České vysoké učení technické v Praze
Fakulta jaderná a fyzikálně inženýrská

Statistické modelování dopravního toku na neřízené křižovatce typu T

Statistical modelling of traffic flow at unsignalized T intersection

Research project

Author: **Bc. Daniel Wohlrath**

Supervisor: **doc. Ing. Tomáš Hobza, Ph.D.**

Advisor: **doc. Mgr. Milan Krbálek, Ph.D.**

Academic year: 2020/2021

Zadání práce

Zadání práce (zadní strana)

Poděkování:
Rád bych zde poděkoval svému školiteli doc. Ing. Tomášovi Hobzovi, Ph.D. za jeho ochotu, vstřícnost, cenné rady, věcné přípomínky a odborné i lidské zázemí při vedení mé práce. Dále děkuji svému konzultantovi doc. Mgr. Milanu Krbálkovi, Ph.D. za jeho ochotu a poskytnuté materiály.

Čestné prohlášení:
Prohlašuji, že jsem tuto práci vypracoval samostatně a uvedl jsem všechnu použitou literaturu.

V Praze dne 13. listopadu 2023 Bc. Daniel Wohlrath

Název práce:

Statistické modelování dopravního toku na neřízené křižovatce typu T

Autor: Bc. Daniel Wohlrath

Obor: Matematické inženýrství

Druh práce: Výzkumný úkol

Vedoucí práce: Doc. Ing. Tomáš Hobza, Ph.D., Katedra matematiky, Fakulta jaderná a fyzikálně inženýrská, České vysoké učení technické v Praze, Trojanova 13, 12000 Praha 2

Konzultant: Doc. Mgr. Milan Krbálek, Ph.D., Katedra matematiky, Fakulta jaderná a fyzikálně inženýrská, České vysoké učení technické v Praze, Trojanova 13, 12000 Praha 2

Abstrakt: Tato práce se zabývá statistickým modelováním kritických světlostí, které jsou hlavním předmětem teorie akceptování mezer, neboli Gap Acceptance. Nejdříve jsme definovali základní matematický model neřízené křižovatky typu T s příslušnou terminologií. Pomocí zavedeného modelu jsme specifikovali problém dílčí distribuce světlostí řádu $k \in \mathbb{N}_0$. Pro gamma rozdělení světlostí a kritických světlostí jsme odvodili řešení tohoto problému a jeho správnost následně numerickými výpočty potvrdili. Pro gamma rozdělené kritické světlosti jsme dále odvodili tvar Sieglochovy funkce a graficky ji znázornili. Na závěr jsme pro empirická data ze tří křižovatek v Německu statisticky a graficky ověřili hypotézu, jestli empirická dílčí distribuce řádu $k \in \{0, 1, 2, 3\}$ pochází z rodiny námi odvozeného rozdělení. K tomu jsme použili Pearsonův χ^2 test dobré shody.

Klíčová slova: časová světlost, dílčí distribuce světlostí, kapacita křižovatky, kritická časová světlost, Pearsonův χ^2 test dobré shody, Sieglochova funkce

Title:

Statistical modelling of traffic flow at unsignalized T intersection

Author: Bc. Daniel Wohlrath

Abstract: This paper deals with statistical modelling of critical clearances which are the main subject of the Gap Acceptance theory. First, we used appropriate terminology to define a mathematical model of an unsignalized T intersection. Using this model we presented the problem of the partial distribution of clearances of order $k \in \mathbb{N}_0$. Assuming gamma distribution of clearances and critical clearances we derived a solution to this problem. We later validated this solution by using numerical computations. Further, with the premise of gamma distributed critical clearances, we analytically derived the Siegloch function. At last, we verified that the empirical partial distribution of clearances of order $k \in \{0, 1, 2, 3\}$ (recorded at three German T intersections) belongs to the family of previously derived partial distribution of clearances. This hypothesis was evaluated by using Pearson's χ^2 goodness of fit test.

Key words: capacity at unsignalized intersections, critical time clearance, partial distribution of clearances, Pearson's χ^2 test, Siegloch function, time clearance

Obsah

Preface

Uncontrolled intersections are the most common type of intersections. Although their capacity is often lower than that of modern solutions, they play an important role in transportation networks. Therefore, it is very important to analyze their characteristics in detail, which primarily includes the capacity of the intersection.

The pioneer in the theory of intersection capacity is W. Siegloch, who developed a method for their calculation. His method is based on the occurrence of time gaps between vehicles on the main road and critical time gaps of drivers who want to enter the intersection from the side road. Critical time gaps are considered random variables and it is useful to estimate the parameters of their distribution using statistical methods, most often the mean and standard deviation.

One of the goals of this thesis was to introduce in detail a mathematical model of a T type intersection and to derive the partial distribution of clearances of order $k \in \mathbb{N}_0$, which corresponds to the probability distribution of time gaps on the main road that were used by exactly k drivers from the side road to turn.

Another goal was to derive the Siegloch function $s(t)$ for the considered model of the intersection, which features in Siegloch's formula for calculating capacity and is often approximated by its linear asymptote in practice.

The last goal of this work was to analyze empirical data from three German intersections and verify the hypothesis whether the observed luminances of the kth order for $k \in \{0, , 1, , 2, , 3\}$ have a distribution given by our derived partial distribution of luminances of the same order. This verification was performed using Pearson's χ^2 test.

Kapitola 1

Theory of Gap Acceptance

At an unsignalized intersection, the general rule of priority applies to vehicles on the main road. Drivers on side roads do not receive any signals when to enter the intersection. Therefore, they must decide for themselves when it is safe to enter the intersection area without endangering or restricting drivers on the main roads. The process by which a driver looks for a gap between vehicles on the main road, the so called *gap*, to safely enter the intersection is the main subject of gap acceptance theory, or *Gap Acceptance*[1], whose application is particularly significant in the analysis of unsignalized intersections.

