorek definujeme vztahem

$$S_j = \{(\tau_k^{(in)}, \tau_k^{(out)}, v_k, d_k) \in T^{(in)} \times T^{(out)} \times V \times D \mid k = (j-1)m + 1, (j-1)m + 2, \dots, jm\},$$

pro $j = 1, 2, \dots, M$. Pro každý datový vzorek pak určíme makroskopické veličiny

- *lokální tok*

$$J_j = \frac{m}{\tau_{jm}^{(out)} - \tau_{(j-1)m+1}^{(in)}}, \quad (3.9)$$

- *průměrná rychlost*

$$\bar{v}_j = \frac{1}{m} \sum_{k=(j-1)m+1}^{jm} v_k,$$

- *hustota* plynoucí z hydrodynamické rovnice

$$\rho_j = \frac{J_j}{\bar{v}_j}, \quad (3.10)$$

- *průměrná časová světlost*

$$\bar{t}_j = \frac{1}{m} \sum_{k=(j-1)m+1}^{jm} t_k, \quad (3.11)$$

- *průměrná prostorová světlost*

$$\bar{r}_j = \frac{1}{m} \sum_{k=(j-1)m+1}^{jm} r_k.$$

Pomocí \bar{t}_j a \bar{r}_j lze pro všechna $k \in \{(j-1)m+1, (j-1)m+2, \dots, jm\}$ definovat *škálované časové světlosti*

$$x_k = \frac{t_k}{\bar{t}_j}, \quad (3.12)$$

a *škálované prostorové světlosti*

$$l_k = \frac{r_k}{\bar{r}_j}.$$

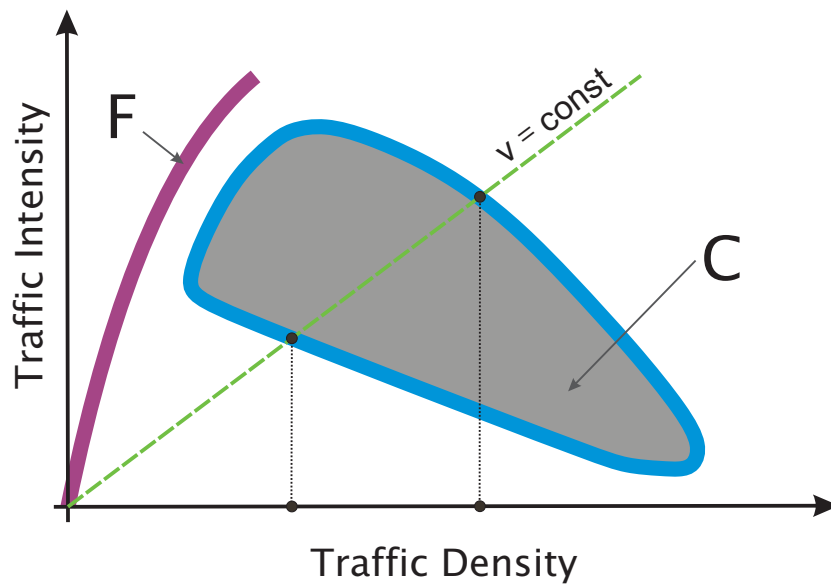
3.3 Fundamentální diagram

K vizualizaci dopravních dat slouží různé typy diagramů, např. vztah mezi rychlostí a hustotou nebo mezi průměrným odstupem a hustotou. Významným zdrojem informací je relace mezi dopravním tokem a hustotou $(\rho, J(\rho))$. Pro ideální dopravní plyn se tato mapa nazývá *fundamentální diagram*. Jeho schématické znázornění nalezneme na obrázku 3.1. Obecně pro malé hustoty splňuje relace $(\rho, J(\rho))$ definici funkce a můžeme hovořit o závislosti toku na hustotě $J = J(\rho)$. Ta je prakticky lineární. U vyšších hustot nelze pojem závislosti použít. Dopravní tok nabývá jediné maximum, a to v oblasti průměrných hodnot hustoty. Z grafu lze vyčíst mj. aktuální průměrnou rychlost, která je definována jako úhel stoupání přímky, která protíná body $(0, 0)$ a $(\rho, J(\rho))$.

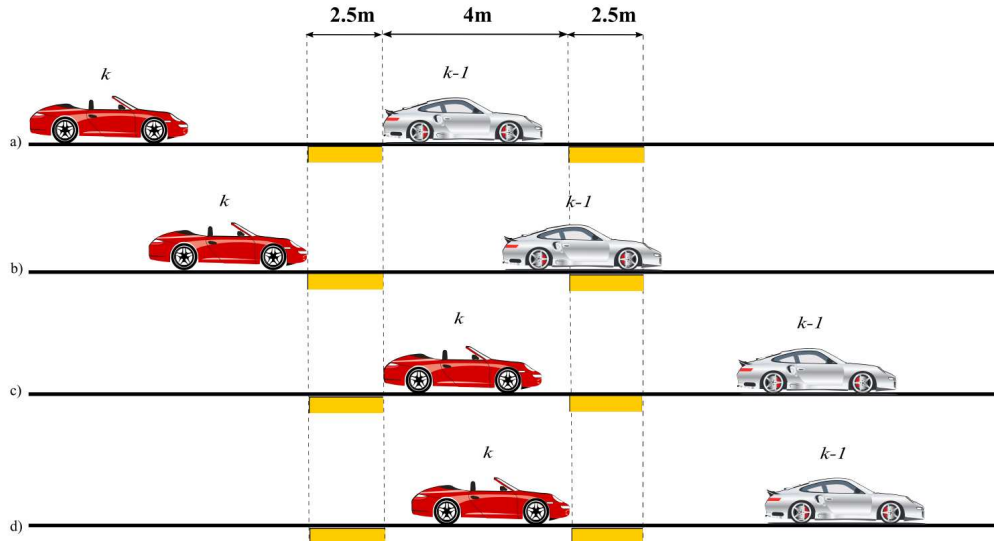
Charakteristický tvar diagramu lze rozdělit na dvě větve (tzv. fáze dopravy). Levá část (ve schématu označena písmenem F) představuje *volnou dopravu*, pravá část (označena písmenem C) odpovídá *kondenzované dopravě*. U volné fáze empirická data víceméně kopírují přímku závislosti $J = J(\rho)$, data příslušná kondenzované fázi jsou v pravé části rozprostřena do dvoudimenzionální oblasti. Kondenzovanou dopravu lze dále členit na synchronizované proudění a rozprostřenou pohybující se kolonu vozidel, rozdíl mezi nimi však již není tak zjevný.

3.4 Detekce dopravních dat

Existuje několik metod sběru dopravních dat, např. letecké snímkování určitého úseku vozovky nebo využití GPS přijímačů. Datové soubory použité v této práci byly získány pomocí dvousmyčkových indukčních detektorů, které se umísťují pod vozovku, viz obrázek 3.2. Fungují na principu změny efektivního napětí LC obvodu - průjezd vozidla způsobí jeho pokles. Ze změny elektromagnetického pole dvousmyčky lze pak vyvodit požadované dopravní mikroveličiny. Detektory dokáží přímo měřit čas protnutí vstupní linie detektoru, čas protnutí výstupní linie detektoru (opuštění detektoru), rychlost vozidel a jejich délku.



Obrázek 3.1: Schématické znázornění fundamentálního diagramu



Obrázek 3.2: Schématické znázornění typických výstupů z měření dopravního proudu prováděných za pomoci magnetické indukční dvousmyčky. (a) $\tau_{k-1}^{(a)}$ čas, kdy zadní nárazník $(k - 1)$ -ho vozidla přetnul výstupní hranu 1. detektoru. (b) $\tau_k^{(b)}$ čas, kdy přední nárazník k -tého vozidla přetnul vstupní hranu 1. detektoru. (c) $\tau_k^{(c)}$ čas, kdy zadní nárazník k -tého vozidla přetnul výstupní hranu 1. detektoru. (d) $\tau_k^{(d)}$ čas, kdy přední nárazník k -tého vozidla přetnul vstupní hranu 2. detektoru. Převzato z [27].

Kapitola 4

Zobecněné inverzní Gaussovo rozdělení GIG a odhadování jeho parametrů

4.1 Dvouparametrická třída GIG

Dopravní mikrostruktura má zjevně stochastický charakter, přičemž stochasticita systému je způsobena zejména individualitou řidiče. Na dopravní mikroskopické veličiny můžeme nahlížet jako na realizace náhodných veličin a přejít k pravděpodobnostnímu popisu. Časová světlost je náhodnou veličinou s absolutně spojitým rozdělením, kterou tedy lze podle [24] popsat pomocí distribuční funkce $F(x)$, resp. hustoty pravděpodobnosti $p(x)$. Hustoty pravděpodobnosti pro časovou světlost sousedních vozidel (agentů) se v teorii dopravy označují jako *clearance distribuce* nebo *headway distribuce*. Z literatury (např. [17, 5, 6, 7]) je známo, že hustotu pravděpodobnosti pro časovou (či prostorovou) světlost nejlépe popisuje rozdělení GIG definované vztahem (2.11). Označme ji $p_{GIG}(x)$. Důvody, proč vyhovuje lépe než jiné matematické modely, jsou tyto:

- $p_{GIG}(x)$ splňuje matematické předpoklady, neboť její chvost má exponenciální průběh, což je pro systém tohoto typu nutná podmínka.
- $p_{GIG}(x)$ splňuje empirické kritérium, které je zřetelné na všech vzorcích dopravních dat, a sice že má plato v $x = 0$. Tato vlastnost se dá matematicky vyjádřit rovnicí

$$\forall m \in \mathbf{N} : \lim_{x \rightarrow 0^+} x^{-m} p_{GIG}(x) = 0,$$

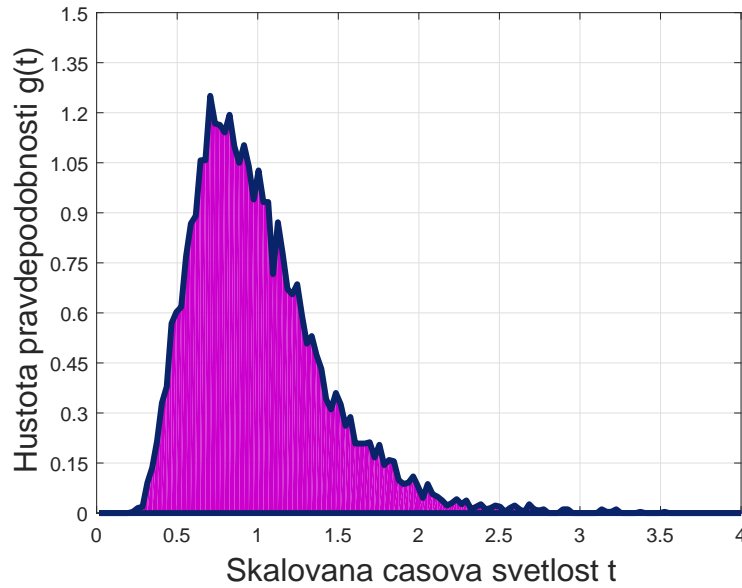
kde proměnná x vyjadřuje časovou světlost. Dá se dokázat, že tento vztah je ve třídě hladkých funkcí ekvivalentní rovnici

$$\forall m \in \mathbf{N}_0 : p_{GIG}^{(m)}(0) = 0,$$

kde $p_{GIG}^{(m)}$ značí m -tou derivaci. Názorná ukázka typické headway distribuce je na obrázku 4.1, kde je plato velmi zřetelné.

- GIG distribuce je rozdělení odvozené analyticky z termodynamického částicového plynu s hyperbolickým repulzním potenciálem, tj. $\varphi(r) = \frac{1}{r}$, kde r označuje vzdálenost. Kromě $p_{GIG}(x)$ neexistuje žádný jiný model, který je odvozen analyticky a zároveň odpovídá realitě (tzn. generuje rozdělení světlostí, které odpovídá naměřeným vzorkům).

Vzhledem k výše uvedenému jsme se zaměřili v další části na distribuční model GIG.



Obrázek 4.1: Ilustrativní zástupce empirických headway distribucí

4.2 Metodika odhadování parametrů distribuce GIG

Nejprve jsme sestavili program pro odhad parametrů dvouparametrické distribuce GIG. Tento program provádí v první fázi pro vstupní soubor dat odhad parametrů (β, D) metodou maximální věrohodnosti (dále jen metodou MLE). Sestavili jsme logaritmickou věrohodnostní funkci, která má podle vztahu (2.16) tvar

$$l(\beta, D) = \sum_{k=1}^n \left(\ln A(\beta, D) - \beta \frac{1}{x_k} - Dx_k \right), \quad (4.1)$$

kde n vyjadřuje počet světlostí v daném souboru a konstanta $A(\beta, D) = \sqrt{\frac{D}{\beta}} \frac{1}{2\mathcal{K}_1[2\sqrt{\beta D}]}$. Parametry (β, D) , které tuto funkci maximalizují, jsou hledaným bodovým odhadem parametrů $(\hat{\beta}, \hat{D})$.

4.3 Testování generátoru pseudonáhodných čísel z dvouparametrické třídy GIG

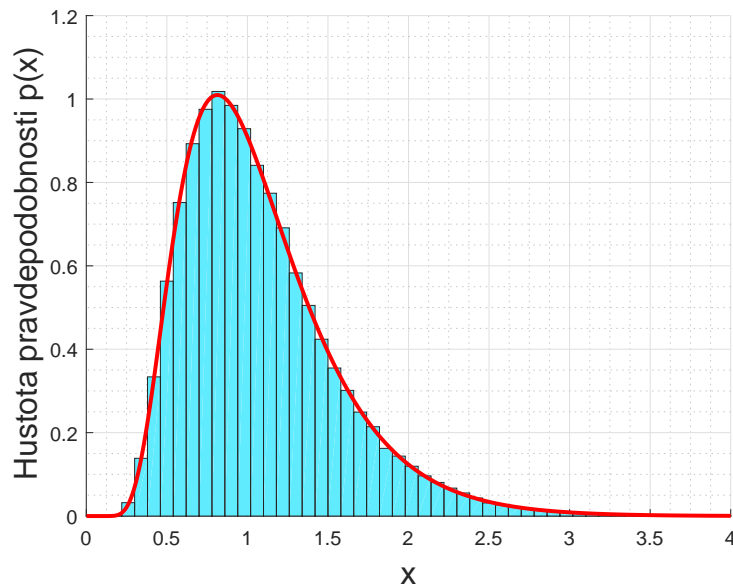
Ve druhé fázi je pro vstupní soubor dat vypočtena distribuční funkce (dále jen CDF - Cumulative Distribution Function) a to dvěma způsoby. Distribuční funkci spočtenou numerickou integrací z odhadnutých parametrů $(\hat{\alpha}, \hat{\beta}, \hat{D})$ označme CDF_T (teoretická). Funkce CDF_E (empirická) je vypočtena přímo ze vstupních dat.

V poslední fázi na základě vypočtených distribučních funkcí CDF_T a CDF_E testujeme na hladině významnosti $h = 0,05$ nulovou hypotézu, že vstupní soubor dat má rozdělení GIG. Ověřování provádíme třemi různými způsoby. Aplikujeme:

- jednovýběrový Kolmogorovův–Smirnovův test dobré shody (dále jen KS test),
- χ^2_{k-1} test dobré shody, přičemž pozorovací prostor byl vždy rozdělen na $k = 10$ buněk, tak aby v každé buňce byl stejný počet dat. Počet stupňů volnosti je roven $k - 1$.

- χ^2_{k-p-1} test dobré shody, kde zohledníme počet odhadovaných parametrů $p = 3$, takže počet stupňů volnosti je roven $k - p - 1$.

K provedení zmíněných testů je v programu MATLAB využíváno právě distribučních funkcí. Graficky můžeme porovnat však i příslušné hustoty pravděpodobnosti (dále jen PDF - Probability Density Function). Srovnání teoretické PDF_T a empirické PDF_E hustoty pravděpodobnosti je na následujícím obrázku 4.2.



Obrázek 4.2: Odhadovaná a empirická hustota pravděpodobnosti pro $N = 100\,000$, $\beta = 2$, $D = 3$

Ověření správné funkčnosti generátoru lze provést, pokud jako vstupní soubor dat použijeme soubory generované dosud neproověřeným generátorem pseudonáhodných čísel, [29], jemuž zadáme pevné parametry. Abychom mohli výsledek kvantifikovat, postupovali jsme následovně:

- Zvolili jsme pevné parametry $\beta = 2$, $D = 3$. Vygenerovali jsme soubor pseudonáhodných dat z rozdělení $GIG(2, 3)$ o počtu prvků $N = 1000$.
- Pro tento soubor jsme odhadli $(\hat{\beta}, \hat{D})$ a určili CDF_T a CDF_E . Pak jsme provedli výše uvedené statistické testy.
- Tento cyklus jsme provedli 1000–krát a pro jednotlivé testy spočítali, v kolika cyklech byla nulová hypotéza zamítnuta.

Celkem jsme otestovali soubory o 8 různých velikostech v rozsahu $N = 1000$ až $N = 1000\,000$. Souhrn výsledků uvádí tabulka 4.1.

Z uvedeného přehledu plyne, že s rostoucí velikostí vzorku povolna vzrůstá počet případů, kdy je nulová hypotéza zamítnuta. Dále je zřejmé, že pokud u χ^2 testu zohledníme počet odhadovaných parametrů, bude hypotéza zamítnuta častěji.

Závěrem lze říci, že, pokud se omezíme na velikosti vzorků podobné těm zde analyzovaným, má testovaný generátor pseudonáhodných čísel kvalitu srovnatelnou s běžně dostupnými standardními generátory (jakými jsou např. `rand` v prostředí MATLAB).

velikost vzorku N	počet cyklů	zamítnutí KS	zamítnutí χ^2_{k-p-1}	zamítnutí χ^2_{k-1}
1 000	1 000	0	58	11
10 000	1 000	0	92	12
50 000	1 000	0	92	19
100 000	1 000	0	80	18
200 000	1 000	0	86	15
500 000	1 000	4	117	24
700 000	1 000	9	119	24
1 000 000	1 000	15	157	39

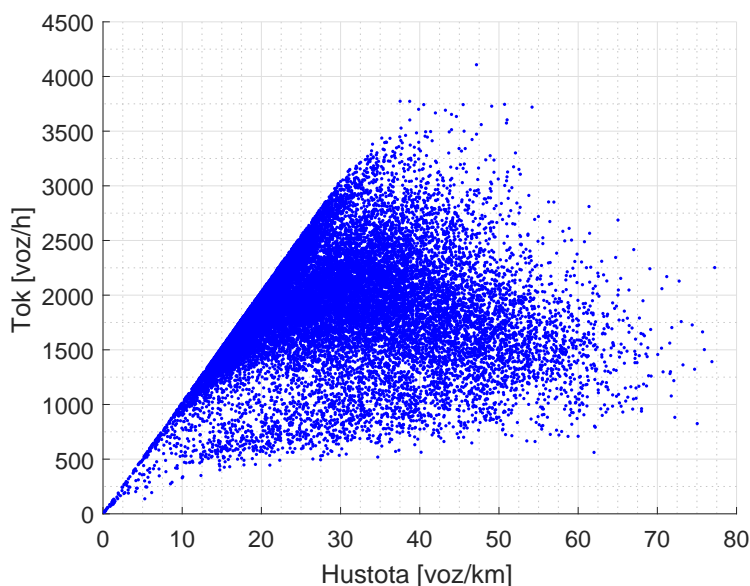
Tabulka 4.1: Počet zamítnutí testem KS, χ^2_{k-p-1} a χ^2_{k-1} pro různé velikosti testovaného souboru

Kapitola 5

Vyhodnocení empirických dat – model pro rozdělení světlostí

5.1 Charakteristika zkoumaného datového souboru

Data, s kterými jsme pracovali, pocházela z datových souborů naměřených na Pražském okruhu v pomalém i rychlém pruhu. Podle práce [27] se v rychlejším pruhu vyskytovalo zhruba 6% nákladních vozidel, v pomalejším pruhu přes 30%. Vyšší počet nákladních vozidel výrazně narušuje statistickou analýzu časových světlostí, protože chování řidiče se v případech, kdy následuje nákladní automobil a osobní automobil, podstatně liší. Nižší průměrná rychlost vozidel v pomalém pruhu má navíc za důsledek slabší interakce mezi jednotlivými vozidly. Uvažovali jsme tedy pouze údaje z levého (rychlejšího) pruhu. Nepracovali jsme s původními soubory naměřenými detektorem, ale využili jsme upravený datový soubor, který zpracovala autorka diplomové práce [27].



Obrázek 5.1: Fundamentální diagram pro analyzovaná data

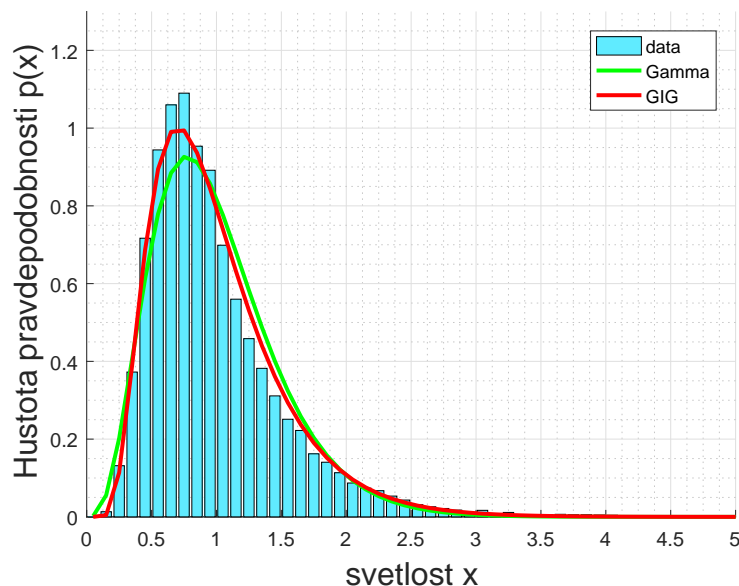
Struktura upraveného souboru je následující: Soubor má zhruba 22 000 řádků. Na každém řádku je uvedena hustota toku ρ , viz (3.10), lokální tok J , viz (3.9), číslo pruhu (rovno 1 ve všech

řádcích) a vektor 50-ti škálovaných časových světlostí x_k , viz (3.12) při vzorkování $m = 50$. Použití škálovaných časových světlostí má hned několik výhod: jak bylo již zmíněno, oproti prostorovým světlostem nejsou časové světlosti zatíženy systematickou chybou. Nemusíme uvažovat délku jednotlivých vozidel (kterou v případě přímo měřených dat zohlednit musíme). Přeskálováním dat zajistíme, že průměrná časová světlost bude v jakémkoliv vzorku rovna 1. Tato úprava nám umožní srovnávat dopravní data z různých dopravních režimů (tj. s různou hustotou toku) nebo z různých zdrojů (různých zemí).

Vykresleme fundamentální diagram pro hodnoty hustoty a toku, viz obrázek 5.1.

Jak již bylo zmíněno, chceme zkoumat interakce mezi vozidly (nejen sousedními). Oblastí našeho zájmu je tedy oblast kondenzované dopravy, neboť ve volné dopravě jsou interakce mezi vozidly velmi slabé. Srovnáme tento diagram se schématem 3.1. Pro další analýzu pak vybereme pouze ta data, která odpovídají pásmu hustot $\rho \in \langle 30, 70 \rangle$.

V předchozí kapitole jsme uvedli, že headway distribuci nejlépe popisuje distribuce *GIG*, což jsme ověřili na následujícím příkladu. Zvolili jsme náhodné pásmo hustot $\rho \in \langle 40, 41 \rangle$. Pro příslušný soubor škálovaných časových světlostí (dále jen světlostí) jsme pomocí metody minimální vzdálenosti (odvozené od L_2 normy, viz (2.20)) našli hustotu pravděpodobnosti distribuce Gamma, resp. *GIG*, viz (2.8), resp. (2.11). Srovnání grafu obou hustot je na obrázku 5.2.



Obrázek 5.2: Srovnání odhadu hustoty pravděpodobnosti distribuce Gamma a GIG

Z grafu je zřejmé, že distribuce *GIG* popisuje headway distribuci mnohem lépe. V oblasti nízkých hodnot světlostí je patrné, že distribuce Gamma postrádá narozdíl od *GIG* plato v 0, které je charakteristické pro všechny dopravní vzorky. Tentýž negativní výsledek bychom dostali i pro exponenciální, Gaussovo a Erlangovo rozdělení.

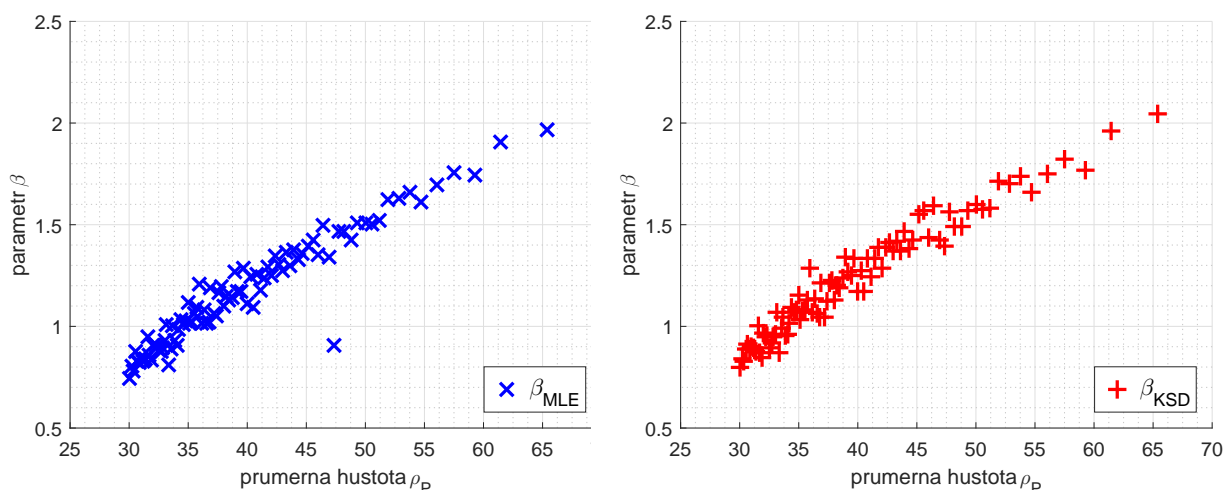
5.2 Metodika odhadování parametrů a statistické testování

Při analýze světlostí předpokládáme, že dopravní data mají rozdělení o hustotě pravděpodobnosti z dvouparametrické třídy *GIG*, viz (2.11). Nejprve jsme zkoumaná data rozdělili na sady o stejné velikosti (stejném počtu prvků) následovně:

- Z výchozího souboru jsme vybrali pouze řádky odpovídající pásmu hustot $\rho \in \langle 30, 70 \rangle$.