Similar processes, where pedestrians or drivers have to decide to use a gap in the main road, occur in various traffic situations. Other examples include:

- crossing a road outside a crosswalk,
- overtaking on country roads where there is a risk of collision with oncoming vehicles,
- entering a roundabout,
- merging from an on ramp onto a highway.

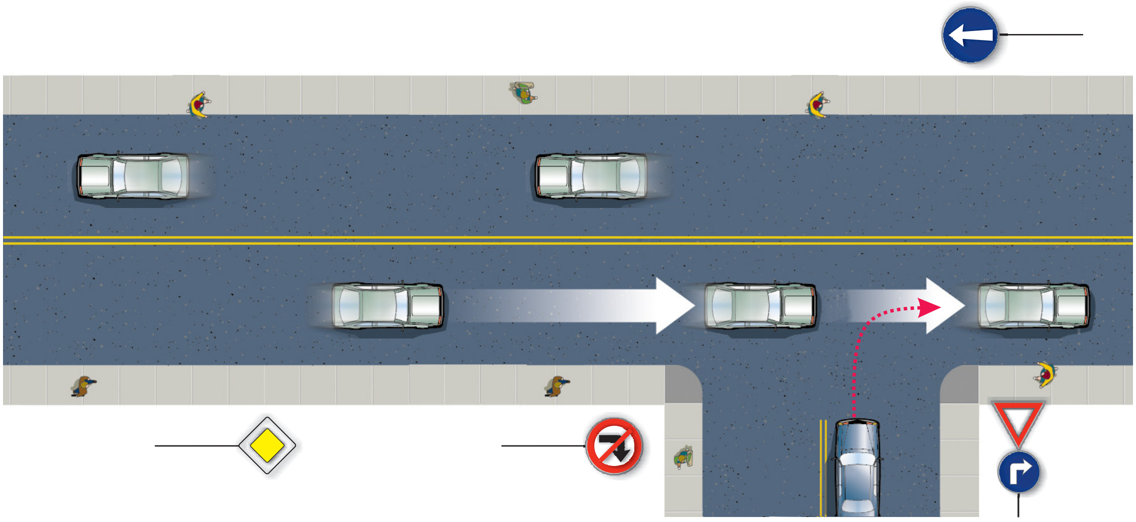
While each of these situations poses its own special problem, they have a very similar basic idea. In each case, there is a certain time t_0 when the driver[2] begins to search among the available gaps on the main road for a safe opportunity to turn. Then vehicles from the main road gradually pass through the intersection. If we denote the times of crossing the imaginary line of entry into the intersection by the front, respectively, the rear bumper of the i th vehicle as t_i^{in}, respectively, t_i^{out}, where $i \in \mathbb{N}$, then the created time intervals $t_{i+1}^{\text{in}} - t_i^{\text{out}}$, $\forall i \in \mathbb{N}$ are called **time clearances** of vehicles on the main road, or briefly clearances. For each driver looking for an acceptable clearance on the main road, there is a set of unused clearances and a clearance used. The smallest time clearance that a specific driver can use is called (individual) **critical time clearance**. If a driver on a side road uses a clearance on the main road to perform a maneuver, we say that the clearance was accepted, or also, that the driver accepted the given clearance.

1.1 Considered Mathematical Model

For a closer study of critical time clearances, let us introduce a simple model of an unsignalized T intersection, which is illustrated in Figure 1.1.

[1]hereafter GA

[2]for clarity, let us assume that we analyze the original problem at an unsignalized intersection



Obrázek 1.1: Illustration of an unsignalized T intersection. The driver with the vehicle on the side road tries to fit between the moving vehicles on the main road. [Illustration by: Milan Krbálek]

Vehicles from the main road pass through the intersection in time clearances, which we consider to be random variables. In accordance with the scientific discipline VHM[3] we assume their independence and identical distribution[4] described by the probability density f_X. Individual time clearances are therefore specific i.i.d. realizations of the random variable $X \sim f_X$. On the side road, there is a queued line of drivers ready to turn and merge between two vehicles on the main road. The driver will make the maneuver when the available time clearance on the main road is greater or equal to his individual critical time clearance.

It is evident that the critical time clearance may vary for each driver at the given intersection. Indeed, consider for example, that for an experienced driver who knows the intersection well, the critical clearance will be very short compared to an inexperienced driver. At the same time, it can be assumed that the critical clearance of a driver will be different if the road is wet or dry. Therefore, from the perspective of probability theory, we consider it to be a random variable with density f_Y, and thus the individual critical clearance is a realization of the random variable $Y \sim f_Y$.

After the first driver with an individual critical clearance y_1, who accepted the clearance $x_1 \geq y_1$, another driver is prepared to enter the intersection, comparing the available clearances on the main road with his critical clearance y_2. The second driver thus first compares the "remaining part"of the clearance used by the first driver with his critical clearance y_2 and either accepts it or waits for another clearance to be created.

It is important to define this *remaining part of the clearance* in a suitable way. In reality, the second vehicle cannot enter the intersection shortly after the first vehicle turns. This fact arises, for example, from non zero lengths of cars and safe distances between vehicles on the side road. In this work, we will define the **remaining time clearance** as the difference between the original clearance and the critical clearance of the preceding vehicle. This approach is considered a primary approximation, however, in a more comprehensive approach, the literature introduces the **following time gap**, which is the time interval between individual entries of vehicles from the side road into the same gap on the main road. The following time gap is then also of a stochastic nature and represents in the calculations of intersection capacity[5] a similar role as the critical clearance. In this work, however, we will focus only on

[3] Vehicular Headway Modeling

[4] independence and identical distribution of random variables are further denoted as i.i.d.