- Tyto řádky jsme seřadili vzestupně podle hustoty ϱ .
- Pro prvních 100 řádků jsme vypočetli průměrnou hustotu ϱ_P , k níž náleží sada 5000 světlostí x_k (velikost vzorkování 50 x 100 hustot).
- Z každé sady světlostí jsme vyřadili hodnoty $x_k > 10$ (tzv. *outliers*). Ty ve většině případů značí chybu detektoru.
- Takto jsme zpracovali každých dalších 100 řádků. Celkem jsme obdrželi 91 sad světlostí x_k .

Dále jsme pro každou sadu světlostí (charakterizované příslušnou ϱ_P) odhadli parametry (β, D) metodou MLE a metodou minimální kolmogorovské vzdálenosti (dále jen metoda KD). U první metody jsme hledali takový bodový odhad parametrů $(\hat{\beta}_{MLE}, \hat{D}_{MLE})$, který maximalizuje logaritmicou věrohodnostní funkci (4.1). U druhé metody jsme hledali takový bodový odhad parametrů $(\hat{\beta}_{KD}, \hat{D}_{KD})$, který naopak minimalizuje kolmogorovskou vzdálenost, viz (2.18). Hodnoty odhadu parametrů jsou shrnuty v tabulce 5.1.



Obrázek 5.3: Závislost parametru $\hat{\beta}_{MLE}$ a $\hat{\beta}_{KD}$ na hustotě ϱ_P

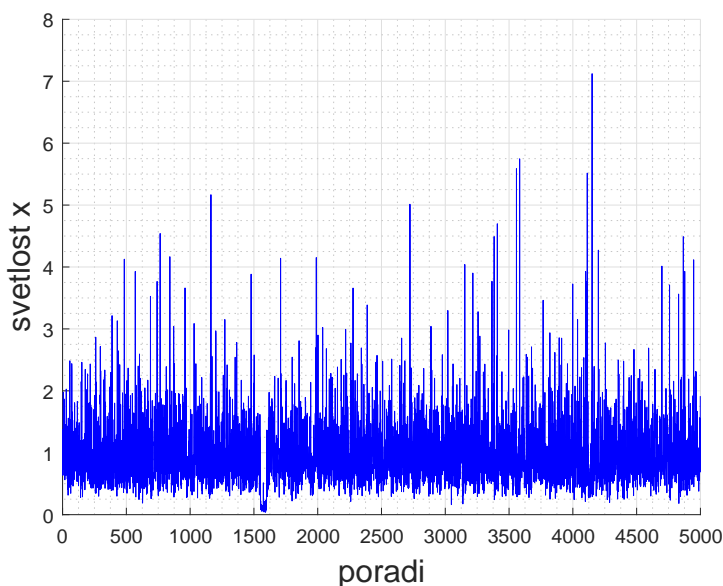
Na obrázku 5.3 vidíme závislost parametrů $\hat{\beta}_{MLE}$ a $\hat{\beta}_{KD}$ na hustotě ϱ_P . Obecně lze říci, že parametr β s rostoucí hustotou ϱ roste, což koresponduje s fyzikální (termodynamickou) interpretací parametru. Obvykle bývá označován termínem *rezistivita*, protože popisuje odolnost systému vůči stochastickému šumu. V grafu pro $\hat{\beta}_{MLE}$ lze detekovat zhruba v polovině rozsahu hustot bod, který se tomuto trendu výrazně vymyká. V grafu pro $\hat{\beta}_{KD}$ však takový bod není. Zřejmě jde o údaj odpovídající $\varrho_P = 47, 35$, viz tabulka 5.1.

Soubor světlostí pro $\varrho_P = 47, 35$ jsme analyzovali na obrázku 5.4, kam jsme vynesli jednotlivé hodnoty světlostí v závislosti na pořadí měření. Postupujeme-li od počátku, mají světlosti zcela náhodnou velikost, ale mezi 1500-tou a 1550-tou hodnotou dojde náhle k výraznému propadu. Detektor zde nejspíš na určitou dobu nefungoval správně. Z výše uvedeného lze vyvodit zajímavý závěr - metoda MLE je vůči takovým náhlým výkyvům více citlivá než metoda KD. Jako vhodnější metoda k odhadu parametru se tedy jeví metoda KD, což ověříme v další části. Na závěr uvádíme srovnání histogramu světlostí a distribuce o parametrech $(\hat{\beta}_{KD}, \hat{D}_{KD})$ pro dvě náhodně zvolené hustoty, $\varrho_P = 34, 57$ a $\varrho_P = 43, 28$, viz obrázek 5.5.

KAPITOLA 5. VYHODNOCENÍ EMPIRICKÝCH DAT – MODEL PRO ROZDĚLENÍ SVĚTLOSTÍ

hustota ϱ_P	$\hat{\beta}_{MLE}$	\hat{D}_{MLE}	$\hat{\beta}_{KD}$	\hat{D}_{KD}	hustota ϱ_P	$\hat{\beta}_{MLE}$	\hat{D}_{MLE}	$\hat{\beta}_{KD}$	\hat{D}_{KD}
30,08	0,74	2,04	0,80	2,23	38,20	1,15	2,48	1,20	2,61
30,21	0,80	2,11	0,84	2,26	38,45	1,13	2,46	1,19	2,60
30,36	0,78	2,08	0,83	2,24	38,69	1,14	2,48	1,24	2,71
30,52	0,88	2,18	0,89	2,30	38,92	1,27	2,62	1,34	2,79
30,68	0,83	2,13	0,92	2,40	39,15	1,17	2,51	1,27	2,70
30,83	0,82	2,13	0,89	2,33	39,43	1,17	2,53	1,25	2,72
30,98	0,85	2,16	0,90	2,30	39,67	1,29	2,63	1,33	2,75
31,13	0,83	2,15	0,90	2,34	39,94	1,11	2,45	1,17	2,60
31,24	0,84	2,14	0,89	2,34	40,26	1,24	2,59	1,28	2,72
31,40	0,84	2,14	0,87	2,28	40,52	1,09	2,42	1,17	2,61
31,57	0,95	2,27	1,00	2,42	40,79	1,25	2,60	1,33	2,74
31,73	0,86	2,16	0,87	2,27	41,12	1,18	2,52	1,25	2,68
31,86	0,84	2,14	0,85	2,26	41,46	1,24	2,58	1,33	2,80
32,01	0,91	2,22	0,96	2,38	41,75	1,29	2,65	1,39	2,84
32,18	0,91	2,23	0,95	2,37	42,04	1,25	2,59	1,28	2,68
32,33	0,90	2,22	0,95	2,40	42,34	1,35	2,70	1,39	2,82
32,50	0,88	2,19	0,89	2,32	42,66	1,30	2,65	1,41	2,86
32,67	0,88	2,18	0,92	2,35	43,00	1,27	2,62	1,37	2,78
32,83	0,90	2,21	0,95	2,38	43,28	1,37	2,72	1,42	2,86
32,99	0,93	2,25	0,96	2,39	43,59	1,30	2,65	1,37	2,82
33,16	1,01	2,33	1,07	2,49	43,94	1,38	2,73	1,47	2,89
33,37	0,81	2,13	0,87	2,34	44,32	1,33	2,67	1,38	2,80
33,55	0,89	2,20	0,99	2,43	44,67	1,36	2,71	1,42	2,84
33,71	1,01	2,34	1,04	2,46	45,12	1,40	2,75	1,55	3,01
33,92	0,93	2,26	0,96	2,39	45,56	1,43	2,79	1,57	3,04
34,07	0,91	2,22	0,96	2,39	45,99	1,35	2,70	1,44	2,91
34,23	0,99	2,31	1,01	2,44	46,45	1,49	2,85	1,59	3,03
34,41	1,03	2,36	1,09	2,49	46,92	1,34	2,70	1,43	2,87
34,57	1,01	2,33	1,07	2,49	47,35	0,90	2,24	1,39	2,86
34,76	1,03	2,36	1,08	2,53	47,76	1,47	2,82	1,56	3,00
34,96	1,12	2,45	1,15	2,56	48,21	1,46	2,83	1,49	2,91
35,14	1,02	2,34	1,04	2,47	48,75	1,42	2,77	1,49	2,92
35,32	1,02	2,34	1,08	2,49	49,32	1,51	2,88	1,57	3,01
35,52	1,06	2,39	1,08	2,52	49,99	1,51	2,87	1,60	3,05
35,72	1,09	2,42	1,13	2,54	50,60	1,50	2,87	1,57	3,02
35,92	1,21	2,55	1,29	2,72	51,21	1,52	2,88	1,58	3,00
36,12	1,01	2,33	1,07	2,50	51,93	1,62	2,99	1,71	3,17
36,34	1,08	2,41	1,14	2,55	52,80	1,63	3,00	1,70	3,16
36,53	1,01	2,34	1,06	2,51	53,72	1,66	3,03	1,74	3,18
36,73	1,02	2,34	1,05	2,47	54,74	1,61	2,97	1,66	3,12
36,92	1,19	2,53	1,21	2,60	56,07	1,69	3,06	1,75	3,20
37,14	1,07	2,39	1,05	2,45	57,55	1,76	3,13	1,82	3,29
37,36	1,05	2,38	1,13	2,57	59,23	1,75	3,12	1,77	3,23
37,59	1,16	2,51	1,21	2,63	61,40	1,91	3,28	1,96	3,43
37,81	1,19	2,53	1,22	2,64	65,40	1,96	3,34	2,04	3,51
37,99	1,10	2,45	1,13	2,55	-	-	-	-	-

Tabulka 5.1: Bodové odhady parametrů ($\hat{\beta}_{MLE}$, \hat{D}_{MLE}) a ($\hat{\beta}_{KD}$, \hat{D}_{KD}) pro světlosti příslušné průměrné hustotě ϱ_P



Obrázek 5.4: Závislost hodnoty světlosti na pořadí měření pro pásmo o $\varrho_p = 47,35$

Model jsme opět prověřili statistickými testy obdobnou metodou jako v případě generátoru pseundonáhodných čísel. Testujeme nulovou hypotézu, že zvolený vzorek světlostí má rozdělení GIG. Využili jsme distribuční funkci CDF_{MLE} , resp. CDF_{KD} spočtenou numerickou integrací z parametrů odhadnutých metodou MLE, resp. KD a empirickou distribuční funkci CDF_E . Postupovali jsme takto:

- Ze souboru světlostí o dané průměrné hustotě ϱ_p jsme náhodně vybrali vzorek o $n = 200$ prvcích.
- Pro tento výběr jsme metodou MLE a KD odhadli parametry $(\hat{\beta}_{MLE}, \hat{D}_{MLE}), (\hat{\beta}_{KD}, \hat{D}_{KD})$ a vypočítali distribuční funkce CDF_E, CDF_{MLE} a CDF_{KD} .
- Otestovali jsme CDF_{MLE} vůči CDF_E třemi způsoby (KS test, χ^2_{k-1} a χ^2_{k-p-1} ; hladina významnosti $\alpha = 0,05$).
- Otestovali jsme CDF_{KD} vůči CDF_E dvěma způsoby (χ^2_{k-1} a χ^2_{k-p-1} test; hladina významnosti $\alpha = 0,05$).
- Tento cyklus jsme provedli 1000–krát a pro každý typ testu určili, v kolika procentech z celkového počtu realizací testu byla nulová hypotéza zamítnuta.
- Výše uvedený postup jsme aplikovali postupně na všech 91 pásmech o hustotách ϱ_p .

Souhrn výsledků pro vzorek $n = 200$ je uveden v tabulkách 5.2 a 5.3. Analogicky jsme provedli testování pro velikost vzorku $n = 300, n = 400$ a $n = 500$. Tabulky s výsledky jsou vzhledem ke své rozsáhlosti uvedeny v přílohách, viz tabulka 1, 2, 3, 4, 5, 6.