[5] the concept of intersection capacity is defined in section 1.2

critical clearances, so we will adhere to the definition given above. The study of subsequent time gaps is addressed, for example, in [1].

Note for completeness that in the presented model we assume that drivers are so called consistent and homogeneous and the intersection is in the so called **saturated state**. A consistent driver with an individual critical clearance y will never use a clearance $x < y$ to turn and will always use a clearance $x \geq y$. A hypothetical population of drivers is homogeneous if each subset of drivers has the same distribution of critical clearances f_Y. The case when a queue is created on the side road is called the **saturated state** of the intersection. This term was first introduced by W. Siegloch and is almost without exception considered in every practical measurement of critical clearances, see [7].

The assumptions of consistency and homogeneity are certainly unrealistic, but very important for analytical derivation. In the work [9] it was proved that if drivers were inconsistent, the capacity of the intersection would increase, and on the other hand, if the population of drivers was not homogeneous, the capacity would decrease. If we considered inconsistent and non homogeneous drivers, according to [9] the resulting difference between the models is minimal, so for simplicity, we will continue to consider the generally accepted assumption that drivers are consistent and homogeneous.

1.2 Intersection Capacity and Estimation of Critical Clearances

The main source of motivation for studying critical clearances is the calculation of intersection capacity. In practice, within the theory of GA, according to [8], the capacity of an intersection is defined by one of the following methods, which, however, converge within the model of this work:

- the capacity of the intersection is the sum of the intensities[6] of all subordinate traffic flows entering the intersection per hour,
- the capacity of the intersection is given by the capacity of the most subordinate entry.

Siegloch formulated in 1973 a formula for calculating the capacity of a subordinate entry, which coincides with the calculation of the intersection capacity for the T intersection model. Equation (1.2) represents the foundation of the entire GA theory. Almost all analytically derived equations for capacity estimates are based on a similar concept, see [3].

Intersection Capacity Equation

Let $s(t)$ be the mean number of vehicles from the secondary road that accept a clearance of size t. The function $s(t)$ is referred to as the Siegloch function and is discussed in more detail in Section 2.3 in Chapter 2. Retaining the notation for the probability density of time clearances f_X from the previous section, according to Siegloch, the capacity increase corresponding to clearances of length t is equal to

$$q\, f_X(t)\, s(t)\,, \tag{1.1}$$

where q is the traffic intensity on the main road in units of vehicles per hour. To obtain the total capacity, expression (1.1) must be integrated over all possible lengths of time clearances. Overall, we obtain the following relationship for the capacity of the intersection c

$$c = q \int_0^{+\infty} f_X(t)\, s(t)\, \mathrm{d}t\,. \tag{1.2}$$

[6]traffic flow intensity is the number of vehicles passing a given point per unit of time

Some authors consider the following assumptions for the simple calculation of intersection capacity:

- Critical clearances and subsequent gaps are constants,
- Clearances on the main road are exponentially distributed with parameter $\lambda > 0$,
- The traffic intensity q is constant.

However, these idealized assumptions are in most cases provably a poor approximation and are studied in [2]. For more general results, it is therefore necessary to take into account a more general distribution of clearances and the stochastic nature of both critical clearances and subsequent gaps.

Estimating Critical Clearances

From the nature of the introduction of the Siegloch function $s(t)$, it follows that it depends on the distribution of critical clearances[7]. Therefore, for the practical calculation of intersection capacity, it is necessary to estimate the parameters of their distribution from measured data, which usually involve the mean and standard deviation. While subsequent gaps are easily observable at a saturated intersection, critical clearances are not directly measurable. Assuming driver consistency, the measurements only tell us that the driver's individual critical clearance is greater than the largest unaccepted clearance and smaller than the one accepted.

Currently, there are many different techniques for estimating these parameters. A detailed comparison of nine different methods is presented in [5], where it was shown that the maximum likelihood method[8] is the most suitable method in many aspects[9].

The disadvantage of the MLE method is the need to assume a specific type of distribution for critical clearances f_Y. According to [5], the estimates are often of good quality even though the specific assumed distribution is unrealistic. Simple and frequently used choices, see for example [6], are exponential or shifted exponential distributions, whose densities are in the form

$$h(x) = \lambda \mathrm{e}^{-\lambda x}, \quad \text{for } x > 0, \quad \text{respectively,} \tag{1.3}$$

$$g_L(x) = \lambda \mathrm{e}^{-\lambda (x-L)}, \quad \text{for } x > L. \tag{1.4}$$

These choices are again too trivial in the second approximation. The distributions only take positive values, but logically do not capture the very low, non zero probability of the occurrence of a driver with arbitrarily small critical clearance. Therefore, theoretically acceptable distributions include well known log normal or gamma distributions. The Generalized Inverse Gaussian (GIG)[10] distribution is also a relevant option, but the problem of the most suitable choice of distribution is still open to further studies.

1.3 Modeling Critical Clearances

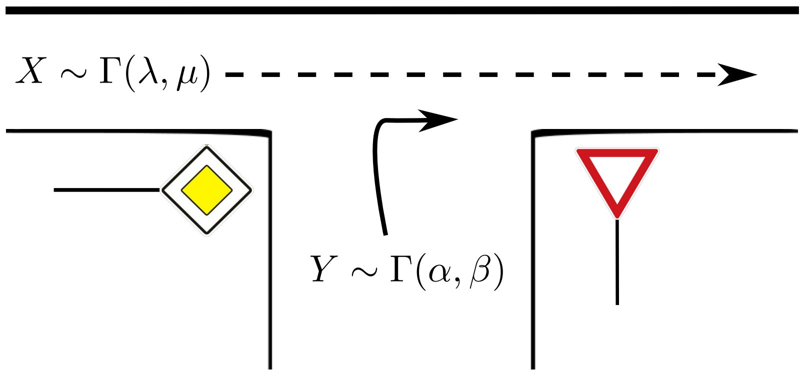
As already outlined in this chapter, it is very important to correctly and accurately estimate the parameters of the distribution of critical clearances, and a suitable method for this is, for example, MLE. Therefore, we will complement the mathematical model of the intersection, which was introduced in Section 1.1, with the assumption of specific types of distributions for time and critical time clearances.