Abychom mohli získané výsledky vyhodnotit, vypočítali jsme pro každý vzorek a každý typ testu průměrnou hodnotu procenta zamítnutí (pro celé pásmo hustot), viz tabulka 5.4. Z tabulky je zřejmé, že pro rostoucí velikost vzorku vzrůstá také počet zamítnutí u všech testů. Dále je zjevné, že počet zamítnutí u libovolného vzorku je pro test χ^2_{k-p-1} vždy vyšší než pro test χ^2_{k-1} . Skutečná

ϱ_p	MLE KS[%]	MLE χ^2_{k-p-1} [%]	MLE χ^2_{k-1} [%]	KD χ^2_{k-p-1} [%]	KD χ^2_{k-1} [%]
30,08	7,4	14,9	8,7	12,6	6,5
30,21	6,7	17,3	11,1	14,3	8,9
30,36	6,3	14,5	9,2	10,5	5,8
30,52	2,1	12,4	6,6	12,4	7,6
30,68	14,2	23,5	16,3	20,2	11,9
30,83	7,2	16,6	10,1	13,2	8,2
30,98	2,4	11,2	6,8	8,1	3,7
31,13	6,5	15,7	9,9	15,4	9,9
31,24	9,9	18,9	12,6	17,2	9,9
31,40	3,4	16,3	9,7	14,5	7,8
31,57	3,6	13,4	7,4	11,7	7,5
31,73	3,2	11,4	7,5	9,1	5,0
31,86	3,3	14,4	8,4	14,6	8,1
32,01	3,6	14,4	9,5	14,5	9,4
32,18	3,4	11,5	7,6	10,6	6,3
32,33	6,0	15,1	9,6	13,7	8,4
32,50	7,0	17,7	11,8	18,9	12,8
32,67	5,4	14,3	9,5	11,7	7,4
32,83	4,7	14,9	8,5	11,9	7,5
32,99	4,7	16,4	8,7	13,8	8,7
33,16	2,0	8,9	4,6	9,0	5,1
33,37	12,7	19,1	11,4	16,5	9,0
33,55	3,7	14,4	8,8	10,0	5,5
33,71	1,8	8,6	3,4	5,9	3,3
33,92	6,4	15,1	9,8	12,0	6,7
34,07	4,3	14,0	8,5	13,2	7,8
34,23	3,3	11,8	6,3	12,0	7,7
34,41	1,7	12,6	7,2	11,0	6,1
34,57	2,0	11,4	6,1	9,4	5,5
34,76	4,9	12,9	7,6	9,9	6,1
34,96	0,9	8,8	5,0	7,5	4,2
35,14	4,7	11,7	7,6	13,6	7,4
35,32	1,5	11,1	6,2	10,7	5,8
35,52	4,7	13,6	8,7	14,5	9,2
35,72	1,3	8,0	5,3	8,6	3,9
35,92	1,0	10,2	4,6	9,7	5,3
36,12	4,6	17,7	10,9	15,2	9,8
36,34	0,7	12,6	6,4	11,1	6,8
36,53	5,8	15,7	9,0	14,5	8,4
36,73	2,5	12,6	6,7	9,2	5,2
36,92	0,1	10,1	6,0	10,8	5,9
37,14	2,1	11,4	6,8	12,1	7,6
37,36	4,6	10,7	6,1	7,9	4,4
37,59	2,9	11,5	6,3	10,3	5,9
37,81	1,8	12,6	7,7	11,4	6,6
37,99	2,3	9,8	5,4	8,1	4,8

Tabulka 5.2: Procento zamítnutí pro test KS, χ^2_{k-p-1} , χ^2_{k-1} a velikost vzorku $n = 200$

5.2. METODIKA ODHADOVÁNÍ PARAMETRŮ A STATISTICKÉ TESTOVÁNÍ

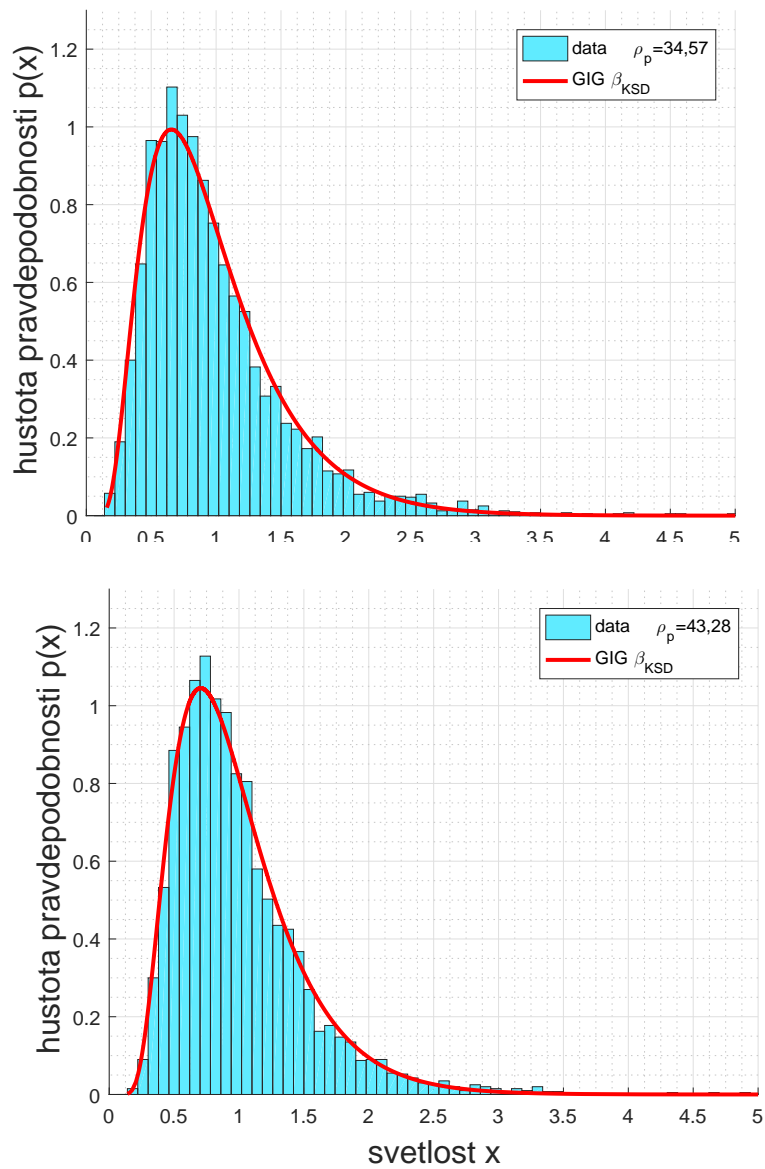
q_p	MLE KS[%]	MLE χ^2_{k-p-1} [%]	MLE χ^2_{k-1} [%]	KD χ^2_{k-p-1} [%]	KD χ^2_{k-1} [%]
38,20	1,6	9,4	6,1	8,7	4,4
38,45	1,0	10,1	6,2	8,5	4,7
38,69	6,9	14,0	8,5	14,0	8,0
38,92	2,8	11,7	6,8	10,7	6,5
39,15	0,6	9,8	5,2	7,6	4,6
39,43	6,6	13,6	8,1	8,0	4,6
39,67	1,1	9,1	5,3	8,7	5,5
39,94	2,2	9,8	6,5	8,8	5,4
40,26	2,4	12,2	7,4	11,2	5,8
40,52	2,1	11,6	7,9	10,8	7,1
40,79	0,5	8,0	4,2	8,6	4,6
41,12	2,6	12,4	7,2	12,6	7,2
41,46	1,6	10,8	4,9	10,5	6,6
41,75	2,6	12,1	7,7	9,1	4,6
42,04	0,2	6,8	3,2	7,4	3,8
42,34	1,1	9,8	5,1	10,3	5,6
42,66	1,1	10,1	5,8	7,4	4,4
43,00	0,6	7,7	4,7	6,9	4,1
43,28	0,6	9,6	4,9	7,9	5,0
43,59	1,3	8,1	4,2	7,0	4,0
43,94	0,4	8,8	5,0	8,4	3,6
44,32	0,6	6,5	3,2	6,7	3,8
44,67	0,6	7,9	3,7	9,3	5,0
45,12	1,2	9,0	5,4	8,4	4,2
45,56	4,0	15,9	9,7	11,8	7,0
45,99	2,0	13,9	8,5	12,6	7,7
46,45	0,4	8,2	4,1	7,6	4,5
46,92	1,6	8,8	4,8	7,4	4,7
47,35	21,0	34,9	26,0	7,1	4,4
47,76	0,5	6,9	4,3	6,6	4,4
48,21	1,9	9,2	5,6	8,5	4,4
48,75	0,8	9,1	5	8,3	5
49,32	1,1	8,1	4,6	7,4	4,3
49,99	1,1	8	4,6	7,4	4,2
50,60	0,8	11,9	6,1	11,2	6
51,21	0,5	7,7	4,5	7,6	4,1
51,93	0,6	11,2	6,9	9,4	5,7
52,80	1,0	10	5,7	9,9	5,4
53,72	0,3	5,9	2,3	6	3,1
54,74	1,1	9,5	4,9	8,9	4,8
56,07	0,7	8,8	5,5	8,1	4,2
57,55	0,4	9,3	5,1	7,6	3,9
59,23	0,6	8	3,6	8,2	4,5
61,40	0,5	8,4	3,9	7,1	4
65,40	0,7	7,4	3,8	6,5	3,5

Tabulka 5.3: Procento zamítnutí pro test KS, χ^2_{k-p-1} , χ^2_{k-1} a velikost vzorku $n = 200$
v pásmu 47-91

vzorek n	MLE	MLE	MLE	KD	KD
	KS[%]	χ^2_{k-p-1} [%]	χ^2_{k-1} [%]	χ^2_{k-p-1} [%]	χ^2_{k-1} [%]
200	3,1	12,0	7,0	10,44	6,1
300	4,8	15,2	9,2	13,27	7,9
400	7,8	19,2	11,9	16,34	10,4
500	11,2	23,2	15,8	19,59	12,7

Tabulka 5.4: Průměrný počet zamítnutí testem KS, χ^2_{k-p-1} a χ^2_{k-1} pro různé velikosti testovaného souboru

hodnota se nachází mezi oběma údaji. Na závěr porovnejme výsledky testu χ^2_{k-1} pro metodu MLE a KD. Hodnoty zamítnutí u metody MLE jsou pro všechny velikosti n vyšší než u metody KD. U testu χ^2_{k-1} pro metody MLE a KD platí totéž. Distribuce získaná metodou KD je tedy zamítána v menším počtu případů. Z toho lze usoudit, že metoda minimalizace kolmogorovské vzdálenosti je v této oblasti vhodnějším nástrojem pro hledání optimálního odhadu, což koresponduje s naším závěrem ze sekce 5.2.



Obrázek 5.5: Srovnání histogramu světlostí a distribuce o parametrech $(\hat{\beta}_{KD}, \hat{D}_{KD})$ pro $\rho_p = 34,57$ a $\rho_p = 43,28$

Kapitola 6

Poruchová konvoluční funkce – nástroj pro detekci statistické závislosti v dopravních datech

6.1 Zavedení pojmu multisvětlost

V předešlých kapitolách jsme zkoumali rozdělení škálovaných časových světlostí x , jež vyjadřují časový odstup dvou sousedních vozidel. Obdobným způsobem můžeme však analyzovat také odstupy nesousedních vozidel. Zavedeme tedy tzv. multisvětlost n -tého řádu $x_k|n$ vztahem

$$x_k|n := \sum_{i=0}^n x_{k+i}, \quad (6.1)$$

kde x_{k+i} označuje $(k + i)$ -tou světlost, přičemž symbol $x_k|n$ chápeme jako kompaktní symbol. Multisvětlost n -tého řádu vyjadřuje vzdálenost $n + 2$ vozidel, mezi nimiž je $n + 1$ světlostí. (Světlost dvou vozidel je tedy vlastně multisvětlost 0-tého řádu.) Stejně jako u souboru světlostí nás u souboru multisvětlostí bude zajímat jeho rozdělení, kde hustotu pravděpodobnosti rozdělení multisvětlostí řádu n označíme $g(x|n)$.

6.2 Statistické odhady pro rozdělení multisvětlostí

Při analýze multisvětlostí předpokládáme, že dopravní data mají rozdělení o hustotě pravděpodobnosti z tříparametrické třídy GIG; viz (2.9). Na začátku jsme ze souboru dopravních dat pro každý řád n vypočetli multisvětlosti. Postup je následující:

- Z výchozího souboru jsme vybrali pouze řádky odpovídající pásmu hustot $\varrho \in \langle 30, 70 \rangle$.
- Tyto řádky jsme seřadili vzestupně podle hustoty ϱ .
- Pro prvních 100 řádků jsme vypočetli průměrnou hustotu ϱ_p . K ní náleží 100 sad světlostí x o 50 prvcích.
- Z každé 50-tice jsme vyřadili outliery ($x > 10$) a určili vektor multisvětlostí aplikací vztahu (6.1). (Například pro řád $n = 2$ sčítáme první, druhou a třetí světlost, pak druhou, třetí a čtvrtou atd.) Těchto 100 sad multisvětlostí jsme pak spojili do jednoho souboru příslušného dané hustotě ϱ_p . Zde je na místě zdůraznit, že jednotlivé početní operace musejí proběhnout přesně v tomto pořadí. Údaje v každé sadě musejí při výpočtu multisvětlostí zůstat v původním pořadí.

- Takto jsme zpracovali každých dalších 100 řádků. Celkem jsme obdrželi pro každý řád n 91 sad multisvětlostí $x_k|n$.