[7] generally also depends on subsequent gaps

[8] hereinafter referred to as MLE

[9] such as estimate consistency, as well as the method's insensitivity to the types of assumed distributions for clearances and critical clearances

[10] generalized inverse Gaussian distribution



Obrázek 1.2: Time clearances on the main road are realizations of a gamma distributed variable X with parameters λ, μ and critical time clearances of drivers on the secondary road are gamma distributed with parameters α, β.

As a generalization of the exponential distribution previously inappropriately used[11], we will assume a gamma distribution for both time and critical time clearances, whose probability density function for parameters α, $\beta > 0$ is in the form

$$f_{Gamma}(y) = \frac{\beta^\alpha}{\Gamma(\alpha)} y^{\alpha-1} \exp(-\beta y)\,, \quad \text{for}\, y > 0\,, \tag{1.5}$$

where $\Gamma(y)$ is the gamma function. Note that for a special choice of parameter $\alpha = 1$, the expression (1.5) transitions to the density of the exponential distribution (1.3), thus it truly acts as a generalization of the existing model used in the literature. However, there is still room in the future for generalizing the shapes of densities for time and critical time clearances that would better reflect reality. A relevant option might be the aforementioned GIG distribution.

For clarity of the above assumptions, a simplified version of the T intersection is shown in Figure 1.2, where the gamma distribution with parameters α, β is abbreviated as $\Gamma(\alpha,\beta)$.

Partial Distributions of Clearances

In GA theory, it is a fundamental assumption that a single time clearance can be used for turning by several vehicles from the secondary road. According to our established model, a gap corresponding to the time clearance x will be accepted by exactly $k \in \mathbb{N}_0$ vehicles, when

$$\sum_{i=1}^{k} y_i \le x \quad \wedge \quad \sum_{i=1}^{k+1} y_i > x\,, \tag{1.6}$$

where $y_1, y_2, \ldots, y_{k+1}$ are the critical clearances of the first $k+1$ drivers. Each time clearance can then be assigned a number $k \in \mathbb{N}_0$, which corresponds to the number of vehicles that have accepted it, and is called the **acceptance order of the clearance**. The set of all time clearances on the main road can then be divided into disjoint subsets $\{\, A_k \,|\, k \in \mathbb{N}_0 \,\}$ according to the acceptance order. In each set A_k, there are only those time clearances that were accepted by exactly k vehicles, i.e., they are realizations of the conditional random variable X assuming the acceptance of exactly k i.i.d. variables Y_i, $i \in \{1, \ldots, k\}$, thus

$$D^{(k)} := X \,\middle|\, \left(\sum_{i=1}^{k} Y_i \le X \wedge \sum_{i=1}^{k+1} Y_i > X \right), \tag{1.7}$$

[11] see Section 3.1 and Figure **??**

where clearances, or critical clearances, are random variables $\{ X_i \,|\, i \in \mathbb{N} \} \overset{i.i.d.}{\sim} f_X$, respectively $\{ Y_i \,|\, i \in \mathbb{N} \} \overset{i.i.d.}{\sim} f_Y$, where f_X and f_Y are given by the prescription (1.5) for generally different parameters λ, μ and α, β. The distribution of the variable $D^{(k)}$ is called for a given $k \in \mathbb{N}_0$ the **partial distribution of clearances of the k th order** and is denoted $D^{(k)} \sim f_{D^{(k)}}$. Specifically for $k = 0$, A_0 is the set of clearances that were not accepted by any vehicle and by substituting into (1.7) then for the variable $D^{(0)}$ it holds

$$D^{(0)} = X \,|\, (Y > X) \; .$$

The realizations of the variable $D^{(k)}$, unlike the realizations of the variable Y, are directly measurable in practice because it is only necessary to record the number of vehicles that accepted each observed clearance. Moreover, these partial distributions of any order are certainly fully characterized by the distribution parameters of X and Y and by the acceptance order k, therefore if we had the form of the density $f_{D^{(k)}}$ available, we could estimate its parameters in practice using standard statistical methods. With these estimates, it is then possible to express retrospectively an estimate for the expected value $\mathrm{E}[Y]$ or the variance $\mathrm{Var}[Y]$ of the critical clearance at a given intersection. For the parameterization of the gamma distribution see (1.5) and for the designation of the distribution parameters of critical clearances $Y \sim \Gamma(\alpha, \beta)$ it holds specifically

$$\mathrm{E}[Y] = \frac{\alpha}{\beta} \, , \quad \mathrm{Var}[Y] = \frac{\alpha}{\beta^2} \, .$$

Kapitola 2

Partial Distributions of Order k and Siegloch Function

In this chapter, we first derive the form of the density of partial distributions of clearances of general kth order for $k \in \mathbb{N}_0$. For empirical data from measurements at intersections, it is possible to estimate the parameters of this distribution and express from them estimates of the parameters of the distribution of critical clearances in the manner described in section 1.3. While the estimation of the parameters of the distribution of partial distributions is discussed in more detail in chapter 3, the reverse expression of the estimated parameters of the distribution of critical clearances is not the subject of this thesis.

Furthermore, in this chapter, we adjust the general form of the Siegloch function to a more suitable form for practical calculations, and specifically for our gamma distribution of clearances and critical clearances, this calculation will be carried out.

Throughout this chapter, let for any $k \in \mathbb{N}_0$, α, β, $\lambda, \mu > 0$ independent random variables represent

$$X \sim \Gamma(\lambda,\mu),$$
$$Y_i \sim \Gamma(\alpha,\beta)\,, \quad \forall\, i \in \{1,\ldots,k+1\}$$

the time clearance, the first k critical clearances, whose drivers accept clearance X and the critical clearance of the $(k+1)$ th driver, who no longer accepted clearance X. The density of the gamma distribution Γ should be given for certainty by the formula (1.5).

2.1 Distribution of clearances Accepted by $k \in \mathbb{N}_0$ Vehicles

The next goal will be to derive the probability density of the partial distribution of clearances of order k, that is (see (1.7)) the density $f_{D^{(k)}}$, which corresponds to the density of the distribution of clearance X accepted by exactly $k \in \mathbb{N}_0$ i.i.d. critical clearances Y_i.

Transformation of Variables

For clarity of the notation, let us first assume $k \in \mathbb{N}$. Let us introduce an auxiliary random variable

$$Z := \sum_{i=1}^{k} Y_i \sim \bigstar_{i=1}^{k} f_Y(z)\,, \tag{2.1}$$

whose distribution, due to the independence of the variables Y_i, $i \in \{1, \ldots, k\}$, is given by k iterations of convolution of density (1.5). The specific choice of distributions for Y_i allows us to determine from the reproductive property of the gamma distribution that $Z \sim \Gamma(k\alpha,\beta)$.

Let us use a regular transformation $h : \mathbb{R}_+^3 \to \mathbb{R}_+ \times \mathbb{R}^2$ specified by:

$$h : \begin{pmatrix} X \\ Z \\ Y_{k+1} \end{pmatrix} \mapsto \begin{pmatrix} U \\ V \\ W \end{pmatrix} = \begin{pmatrix} X \\ X - Z \\ Z + Y_{k+1} - X \end{pmatrix}, \tag{2.2}$$

whose determinant of the matrix of first derivatives, or Jacobian, is equal to 1. By using the theorem on the transformation of densities of random variables and assuming the independence of the variables X, Z, Y_{k+1}, we obtain the form of the joint density of the transformed variable as

$$f_{U,V,W}(u, v, w) = f_X(u) f_Z(u - v) f_Y(v + w)$$

for $u > 0$, $u - v > 0$ and $v + w > 0$.

In the new variables U, V, W, the condition of conditioning in the formula (1.7) is met when $V > 0$ and at the same time $W > 0$. Therefore, we will look for the conditional density $f_{U|V>0,W>0}(u)$:

$$\begin{aligned} f_{U|V>0,W>0}(u) &= \frac{\mathrm{d}}{\mathrm{d}u} F_{U|V>0,W>0}(u) = \frac{\mathrm{d}}{\mathrm{d}u} \mathsf{Pr}\left[U \le u \,|\, V > 0, W > 0\right] \\ &= \underbrace{\frac{1}{\mathsf{Pr}\left[V > 0, W > 0\right]}}_{\text{denoted as } N} \cdot \underbrace{\frac{\mathrm{d}}{\mathrm{d}u} \mathsf{Pr}\left[U \le u, V > 0, W > 0\right]}_{\text{denoted as } M} . \end{aligned} \tag{2.3}$$

Adjustment of the expression M

Using Fubini's theorem and subsequent differentiation of the integral as a function of its upper limit, we modify the expression M as seen in (2.3):

$$\begin{aligned} M &= \frac{\mathrm{d}}{\mathrm{d}u} \int_0^u \mathrm{d}\tilde{u} \int_0^{+\infty} \mathrm{d}v \int_0^{+\infty} \mathrm{d}w \, f_{U,V,W}(\tilde{u}, v, w) \\ &= \int_0^{+\infty} \mathrm{d}v \int_0^{+\infty} \mathrm{d}w \, f_{U,V,W}(u, v, w) \qquad (2.4) \\ &= \int_0^u \mathrm{d}v \int_0^{+\infty} \mathrm{d}w \, f_X(u) f_Z(u - v) f_Y(v + w) \\ &= f_X(u) \int_0^u \mathrm{d}v \, f_Z(u - v) \underbrace{\int_0^{+\infty} \mathrm{d}w \, f_Y(v + w)}_{\text{denoted as } \mathcal{I}} . \qquad (2.5) \end{aligned}$$

First, we compute the above denoted integral $\mathcal{I}$ over the variable w:

$$\int_0^{+\infty} f_Y(v + w) \, \mathrm{d}w = \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^{+\infty} (v + w)^{\alpha - 1} \exp\left(-\beta(v + w)\right) \, \mathrm{d}w \, .$$

Now use the binomial theorem for the expression $(v + w)^{\alpha-1}$, which holds for all nonnegative integer values of the exponent $\alpha - 1$. Therefore, we further restrict ourselves only to $\alpha \in \mathbb{N}$, i.e., let the variables

Y_i have Erlang distribution with parameters α, β and continue with the modifications of $\mathcal{I}$:

$$\begin{aligned}
\int_0^{+\infty} f_Y(v+w)\,\mathrm{d}w &= \frac{\beta^\alpha e^{-\beta v}}{\Gamma(\alpha)} \sum_{j=0}^{\alpha-1} \binom{\alpha-1}{j} v^{\alpha-1-j} \underbrace{\int_0^{+\infty} w^j e^{-\beta w}\,\mathrm{d}w}_{=\Gamma(j+1)/\beta^{j+1}} \\
&= e^{-\beta v} \sum_{j=0}^{\alpha-1} \frac{(v\beta)^{\alpha-1-j}}{(\alpha-1-j)!} \\
&= e^{-\beta v} \sum_{j=0}^{\alpha-1} \frac{(v\beta)^j}{j!}\,.
\end{aligned}$$