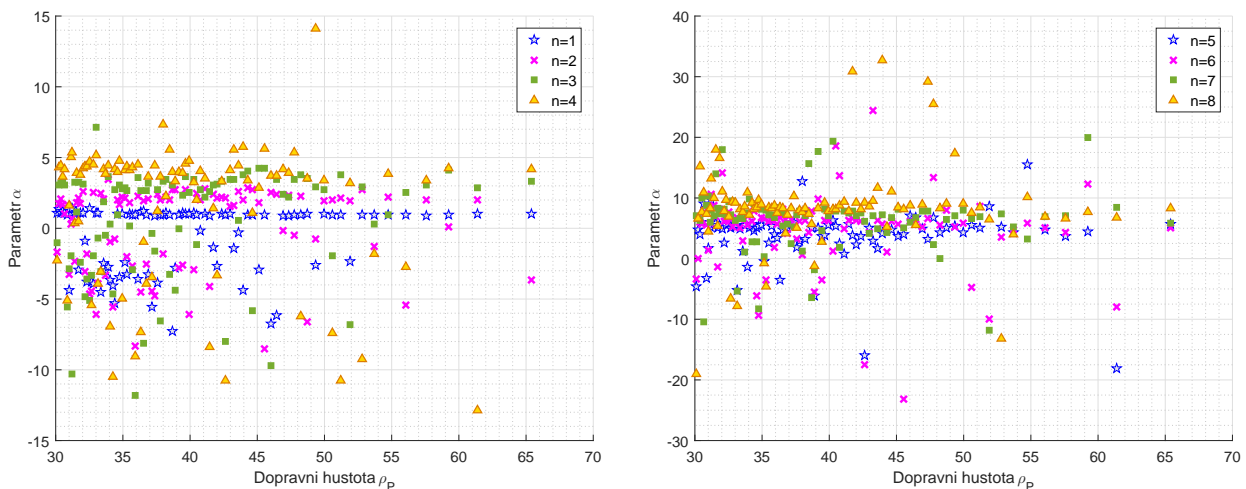
Následně jsme pro každý řád n a každou sadu multisvětlostí charakterizovanou příslušnou hustotou q_P odhadli parametry (α, β, D) metodou KD. Nalezli jsme bodový odhad parametrů $(\hat{\alpha}_{KD,n}, \hat{\beta}_{KD,n}, \hat{D}_{KD,n})$, který minimalizuje kolmogorovskou vzdálenost, viz (2.18). Při výpočtu odhadu touto metodou v programu MATLAB lze zadat inicializační hodnoty minimalizace (vektor hodnot e_1, e_2, e_3 , poblíž kterých se program snaží najít odhad parametrů). V tomto případě jsme za inicializační hodnoty dosazovali hodnoty plynoucí ze vztahu (2.12) pro konvoluci dvou GIG rozdělení o stejných parametrech a ze vztahu

$$p_n(x) = p_{k-1}(x) \star p_0(x). \quad (6.2)$$

Ten popisuje, jak získat z hustoty pravděpodobnosti světlosti $p_0(x)$ hustotu pravděpodobnosti multisvětlosti n -tého řádu. Pro hodnoty e_1, e_2, e_3 odtud platí

$$\begin{aligned} e_1 &= n, \\ e_2 &= (n + 1)^2 \hat{\beta}, \\ e_3 &= \hat{D}, \end{aligned}$$

kde $\hat{\beta}, \hat{D}$ jsou odhady parametrů odpovídající téže hustotě vypočtené v páté kapitole, viz tabulka 5.1. Závislost parametru $\hat{\alpha}_{KD}$, resp. $\hat{\beta}_{KD}$ na hustotě q_P pro různé řády jsme vykreslili do obrázku 6.1, resp. 6.2.

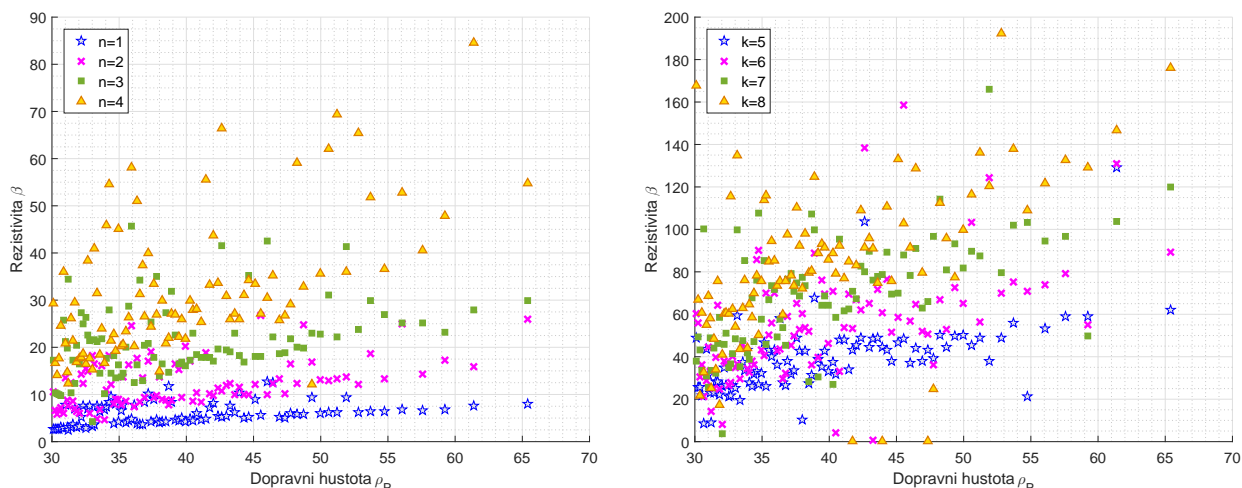


Obrázek 6.1: Závislost parametru $\hat{\alpha}_{KD}$ na hustotě q_P pro řád multisvětlosti $n = 1$ až $n = 4$, resp. $n = 5$ až $n = 8$

Na obrázku 6.1 vidíme, že závislost parametru α na hustotě nevykazuje žádný typický trend, hodnoty fluktuují kolem nuly. Z obrázku 6.2 jsou patrné dvě zásadní skutečnosti:

- Pro daný řád n rezistivita $\hat{\beta}_{KD}$ vzrůstá s rostoucí dopravní hustotou q_P .
- Pro fixovanou hodnotu q_P rezistivita $\hat{\beta}_{KD}$ pro zvyšující se hodnotu řádu n roste.

V předešlém textu bylo již uvedeno, že škálované časové světlosti mají $E(X) = 1$. Pak mají multisvětlosti řádu n střední hodnotu $E(X) = n + 1$. Parametr D je i v případě multisvětlostí funkcí parametrů α, β podle vztahu (2.10). Graf závislosti parametru D na hustotě q_P tedy neuvádíme.



Obrázek 6.2: Závislost rezistivity $\hat{\beta}_{KD}$ na hustotě ρ_P pro řád multisvětlosti $n = 1$ až $n = 4$, resp. $n = 5$ až $n = 8$

6.3 Koncept poruchové konvoluční funkce a analýza jejího průběhu

Základní úlohou této diplomové práce je stanovení dostahu interakčních sil mezi vozidly pohybujícími se na reálných komunikacích. Zajímá nás, na kolik předcházejících vozidel řidič reaguje, a chceme vytvořit nástroj, s jehož pomocí lze tento počet vozidel (tzv. *interakční dosah*) určit.

Můžeme vyjít z představy, že řidič je při řízení ovlivněn pouze chováním předcházejícího (sousedního) vozidla. To by v řeci dopravních veličin znamenalo, že sousední světlosti jsou zcela nezávislé náhodné veličiny. Matematickými prostředky je pak možné otestovat, zda je tomu skutečně tak. My však máme zcela logické podezření založené na empirické zkušenosti z reálného dopravního provozu, že řidič reaguje na více vozidel. Předpokládáme tedy, že světlosti jsou více či méně závislé a snažíme se tuto závislost nějakým způsobem změřit. Dosud jsme se zabývali stanovením parametrů rozdělení světlostí a multisvětlostí. Jako vhodný prostředek ke zkoumání závislosti se tedy nabízí rozdíl mezi dvěma hustotami pravděpodobnosti $r(x)$ a $q(x)$, resp. distribučními funkcemi, z nichž jedna reprezentuje ideální případ, kdy jsou světlosti nezávislé, a druhá představuje rozdělení reálných dat. Pokud jsou světlosti nezávislé, vypočteme hustotu pravděpodobnosti jejich součtu $r(x)$ podle vztahu (2.14) jako konvoluci hustot příslušných dílčím světlostem. Hustotou pravděpodobnosti multisvětlostí $q(x)$ z reálného dopravního toku a odhadem jejich parametrů jsme se zabývali v předešlém textu.

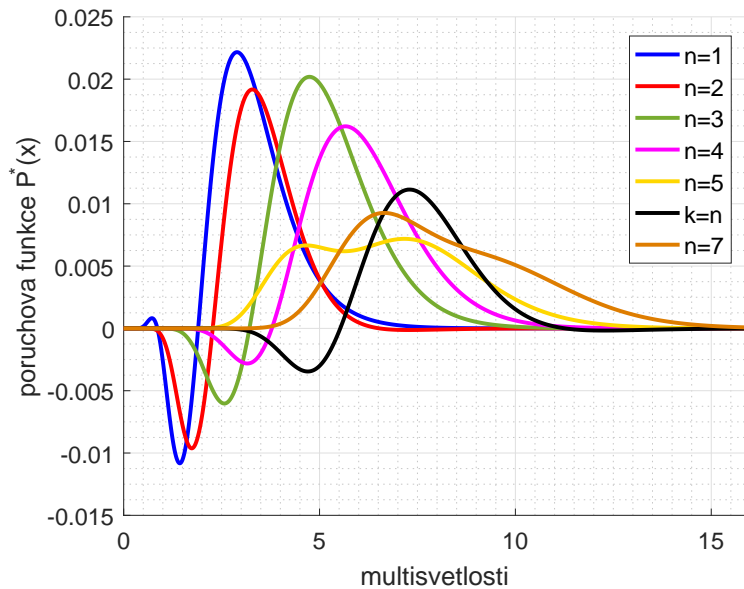
Zavedeme pro výše zmíněné hustoty $r(x)$, $q(x)$ *poruchovou konvoluční funkci* $P^*(x)$ (dále jen poruchová funkce) vztahem

$$P^*(x) = r(x) - q(x). \quad (6.3)$$

Pokud by platila závislost světlostí, pro velké počty pozorování by se hodnota $P^*(x)$ měla blížit nule. Nenulová hodnota tedy může ukazovat na nezávislost světlostí.

Průběh poruchové funkce si ukážeme na ilustrativním případě, viz obrázek (6.3), kde je pro dopravní vzorek v daném pásmu hustot vykreslena poruchová funkce pro osm různých řádů multisvětlostí.

Poruchová funkce v sobě skrývá kýženou informaci o nezávislosti náhodných veličin. Jedná se však o zcela nový přístup, žádná teorie o poruchové konvoluční funkci nebyla dosud publikována. Naším cílem tedy bude rozhodnout, zda je navržený koncept vhodný a může fungovat jako prostředek k detekci dosahu interakčních sil v dopravních systémech. Abychom mohli



Obrázek 6.3: Průběh poruchové funkce $P^*(x)$ v pásmu o průměrné hustotě $\rho_P = 35,14$ pro řád multisvětlosti $n = 1$ až $n = 8$

koncept poruchové funkce aplikovat, potřebujeme informaci o závislosti obsaženou v poruchové funkci nějak kvantifikovat. Zvolíme si *funkcionální normu* $\|\cdot\|$ (míru závislosti), která danou poruchovou funkci charakterizuje. Nabízí se vzdálenost funkce $P^*(x)$ od nulové funkce, která odpovídá nezávislým náhodným veličinám. Čím vyšší hodnotu má tato vzdálenost, tím výraznější je závislost analyzovaných náhodných veličin. Vhodného kandidáta budeme logicky volit z množiny známých norem. Princip rostoucí vzdálenosti (normy) při rostoucí závislosti by měl platit zcela obecně bez ohledu na volbu normy. Volíme tedy dvě různé normy, které již byly definovány v předešlé kapitole. První je L_1 norma daná vztahem (2.19), druhou je kolmogorovská vzdálenost (dále kolmogorovská norma) daná vztahem (2.18), do níž se místo hustot pravděpodobnosti dosazují příslušné distribuční funkce.

Pro obě varianty chceme ukázat, že norma poruchové funkce klesá v daném pásmu dopravních hustot s rostoucím řádem multisvětlosti n . Tento trend by potvrdil náš intuitivní předpoklad, že řidič na vzdálenější vozidla reaguje méně. Od jistého řádu n bude norma nulová, což koresponduje s představou, že pohyb $(n + 1)$ -ho předcházejícího vozidla již na rozhodování řidiče nemá žádný vliv. Podaří-li se nám prokázat klesající trend normy $\|P^*(x)\|$ popsany výše, můžeme funkci $P^*(x)$ pokládat za nástroj vhodný k detekci dosahu interakce.

Dosud jsme ho popisovali pouze kvalitativně. Víme, že nižší hodnotě normy odpovídá slabší interakce vozidel. Nemůžeme však zatím rozhodnout, od kolikátého předcházejícího vozidla (tj. od kterého řádu multisvětlosti n) budeme považovat interakci za nulovou. Nedílnou součástí navrhovaného aparátu proto musí být soustava kritických hodnot řádu n (*rozhodovacích kritérií*). U reálných dopravních vzorků je počet dat zřejmě vždy konečný, takže i nezávislým světlostem může příslušet nenulová norma $\|P^*(x)\|$ (i když její hodnota bude velmi malá). Další etapou by tedy mělo být stanovení kritických hodnot tak, aby se norma poruchové funkce pod kritickou hranicí lišila od nulové je nepatrně. Od této kritické hodnoty lze světlosti prohlásit za nezávislé. Prakticky to znamená, že vozidlo odpovídající kritické hodnotě (řádu) a všechna vozidla jemu předcházející už s uvažovaným vozidlem neinteragují. Výpočet kritických hodnot by mohl být založen například na principu kolmogorovské vzdálenosti. Pro jejich stanovení by bylo nutné vytvořit model generující dopravní data, v němž volbou interakčního potenciálu určíme, od kterého řádu budou světlosti

nezávislé. Vypočtená norma poruchové funkce pak bude rovna kritické hodnotě. Takto lze sestavit souhrn kritických hodnot pro různé hustoty. Potvrdí-li se tedy naše hypotéza o chování poruchové funkce, mohl by se tento model stát předmětem dalšího zkoumání.