Substituting back into equation (2.5) we obtain:

$$\begin{aligned}
M &= f_X(u) \sum_{j=0}^{\alpha-1} \frac{\beta^j}{j!} \int_0^u f_Z(u-v) e^{-\beta v} v^j\,\mathrm{d}v \\
&= f_X(u) e^{-\beta u} \sum_{j=0}^{\alpha-1} \frac{\beta^j}{j!} \frac{\beta^{k\alpha}}{\Gamma(k\alpha)} \underbrace{\int_0^u (u-v)^{k\alpha-1} v^j\,\mathrm{d}v}_{\text{denoted as } \mathcal{J}}\,. \qquad (2.6)
\end{aligned}$$

To solve the integral $\mathcal{J}$ we will use the fact that for $a \neq 0$ and $m, n \in (0, +\infty)$ the following holds

$$\begin{aligned}
\frac{1}{a^{m+n-1}} \int_0^a y^{m-1}(a-y)^{n-1}\mathrm{d}y &= \frac{1}{a^{m+n-1}} \int_0^a a^{m+n-2} \left(\frac{y}{a}\right)^{m-1} \left(1-\frac{y}{a}\right)^{n-1} \mathrm{d}y \\
&= \frac{1}{a} \int_0^a \left(\frac{y}{a}\right)^{m-1} \left(1-\frac{y}{a}\right)^{n-1} \mathrm{d}y \qquad (2.7) \\
&= \int_0^1 x^{m-1}(1-x)^{n-1}\mathrm{d}x \\
&= \mathrm{B}(m,n)\,, \qquad (2.8)
\end{aligned}$$

where in step (2.7) we performed the substitution $y/a = x$ and $\mathrm{B}(m,n)$ denotes the beta function at the point (m,n). Then, by applying this identity, we modify the expression $\mathcal{J}$:

$$\int_0^u (u-v)^{k\alpha-1} v^j\,\mathrm{d}v = u^{j+k\alpha} B(j+1, k\alpha)\,, \qquad (2.9)$$

where, with the same notation as in (2.8), we chose

$$a = u, \quad y = v, \quad m = j+1, \quad n = k\alpha\,.$$

By substituting the result (2.9) into (2.6) and using the wellknown relation $\mathrm{B}(p,q) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}$ we obtain

$$M = f_X(u) e^{-\beta u} \frac{\beta^{k\alpha}}{\Gamma(k\alpha)} \cdot \sum_{j=0}^{\alpha-1} \frac{\beta^j}{j!} \frac{\Gamma(k\alpha)\Gamma(j+1)}{\Gamma(j+k\alpha+1)} u^{j+k\alpha}\,.$$

Since during the calculation we limited ourselves only to positive integer values of the parameter α, we can continue to modify using $\Gamma(\alpha) = (\alpha-1)!$

$$M = f_X(u) e^{-\beta u} \sum_{j=0}^{\alpha-1} \frac{(u\beta)^{k\alpha+j}}{(k\alpha+j)!}$$

and by substituting the formula for the density of the random variable X:

$$M = \frac{\mu^\lambda}{\Gamma(\lambda)} u^{\lambda-1} e^{-(\beta+\mu)u} \sum_{j=0}^{\alpha-1} \frac{(u\beta)^{k\alpha+j}}{(k\alpha+j)!} \,.$$

Expression Adjustment N

Let us now adjust the expression N designated in equation (2.3). If we express the marginal density $f_{V,W}$ as an integral of the joint density $f_{U,V,W}$ over all possible values of the variable U, i.e.,

$$f_{V,W}(v,w) = \int_0^{+\infty} f_{U,V,W}(u,v,w)\,\mathrm{d}u\,,$$

then it is clear from the relationship (2.4) that the reciprocal value $1/N$ is obtained by integrating the expression M with respect to the variable u over the interval $(0,+\infty)$, thus we have

$$\begin{aligned}
N^{-1} &= \mathbb{P}(V>0, W>0) = \int_0^{+\infty} \mathrm{d}v \int_0^{+\infty} \mathrm{d}w\ f_{V,W}(v,w) \\
&= \int_0^{+\infty} \mathrm{d}u \int_0^{+\infty} \mathrm{d}v \int_0^{+\infty} \mathrm{d}w\ f_{U,V,W}(u,v,w) \\
&= \int_0^{+\infty} \frac{\mu^\lambda}{\Gamma(\lambda)} u^{\lambda-1} e^{-(\beta+\mu)u} \sum_{j=0}^{\alpha-1} \frac{(u\beta)^{k\alpha+j}}{(k\alpha+j)!}\,\mathrm{d}u \\
&= \frac{\mu^\lambda}{\Gamma(\lambda)} \sum_{j=0}^{\alpha-1} \frac{\beta^{k\alpha+j}}{(k\alpha+j)!} \int_0^{+\infty} e^{-(\beta+\mu)u} u^{k\alpha+j+\lambda-1}\,\mathrm{d}u
\end{aligned}$$

and using the substitution $(\beta+\mu)u = y$ we get overall

$$N^{-1} = \frac{\mu^\lambda}{\Gamma(\lambda)} \sum_{j=0}^{\alpha-1} \frac{\beta^{k\alpha+j}}{(k\alpha+j)!} \frac{\Gamma(k\alpha+j+\lambda)}{(\beta+\mu)^{k\alpha+j+\lambda}}\,.$$

By substituting the abovederived relationships for M and N into the original formula (2.3) we obtain the desired probability density