6.4 Empirické chování normy poruchové funkce

V této části se budeme věnovat výpočtu normy poruchové funkce pro empirická data. Testujeme předpoklad, že hodnota normy bude s rostoucím řádem klesat. Použijeme 91 souborů multisvětlostí, jež jsme zavedli v sekci 6.2. Výpočet provedeme v každém z 91 pásem pro L_1 normu i kolmogorovskou normu, jak bylo popsáno v předešlé sekci. Sestavme nejprve pro oba případy poruchovou funkci podle vztahu (6.3). kolmogorovskou normou (2.18) budeme pro daný řád n hodnotit funkci

$$P_n^*(x) = CDF_{KD,n} - CDF_{E,n}.$$

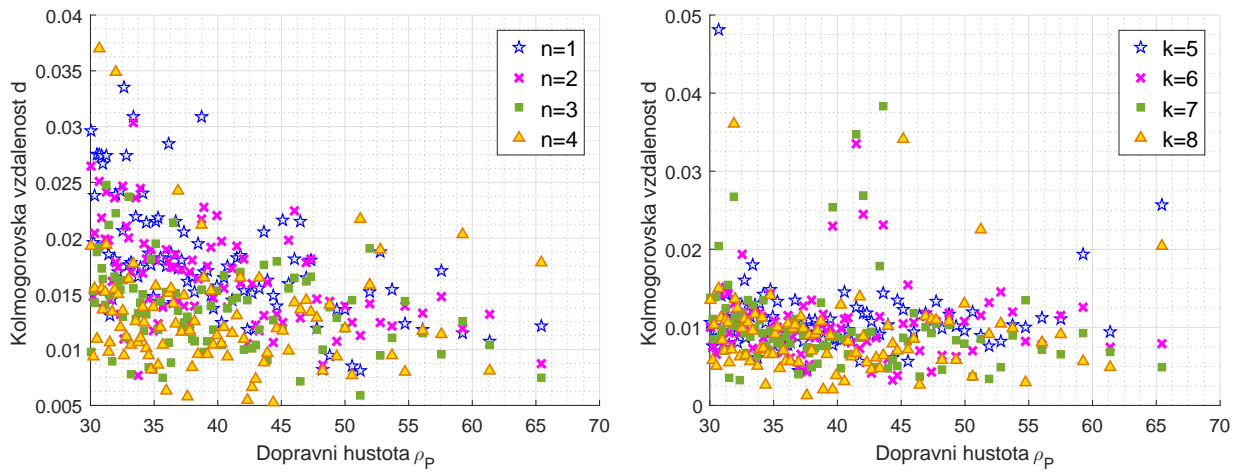
Symbolem $CDF_{KD,n}$ rozumíme distribuční funkci vypočtenou z odhadnutých parametrů $(\hat{\alpha}_{KD,n}, \hat{\beta}_{KD,n}, \hat{D}_{KD,n})$, viz sekce 6.2. Ta reprezentuje rozdělení reálných dat. Funkce $CDF_{E,n}$ reprezentuje nezávislé světlosti, je to analytická funkce vypočtená numericky z hustoty pravděpodobnosti $p_n(x)$ dané vztahem (6.2). Normou L_1 danou vztahem (2.19) vyhodnotíme funkci

$$P_n^*(x) = PDF_{KD,n} - PDF_{E,n},$$

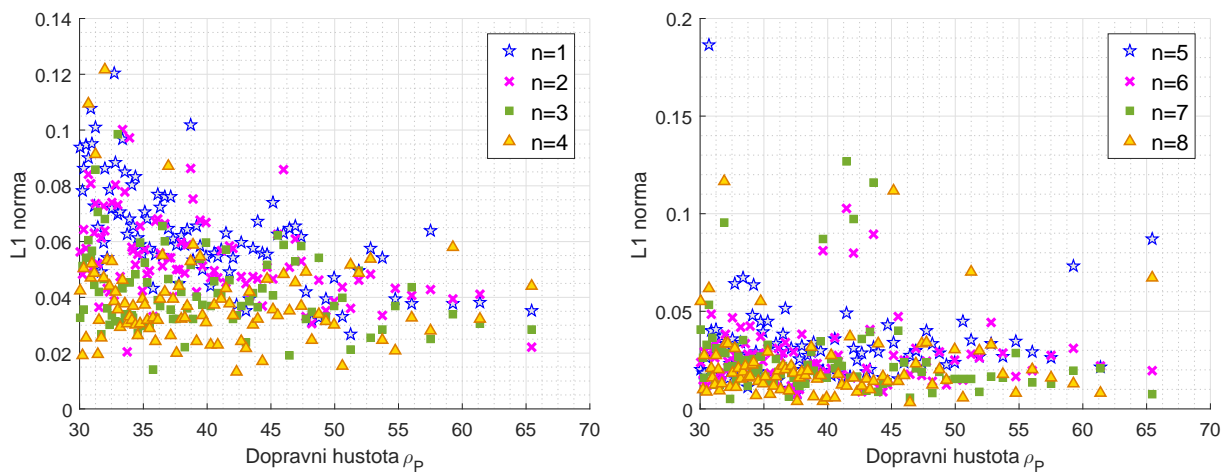
kde analogicky symbolem $PDF_{KD,n}$ rozumíme hustotu pravděpodobnosti vypočtenou z odhadnutých parametrů $(\hat{\alpha}_{KD,n}, \hat{\beta}_{KD,n}, \hat{D}_{KD,n})$. Symbol $PDF_{E,n}$ reprezentuje nezávislé světlosti, je to analytická hustota pravděpodobnosti $p_n(x)$ daná vztahem (6.2) vypočtená numericky.

Výsledky našich výpočtů jsme sumarizovali graficky. Do obrázku (6.4) jsme vynesli hodnoty kolmogorovské normy pro všech 8 řádů multisvětlostí v závislosti na dopravní hustotě ρ_p . Je zřejmé, že hodnota normy klesá s rostoucí dopravní hustotou, tento trend je společný pro všechny řády. Tento výsledek koresponduje s empirickou zkušeností - při nižší hustotě provozu řidič zohledňuje chování několika vozidel před sebou, při vysoké hustotě (zácpě) už se řídí prakticky pouze svým předchůdcem. Z obrázku (6.5) lze vyvodit identické závěry. L_1 norma s rostoucí dopravní hustotou klesá pro všechny řády multisvětlosti.

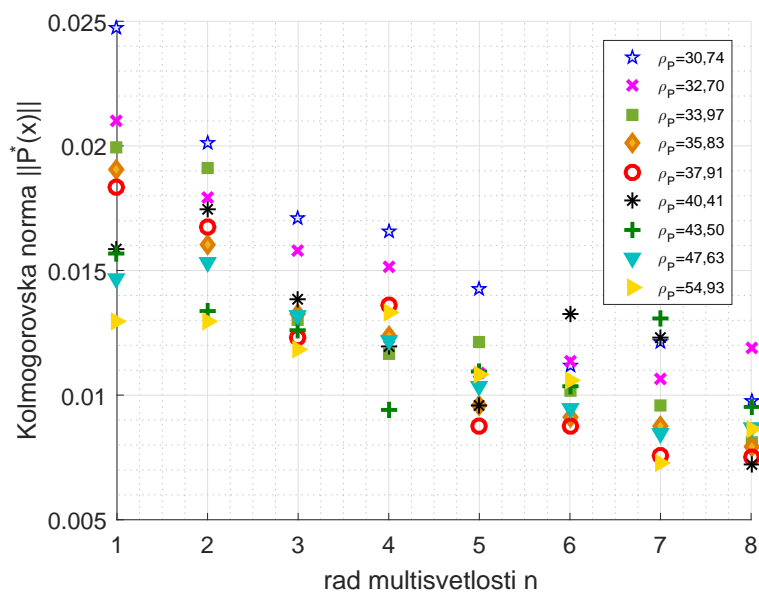
Dále jsme vypočetli v pásmu prvních 10 hustot pro všechny řády průměrné hodnoty hustoty ρ_p i kolmogorovské a L_1 normy. Totéž jsme učinili se zbývajícimi hustotními pásmy. Výsledek jsme opět graficky sumarizovali do obrázků (6.6), resp. (6.7), v nichž je vynesena závislost kolmogorovské normy, resp. L_1 normy poruchové funkce na řádu multisvětlosti pro 9 různých hustotních zón. Vyplývají z nich naprosto zásadní závěry. Norma poruchové funkce v obou případech zcela jasně pro dané pásmo dopravních hustot klesá s rostoucím řádem multisvětlosti n . Potvrzuje se tedy předpoklad, že závislost světlostí s rostoucím řádem klesá a interakce vozidel je slabší. Podařilo se nám prokázat klesající trend normy poruchové funkce, poruchovou funkci $P^*(x)$ tedy lze pokládat za nástroj vhodný k detekci dosahu interakčních sil v dopravních systémech. Pokud analyzujeme kolmogorovskou normu, resp. L_1 normu pro fixovanou hodnotu řádu n , její hodnota je zjevně vyšší při nízké hustotě provozu a klesá s rostoucí hustotou provozu, což jsme pozorovali již na obrázku (6.4), resp. (6.5).



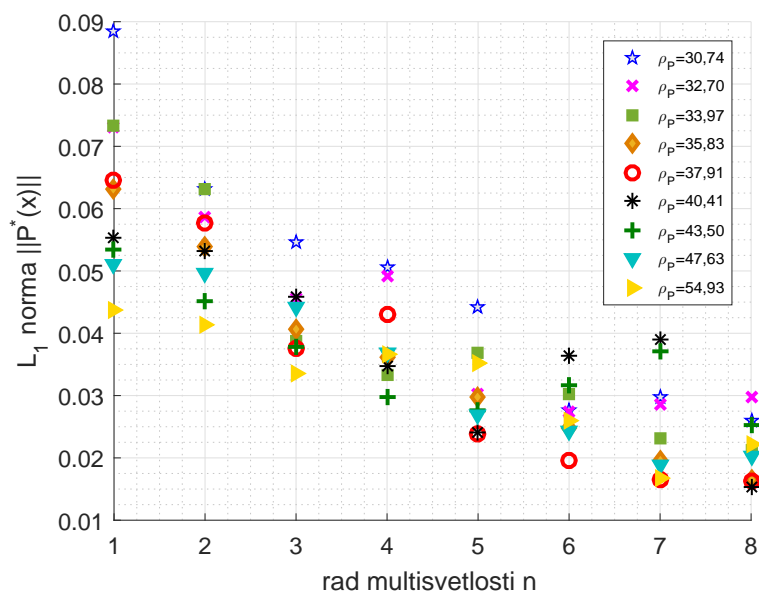
Obrázek 6.4: Závislost kolmogorovské normy $\|P^*(x)\|$ na hustotě ρ_P pro řád multisvětlosti $n = 1$ až $n = 4$, resp. $n = 5$ až $n = 8$



Obrázek 6.5: Závislost L_1 normy $\|P^*(x)\|$ na hustotě ρ_P pro řád multisvětlosti $n = 1$ až $n = 4$, resp. $n = 5$ až $n = 8$



Obrázek 6.6: Závislost kolmogorovské normy poruchové funkce $\|P^*(x)\|$ na řádu multisvětlosti n pro hustotní pásma $\rho_P = 30,74$ až $\rho_P = 54,93$



Obrázek 6.7: Závislost L_1 normy poruchové funkce $\|P^*(x)\|$ na řádu multisvětlosti n pro hustotní pásma $\rho_P = 30,74$ až $\rho_P = 54,93$

Závěr

Tato diplomová práce se zabývá modelováním distribucí fundamentálních dopravních mikrove-
ličin, světlostí vozidel. Klade si za cíl matematickými prostředky stanovit na základě znalosti
souboru světlostí dosah interakčních sil mezi vozidly pohybujícími se v reálném provozu. V úvodní
kapitole jsme se seznámili s matematickým aparátem, jehož znalost je nezbytná pro další odvo-
zování a výpočty. Definovali jsme obecný pojem hustoty a operaci konvoluce na třídě hustot a
sestavili souhrn poznatků z teorie funkcionálních konvolucí a integrálních transformací.

Ve druhé kapitole jsme navázali na tento přehled důležitými pojmy z teorie pravděpodobnosti
(náhodná veličina, hustota pravděpodobnosti) a analyzovali vlastnosti konvoluce v kontextu teorie
náhodných veličin. Seznámili jsme se s několika typy absolutně spojitých rozdělení a jejich vlast-
nostmi, přičemž důraz byl kladen zejména na třídu dvou- a tříparametrických GIG distribucí, pro-
tože tato třída se v současné době pro distribuci světlostí jeví jako nejvhodnější. Závěr kapitoly
byl věnován statistickým metodám odhadu parametrů rozdělení, které byly aplikovány v dalších
částech práce, tedy metodě minimální vzdálenosti a metodě maximální věrohodnosti. Pro účely
analýzy poruchové konvoluční funkce, již jsme se zabývali v poslední kapitole, jsme zde před-
stavili dvě funkcionální normy (kolmogorovskou a L_1 normu.)

Na začátku třetí kapitoly jsme se seznámili se základními mikroskopickými a makroskopický-
mi veličinami teorie dopravního proudění. Zavedli jsme klíčové pojmy škálované časové svět-
losti, dopravního toku a dopravní hustoty, jejichž prostřednictvím lze dopravní mikrostrukturu
popisovat, a jejich souvislost s přímo měřitelnými (primárními) dopravními veličinami. Kapitole
završuje obecný výklad fundamentálního diagramu a způsobů detekce dopravních dat včetně ná-
zorné ukázky měření dopravního proudu.

Na časovou světlost lze nahlížet jako na náhodnou veličinu s absolutně spojitým rozdělením,
dopravní struktura má zjevně stochastický charakter. K popisu dopravního systému lze tedy použít
hustotu pravděpodobnosti světlostí (tzv. headway distribuci). Ve čtvrté kapitole jsme proto zfor-
mulovali model dvouparametrického GIG rozdělení, které podle [17, 5, 6, 7] popisuje rozdělení
světlostí nejlépe. V prostředí MATLAB byl sestaven program, který pro zadaný soubor dat vy-
počítá odhad parametrů GIG distribuce dvěma různými metodami - metodou minimální vzdálenosti
a metodou maximální věrohodnosti. Pro obě metody jsme následně model prověřili na generátoru
pseudonáhodných čísel z dvouparametrické třídy GIG pomocí standardních statistických testů, tj.
Kolmogorova-Smirnovova testu a χ^2 -testu dobré shody. Měřili jsme procento zamítnutí nulové hy-
potézy, která tvrdí, že soubor generovaných pseudonáhodných čísel má rozdělení GIG. Došli jsme
mj. k závěru, že s rostoucí velikostí vzorku povolna vzrůstá počet případů, kdy je nulová hypotéza
zamítnuta.

V páté kapitole jsme výše popsany model aplikovali na dostupný vzorek dopravních dat (svět-
lostí) naměřených na Pražském okruhu a poskytnutých Ředitelstvím silnic a dálnic ČR. Nejprve
jsme diskutovali charakteristiku analyzovaných datových souborů a prověřili, že headway dis-
tribuci nejlépe odpovídá dvouparametrická GIG distribuce. Navrhli jsme originální metodiku pro
členění dat ze vstupního souboru a stanovili opět pomocí obou statistických metod relevantní
odhady parametrů. Model jsme prověřili obdobně jako v předchozí kapitole standardními statistick-

kými testy. Porovnáním obou užitých metod jsme došli k závěru, že metoda minimální vzdálenosti je odolnější vůči výraznějším výkyvům hodnot. V dalších výpočtech jsme proto od metody maximální věrohodnosti upustili. Následně jsme diskutovali závislost parametrů odhadované GIG distribuce na dopravní hustotě, z nichž nejzajímavější je závislost rezistivity na hustotě, která vykazuje pro rostoucí hodnoty hustoty téměř lineární nárůst. Výsledky této části práce posloužily jako základ připravovaného impaktovaného článku.

V šesté kapitole jsme využili všechny poznatky získané v předchozích kapitolách. Zavedli jsme veličinu multisvětlost, která popisuje odstupy dvou nesousedních vozidel. Dále jsme analyzovali statistické vlastnosti multisvětlostí v empirických datech naměřených na Pražském okruhu, přičemž jsme vycházeli z tříparametrické třídy GIG distribucí. Navrhli jsme metodiku pro výpočet multisvětlostí ze vstupního souboru a provedli statistické odhady parametrů pro rozdělení multisvětlostí. Opět jsme diskutovali závislost parametrů odhadované GIG distribuce na dopravní hustotě. Zajímavým výsledkem je, že pro daný řád hodnota rezistivity roste s rostoucí hustotou, a pro pevnou hodnotu hustoty hodnota rezistivity roste se zvyšujícím se řádem multisvětlosti. Závěrečná a nejzásadnější část práce je věnována konceptu poruchové konvoluční funkce coby efektivního nástroje ke stanovení dosahu interakčních sil v dopravních systémech. Na základě předpokladu, že světlosti jsou závislé náhodné veličiny, jsme zavedli tzv. poruchovou konvoluční funkci, v níž se odráží míra jejich závislosti. Je vyjádřena jako odchylka mezi hustotou pravděpodobnosti multisvětlosti a konvolucí hustot příslušných světlostí. Zvolili jsme dvě známé funkcionální normy, jejichž prostřednictvím jsme kvantifikovali informaci o závislosti náhodných veličin obsaženou v poruchové funkci.

Prozkoumali jsme empirické chování kolmogorovské i L_1 normy poruchové funkce. Zjistili jsme, že v obou případech hodnota normy u všech uvažovaných řádů multisvětlostí klesá s rostoucí dopravní hustotou. Mnohem významnější je však další zjištěný fakt, a to že norma poruchové funkce pro danou dopravní hustotu klesá s rostoucím řádem multisvětlosti. Závislost světlostí s rostoucím řádem multisvětlosti klesá, což odpovídá slabší interakci vzdálenějších vozidel. Stejného výsledku docílila i studie interakčního dosahu [20] založená na analýze trendu příslušných korelačních koeficientů. Tímto zjištěním jsme potvrdili, že koncept poruchové konvoluční funkce může být vhodným nástrojem pro stanovení interakčního dosahu vozidel.

Literatura

- [1] Li, L., Chen, X.M., 2017. *Transportation Research Part C* 76, 170.
- [2] Helbing, D., 2001. *Rev. Mod. Phys.* 73, 1067.
- [3] Treiber, M., Kesting, A., 2013. *Traffic Flow Dynamics*, Berlin: Springer.
- [4] Krbálek, M., Helbing, D., 2004. *Physica A* 333, 370.
- [5] Krbálek, M., 2007. *J. Phys. A: Math. Theor.* 40, 5813.
- [6] Krbálek, M., 2008. *J. Phys. A: Math. Theor.* 41, 205004.
- [7] Krbálek, M., Šleis, J., 2015. *J. Phys. A: Math. Theor.* 48, 015101.
- [8] Adams, W.F., 1936. *J. Inst. Civil Engin.* 4 (1), 121.
- [9] May, A.D., 1990. *Traffic Flow Fundamentals*, Prentice Hall, Englewood Cliffs, New Jersey.
- [10] Gerlough, D.L., 1955. *Eno Found. Highway Traffic Control*.
- [11] Greenberg, I., 1966. *Austral. Road Res.* 2 (7), 14.
- [12] Branston, D., 1976. *Transport. Sci.* 10 (2), 125.
- [13] Hoogendoorn, S., Bovy, P.H.L., 1998. *Transp. Res. Rec.* 1646, 18–28.
- [14] Krbálek, M., Šeba, P., Wagner, P., 2001. *Phys. Rev. E* 64 (6), 066119.
- [15] Jin, X., Zhang, Y., Wang, F., Li, L., Yao, D., Su, Y., Wei, Z., 2009. *Transport. Res. Part C: Emerg. Technol.* 17 (3), 318.
- [16] Krbálek, M., Šeba, P., 2009. *J. Phys. A: Math. Theor.* 42, 345001.
- [17] Krbálek, M., 2013. *J. Phys. A: Math. Theor.* 46, 445101.
- [18] Jörgensen, B., 1982. *Statistical properties of the generalized inverse Gaussian distribution*, Springer-Verlag.
- [19] Johnson, N.L., Kotz, S., Balakrishnan, N., 1994. *Continuous Univariate Distributions*, Vol. 1, Second Edition, John Wiley & Sons.
- [20] Krbálek, M., Apeltauer, J., Apeltauer, T., Szabová, Z., 2018. *Physica A* 491, 112.
- [21] Krbálek, M., 2014. *Teorie míry a Lebesgueova integrálu*, Česká technika – nakladatelství ČVUT, Praha.

- [22] Krbálek, M., 2009. *Matematická analýza IV (druhé rozšířené vydání)*, Česká technika – nakladatelství ČVUT, Praha.
- [23] Lerch, M., 1903. *Sur un point de la théorie des fonctions génératrices d'Abel*, Acta Mathematica, 27, 339-351.
- [24] Zvára, K., Štěpán, J., 2015. *Pravděpodobnost a matematická statistika*, MATFYZPRESS, Praha.
- [25] Wolfowitz, J., 1957. *The minimum distance method*, The Annals of Mathematical Statistics. 28 (1), 75-78.
- [26] Vacková, J., 2017. *Poruchová teorie statistické rigidity částicových systémů*, diplomová práce, FJFI ČVUT v Praze.
- [27] Szabová, Z., 2016. *Studie korelací v souborech empirických dopravních dat*, bakalářská práce, FJFI ČVUT v Praze.
- [28] Mikyška, J., 2008. *Asymptotické metody*, Česká technika – nakladatelství ČVUT, Praha.
- [29] Devroye, L., 2014. *Random variate generation for the generalized inverse Gaussian distribution*, Statistics and Computing, 24, 239-246.

Přílohy

q_p	MLE KS[%]	MLE χ^2_{k-p-1} [%]	MLE χ^2_{k-1} [%]	KD χ^2_{k-p-1} [%]	KD χ^2_{k-1} [%]
30,08	11,6	20,7	13,7	17,8	11,1
30,21	7,9	20,1	12,7	18,7	12,1
30,36	9,6	20,4	13,7	12,8	7,9
30,52	6,4	21,4	13,2	20,5	12,9
30,68	25,1	36,5	27,0	29,0	19,7
30,83	11,2	22,9	15,0	17,6	10,7
30,98	2,5	13,3	6,8	11,7	5,8
31,13	9,1	21,5	14,8	23,3	15,4
31,24	20,4	30,0	20,3	26,2	17,4
31,40	7,6	20,7	12,3	17,8	10,7
31,57	4,7	16,9	10,9	14,8	9,1
31,73	4,4	15,5	8,6	10,6	6,0
31,86	6,4	20,6	12,0	21,4	13,9
32,01	4,6	17,5	10,0	16,9	10,6
32,18	4,9	13,1	7,3	12,5	7,4
32,33	8,9	18,3	11,8	17,8	10,7
32,50	13,0	26,5	18,3	24,9	16,6
32,67	7,1	17,3	10,3	14,7	8,6
32,83	6,7	20,3	10,7	17,9	11,0
32,99	8,9	20,5	13,1	21,1	13,8
33,16	3,2	11,4	7,4	11,3	6,8
33,37	19,6	26,8	17,5	22,1	14,6
33,55	6,7	19,1	12,9	13,9	7,9
33,71	1,1	10,5	5,6	8,7	4,6
33,92	5,5	16,4	9,7	14,3	8,5
34,07	8,3	17,8	11,7	14,3	8,0
34,23	5,0	15,0	9,2	14,8	9,7
34,41	3,3	16,5	9,8	14,0	7,0
34,57	5,0	13,2	7,3	10,5	5,6
34,76	7,1	15,4	9,8	13,4	8,4
34,96	3,2	11,9	5,9	9,7	5,5
35,14	6,8	15,1	8,9	16,0	11,4
35,32	2,8	14,3	7,9	11,2	7,1
35,52	7,6	18,6	11,9	19,5	12,8
35,72	1,9	8,7	4,6	10,8	6,1
35,92	2,1	12,4	6,6	11,3	6,4
36,12	8,1	21,9	14,6	17,9	11,5
36,34	1,4	13,0	6,7	12,6	7,6
36,53	9,8	20,2	13,0	17,1	10,8
36,73	6,3	14,8	9,1	12,1	7,1
36,92	1,1	12,7	8,0	10,8	6,9
37,14	4,7	15,9	10,0	15,9	8,3
37,36	4,0	11,1	6,8	10,8	5,7
37,59	2,3	13,3	7,7	13,3	8,8
37,81	2,5	17,7	11,2	16,1	10,1
37,99	1,3	11,4	6,4	10,8	4,9

Tabulka 1: Procento zamítnutí pro test KS, χ^2_{k-p-1} , χ^2_{k-1} a velikost vzorku $n = 300$ v pásmu 1-46

Q_p	MLE KS[%]	MLE χ^2_{k-p-1} [%]	MLE χ^2_{k-1} [%]	KD χ^2_{k-p-1} [%]	KD χ^2_{k-1} [%]
38,20	1,9	13,6	7,9	12,9	7,0
38,45	1,5	10,6	6,3	9,6	5,2
38,69	11,1	20,6	13,3	18,5	11,4
38,92	2,0	14,4	8,6	13,3	8,7
39,15	0,8	10,7	5,6	10,4	5,2
39,43	5,9	18,6	11,0	13,5	8,6
39,67	1,1	11,8	6,4	10,4	5,1
39,94	2,4	10,5	5,9	9,5	6,1
40,26	3,0	13,6	7,6	13,6	7,7
40,52	3,9	17,1	9,5	15,8	9,5
40,79	1,0	8,0	3,8	7,7	3,9
41,12	3,7	17,1	10,5	16,4	10,9
41,46	4,5	13,4	8,3	12,9	8,3
41,75	1,5	12,1	7,4	13,0	8,8
42,04	0,6	8,9	4,7	8,1	4,4
42,34	2,3	12,0	6,7	13,0	6,9
42,66	2,8	9,4	5,9	7,6	4,4
43,00	0,8	7,0	3,6	7,9	4,6
43,28	2,0	11,6	6,6	9,9	5,2
43,59	2,0	11,2	5,9	10,0	5,1
43,94	0,8	10,4	5,7	10,3	5,7
44,32	0,6	8,0	3,9	6,4	3,6
44,67	1,4	8,7	4,7	9,5	5,1
45,12	1,9	9,7	6,0	8,7	5,4
45,56	5,5	22,1	13,9	17,3	10,5
45,99	8,1	22,1	14,7	17,8	11,8
46,45	1,4	12,0	7,4	12,1	7,2
46,92	1,6	11,0	6,0	8,6	4,7
47,35	27,3	44,1	33,6	8,2	4,5
47,76	1,7	8,9	5,1	6,9	4,1
48,21	1,0	12,7	7,0	10,3	6,4
48,75	1,0	8,3	4,7	8,1	4,4
49,32	1,1	10,4	5,3	9,1	5,0
49,99	1,2	6,5	3,4	5,7	2,2
50,60	0,7	8,0	4,4	7,4	4,2
51,21	0,5	9,0	4,6	8,0	4,8
51,93	1,9	12,1	6,2	9,3	5,6
52,80	1,3	11,3	6,1	10	5,3
53,72	0,4	7,8	3,6	7,2	4,1
54,74	2,3	15,2	8,9	12,7	7,2
56,07	1,3	8,7	4,8	9,7	5,5
57,55	1,6	11,9	6,9	11,6	6,4
59,23	2,5	13,1	7,7	10,7	6,2
61,40	1,5	11,5	5,3	9,6	5,2
65,40	1,4	7,1	4,6	8,2	4,6