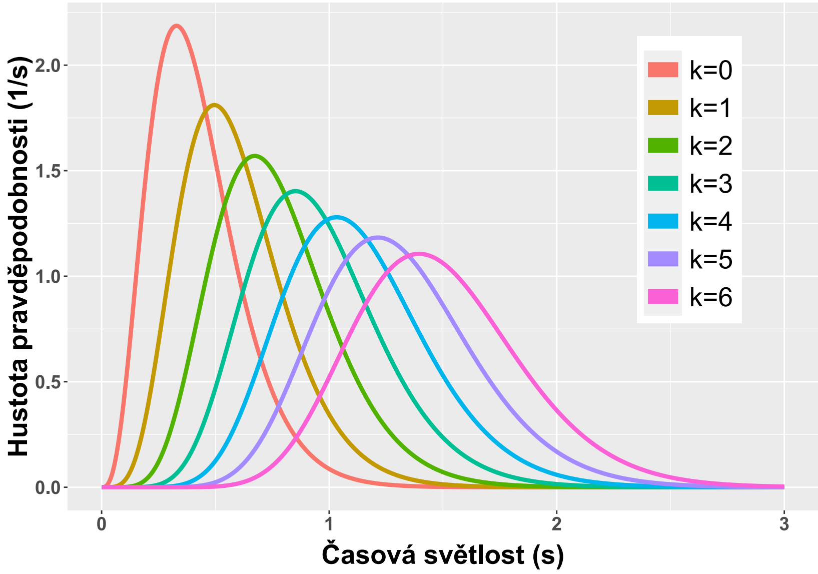
$$\begin{aligned}
f_{D^{(k)}}(u) = f_{U|V>0,W>0}(u) &= u^{\lambda-1} e^{-(\beta+\mu)u} \left(\sum_{j=0}^{\alpha-1} \frac{(u\beta)^{k\alpha+j}}{(k\alpha+j)!} \right) \cdot \left(\sum_{l=0}^{\alpha-1} \frac{\beta^{k\alpha+l}}{(k\alpha+l)!} \frac{\Gamma(k\alpha+l+\lambda)}{(\beta+\mu)^{k\alpha+l+\lambda}} \right)^{-1} \\
&= (\beta+\mu)^{k\alpha+\lambda}\, u^{k\alpha+\lambda-1} e^{-(\beta+\mu)u} \left(\sum_{j=0}^{\alpha-1} \frac{(u\beta)^{j}}{(k\alpha+j)!} \right) \cdot \left(\sum_{l=0}^{\alpha-1} \frac{\beta^{l}}{(\beta+\mu)^l} \frac{\Gamma(k\alpha+l+\lambda)}{(k\alpha+l)!} \right)^{-1},
\end{aligned} \tag{2.10}$$

for $u>0$.

Distribution Not Accepting Any Vehicle

The procedure introduced for deriving the distribution X accepted by exactly $k \in \mathbb{N}$ vehicles can be simply modified for the value $k=0$, which the result (2.10) generally admits. Indeed, instead of the transformation (2.2) consider only the twodimensional transformation

$$\mathbb{R}_+^2 \to \mathbb{R}_+ \times \mathbb{R} : (X,Y) \longmapsto (U,V) := (X, Y-X)$$



Obrázek 2.1: Densities of partial distributions of clearance orders $k \in \{0,\ 1,\ \dots, 8\}$ for a fixed choice of parameters of gamma distribution of clearance and Erlang distribution of critical clearance λ, μ, α, β, see relationship (2.11).

and then proceed only with the condition $V > 0$ analogously. We then arrive at the form of the density of unaccepted clearances

$$f_{D^{(0)}}(u) = (\beta+\mu)^{\lambda}\, u^{\lambda-1} e^{-(\beta+\mu)u} \left(\sum_{j=0}^{\alpha-1} \frac{(u\beta)^j}{j!}\right) \cdot \left(\sum_{l=0}^{\alpha-1} \frac{\beta^l}{(\beta+\mu)^l} \frac{\Gamma(l+\lambda)}{l!}\right)^{-1}.$$

2.2 Verification of Derivation by Simulations

Graphical Visualization

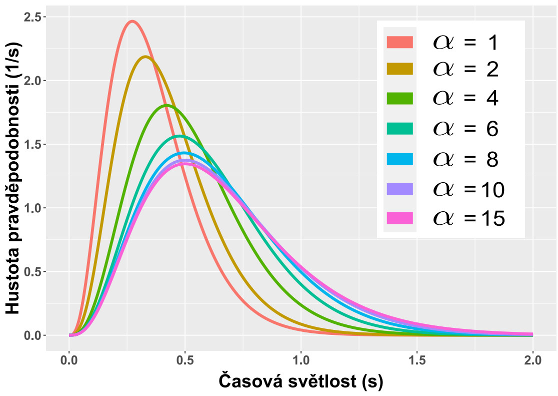
The depiction of the dependence of the density shape (2.10) on the number of accepted vehicles k is shown in Figure 2.1, where the parameters were fixedly chosen as

$$\alpha = 2,\ \beta = 5,\ \lambda = 4,\ \mu = 6\,. \tag{2.11}$$

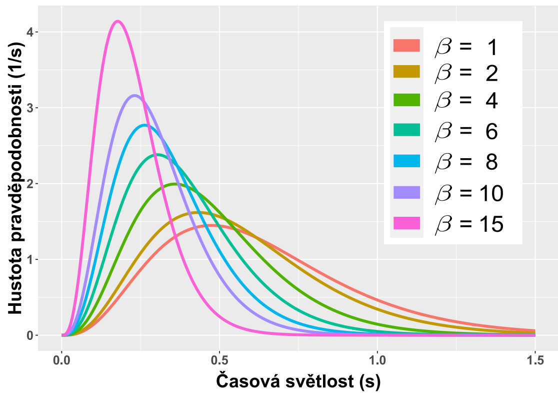
Let's introduce the labeling of these parameter values as **reference** and mention that such specific values were chosen only for the clarity of the graph and are not significant by themselves. From the figure, it is visible that as the acceptance order of clearance k increases, so do their mean value and variance.