Tabulka 2: Procento zamítnutí pro test KS, χ^2_{k-p-1} , χ^2_{k-1} a velikost vzorku $n = 300$ v pásmu 47-91

ϱ_p	MLE KS[%]	MLE χ^2_{k-p-1} [%]	MLE χ^2_{k-1} [%]	KD χ^2_{k-p-1} [%]	KD χ^2_{k-1} [%]
30,08	22,80	30,1	21,0	24,1	15,1
30,21	12,10	28,2	19,6	25,2	17,0
30,36	19,10	26,9	18,8	19,9	13,1
30,52	10,70	26,0	17,5	23,9	15,7
30,68	39,50	45,3	34,1	38,7	27,9
30,83	20,90	29,4	21,8	23,0	16,3
30,98	5,10	14,5	9,2	13,1	7,7
31,13	15,30	31,1	21,0	31,5	21,8
31,24	31,00	40,6	30,5	34,9	25,2
31,40	12,30	24,1	15,3	20,9	13,6
31,57	7,50	22,4	14,5	21,7	13,5
31,73	9,30	19,1	11,6	11,6	6,7
31,86	12,10	25,9	16,7	27,3	19,1
32,01	9,00	23,0	15,8	20,0	12,7
32,18	7,60	18,0	11,5	17,1	11,0
32,33	14,50	25,4	16,3	21,4	14,4
32,50	24,20	35,2	24,1	31,2	21,9
32,67	13,10	21,4	13,4	17,9	11,3
32,83	12,70	27,9	18,5	25,5	16,8
32,99	13,30	27,7	19,2	25,8	17,4
33,16	5,40	13,8	8,6	11,9	6,6
33,37	31,40	37,6	27,1	30,2	21,6
33,55	11,30	22,3	14,9	17,7	11,1
33,71	2,40	13,0	7,2	11,2	6,2
33,92	10,00	19,7	11,2	16,3	9,0
34,07	12,10	25,6	16,7	21,4	13,9
34,23	7,00	17,1	11,2	19,0	11,4
34,41	6,30	24,1	15,8	17,8	11,7
34,57	7,20	16,5	9,2	12,4	7,9
34,76	9,90	17,5	11,2	13,4	7,9
34,96	2,90	13,0	8,2	10,6	6,4
35,14	11,50	20,7	13,4	22,4	14,1
35,32	3,10	17,0	10,8	13,2	8,8
35,52	14,20	27,4	18,3	25,5	17,4
35,72	2,00	9,9	6,1	10,3	6,0
35,92	3,70	14,6	8,2	16,6	10,5
36,12	13,50	27,5	18,9	25,8	18,9
36,34	3,10	17,9	10,2	17,9	12,1
36,53	15,00	28,6	19,4	23,5	15,5
36,73	9,80	20,3	11,4	15,6	9,2
36,92	1,70	17,1	9,9	14,8	8,5
37,14	7,80	19,7	12,2	19,1	13,8
37,36	5,40	14,8	9,2	13,8	8,3
37,59	2,80	16,8	9,6	13,9	9,3
37,81	4,70	20,4	12,6	19,4	11,6
37,99	1,90	14,5	8,5	11,1	5,8

Tabulka 3: Procento zamítnutí pro test KS, χ^2_{k-p-1} , χ^2_{k-1} a velikost vzorku $n = 300$ v pásmu 1-46

q_p	MLE KS[%]	MLE χ^2_{k-p-1} [%]	MLE χ^2_{k-1} [%]	KD χ^2_{k-p-1} [%]	KD χ^2_{k-1} [%]
38,20	2,60	16,4	9,8	14,1	8,9
38,45	1,80	11,2	6,3	11,1	6,4
38,69	18,70	28,4	18,6	25,9	17,5
38,92	4,80	19,3	11,5	15,6	10,2
39,15	1,60	10,9	6,3	9,9	5,8
39,43	9,00	20,8	14,4	14,3	10,0
39,67	2,70	17,1	10,3	14,5	8,9
39,94	3,60	13,4	7,2	12,0	7,9
40,26	5,30	17,4	11,1	17,0	10,3
40,52	8,90	22,7	15	22,6	15,4
40,79	1,70	12,0	6,6	10,8	6,4
41,12	7,10	23,0	14,4	22,2	14,5
41,46	5,90	15,4	9,8	15,0	8,4
41,75	3,50	15,3	8,8	15,1	8,4
42,04	0,60	9,8	5,9	8,9	5,7
42,34	2,20	15,8	9,1	16,5	9,9
42,66	3,50	12,0	7,1	10,0	6,2
43,00	0,60	9,7	5,5	8,2	4,6
43,28	2,30	14,6	7,2	11,9	7,2
43,59	3,00	10,7	6,0	8,7	5,3
43,94	2,30	13,4	9,6	13,7	8,6
44,32	1,20	9,4	4,8	9,1	4,4
44,67	1,70	10,8	5,8	9,7	5,7
45,12	2,70	13,1	7,5	9,7	5,7
45,56	8,80	28,0	18,6	19,6	12,3
45,99	10,50	25,4	15,9	21,4	13,3
46,45	2,10	13,0	8,4	12,2	7,4
46,92	2,10	11,0	6,0	9,6	5,2
47,35	41,30	55,8	47	8,6	5,0
47,76	1,70	8,9	4,2	6,9	3,7
48,21	2,10	16,1	9,8	16,5	10,3
48,75	2,00	10,6	5,8	8,3	5,6
49,32	1,60	12,4	7,6	11,9	7,0
49,99	1,20	9,8	5,9	8,0	3,5
50,60	1,30	10,4	5,6	7,6	4,5
51,21	1,20	11,3	7,2	8,4	3,9
51,93	3,20	16,0	9,4	13,3	6,8
52,80	2,70	12,8	7,9	12,9	7,2
53,72	0,80	9,2	5,4	7,8	5,0
54,74	3,40	17,6	10,3	13,2	8,3
56,07	1,90	11,2	6,9	10,9	6,5
57,55	2,00	14,0	8,9	12,9	7,4
59,23	2,40	12,9	7,2	10,7	7,5
61,40	2,10	13,5	7,5	11,0	6,7
65,40	0,90	9,7	4,2	9,1	5,1

Tabulka 4: Procento zamítnutí pro test KS, χ^2_{k-p-1} , χ^2_{k-1} a velikost vzorku $n = 300$ v pásmu 47-91

q_p	MLE KS[%]	MLE χ^2_{k-p-1} [%]	MLE χ^2_{k-1} [%]	KD χ^2_{k-p-1} [%]	KD χ^2_{k-1} [%]
30,08	31,9	37,5	26,6	28,0	19,3
30,21	18,9	40,1	27,4	31,3	22,9
30,36	23,7	33,8	22,0	18,8	13,0
30,52	16,1	30,5	20,4	29,2	21,1
30,68	54,0	56,5	44,9	46,8	35,9
30,83	29,7	39,9	28,3	28,4	19,4
30,98	8,9	21,6	14,7	14,8	8,6
31,13	22,5	38,0	26,5	36,3	25,9
31,24	44,1	52,7	39,6	45,4	33,1
31,40	19,0	35,0	25,0	31,1	22,8
31,57	10,2	24,1	16,5	22,5	15,7
31,73	13,1	22,3	14,2	15,4	10,1
31,86	18,4	34,9	25,9	33,3	23,8
32,01	13,3	27,9	18,1	24,4	14,4
32,18	12,4	20,9	14,7	20,9	14,1
32,33	20,5	29,9	20,8	27,0	19,5
32,50	34,5	41,8	31,9	39,7	27,9
32,67	19,6	25,7	16,9	18,3	12,0
32,83	20,8	33,7	23,1	30,3	21,2
32,99	20,2	37,3	24,5	33,8	23,1
33,16	7,9	18,5	11,5	16,3	9,2
33,37	41,0	48,1	34,6	37,6	25,3
33,55	18,1	30,6	20,2	21,6	14,9
33,71	3,1	12,8	8,3	13,2	8,2
33,92	15,0	22,3	15,5	18,6	10,5
34,07	19,3	32,9	22,6	23,4	14,6
34,23	11,8	22,0	14,8	22,7	16,0
34,41	7,1	26,5	18,9	23,5	16,6
34,57	9,1	19,9	12,7	16,4	10,3
34,76	15,7	20,0	13,5	18,0	12,1
34,96	5,1	17,5	10,8	14,8	8,8
35,14	19,8	26,0	17,7	25,1	15,9
35,32	6,2	21,1	14,1	18,1	11,1
35,52	21,9	30,8	21,6	30,4	21,6
35,72	4,0	12,0	7,5	12,7	7,8
35,92	4,5	19,1	12,2	17,4	10,1
36,12	20,1	37,4	25,9	33,9	25,6
36,34	4,4	19,4	10,8	19,5	11,9
36,53	22,1	35,2	24,2	28,7	20,5
36,73	12,7	21,4	15,0	16,8	9,0
36,92	2,1	17,5	11,4	16,0	8,6
37,14	12,6	27,1	17,9	24,1	16,4
37,36	10,2	20,2	12,7	20,1	12,4
37,59	5,0	19,9	13,3	16,4	10,5
37,81	7,5	26,7	18,2	24,6	16,4
37,99	3,6	17,2	11,3	15,9	9,2

Tabulka 5: Procento zamítnutí pro test KS, χ^2_{k-p-1} , χ^2_{k-1} a velikost vzorku $n = 300$ v pásmu 1-46

q_p	MLE KS[%]	MLE χ^2_{k-p-1} [%]	MLE χ^2_{k-1} [%]	KD χ^2_{k-p-1} [%]	KD χ^2_{k-1} [%]
38,20	4,0	17,8	10,8	16,3	9,9
38,45	2,8	12,7	8,5	11,7	6,2
38,69	24,9	33,3	23,3	29,4	20,4
38,92	5,8	23,5	14,9	20,9	12,8
39,15	3,9	15,2	8,6	13,1	7,3
39,43	11,9	26,3	17,7	16,2	8,9
39,67	2,6	16,1	10,4	14,1	8,7
39,94	5,6	16,4	10,1	14,0	8,2
40,26	6,8	18,5	11,1	18,9	12,5
40,52	12,6	28,8	20,3	25,1	17,5
40,79	2,1	13,6	7,4	13,0	8,0
41,12	5,8	26,1	17,6	25,6	16,6
41,46	9,8	18,6	11,9	18,8	11,2
41,75	5,6	19,8	13,1	18,7	11,8
42,04	1,5	9,1	4,9	8,7	4,7
42,34	4,0	19,7	12,5	20,1	12,5
42,66	3,8	12,4	6,6	11,1	6,8
43,00	0,3	10,4	5,2	10,9	5,7
43,28	3,5	15,4	9,4	14,3	9,0
43,59	3,1	11,0	6,3	9,5	5,3
43,94	2,1	16,5	9,0	13,9	7,7
44,32	1,3	7,2	4,1	7,2	4,1
44,67	1,0	10,5	6,8	10,5	5,4
45,12	3,3	12,5	7,0	11,8	6,0
45,56	15,4	38,7	28,7	28,2	20,1
45,99	14,1	29,9	20,4	26,8	19,7
46,45	1,9	15,5	10,8	13,9	7,5
46,92	4,2	13,1	9,1	11,7	7,2
47,35	50,5	67,0	57,8	9,9	5,5
47,76	1,9	11,4	6,3	8,2	4,2
48,21	1,9	18,3	11,3	18,0	11,2
48,75	2,6	10,9	6,3	9,2	5,9
49,32	3,3	16,0	9,9	15,9	9,5
49,99	4,0	12,7	7,0	9,5	5,6
50,60	1,9	12,4	8,7	9,2	5,5
51,21	1,1	9,2	5,4	8,9	4,2
51,93	5,5	17,0	11,4	14,2	8,8
52,80	2,4	13,1	8,2	13,0	7,6
53,72	1,0	7,8	3,9	7,4	3,9
54,74	5,5	20,7	13,8	16,3	9,7
56,07	2,2	12,6	6,8	12,2	6,5
57,55	4,2	17,6	9,8	14,8	9,8
59,23	4,2	16,9	11,2	17,1	9,6
61,40	3,1	17,6	10,0	14,4	9,1
65,40	2,3	10,8	6,6	8,7	4,0

Tabulka 6: Procento zamítnutí pro test KS, χ^2_{k-p-1} , χ^2_{k-1} a velikost vzorku $n = 300$ v pásmu 47-91