Furthermore, Figure 2.2 illustrates the dependence of the density on changes in the parameters α, β, λ, μ, with the acceptance order $k = 0$ chosen for simplicity. In each of the graphs, only one parameter is variable, namely taking values from the set $\{1, 2, 4, 6, 8, 10, 15\}$. For the fixed parameters, the reference values were always chosen, see relationship (2.11).

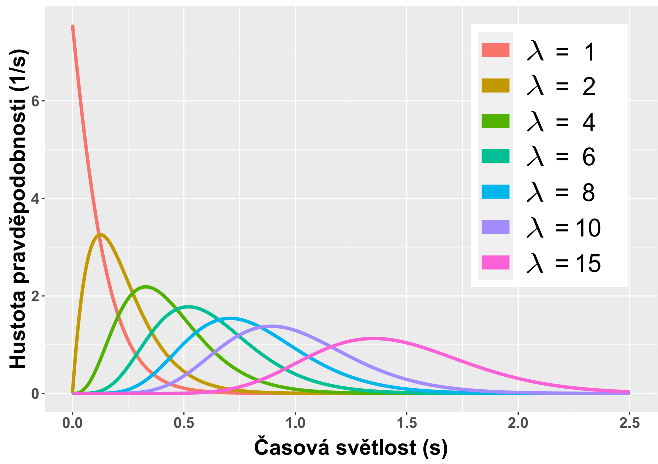
To graphically verify the correctness of the derived density shape $f_{D^{(k)}}$, clearances $\{x_i \,|\, i \in \mathbb{N}\} \sim \Gamma(\lambda,\mu)$ and critical clearances $\{y_i \,|\, i \in \mathbb{N}\} \sim \mathrm{Erlang}(\alpha,\beta)$ were simulated for the reference parameter values. According to rule (1.6), the clearances were divided according to their acceptance order and their histograms were graphically compared with the graphs of the theoretically derived density, see Figure 2.3. Each of the mentioned histograms is created from 10^5 realizations of the given random variable.



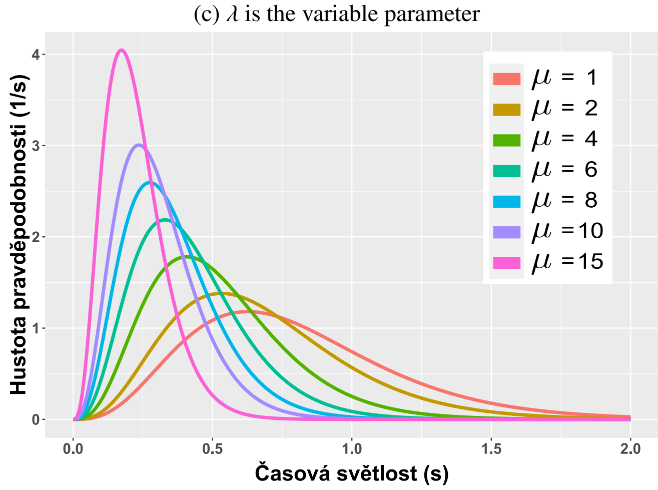
(a) α is the variable parameter



(b) β is the variable parameter

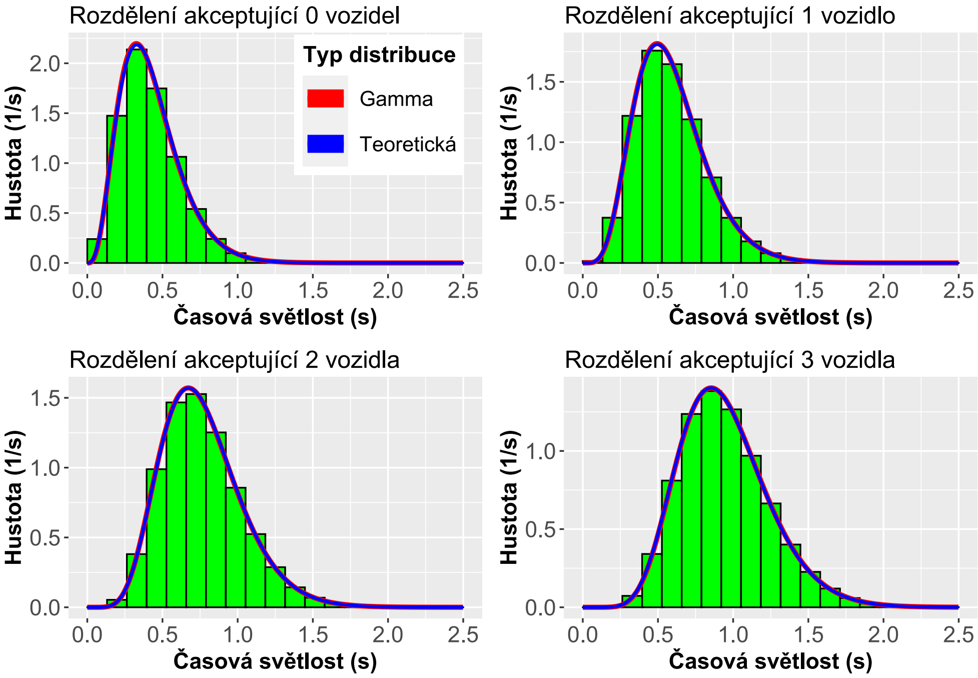


(c) λ is the variable parameter



(d) μ is the variable parameter

Obrázek 2.2: Densities of the zeroorder partial distribution, where 1 out of 4 parameters is variable and the others are chosen according to the reference selection, see (2.11).



Obrázek 2.3: Comparison of the histogram estimate of the density of the partial distribution of order $k \in \{0, 1, 2, 3\}$ created from simulations with the analytically derived density with reference parameters (2.1) and gamma density with estimated parameters using the MLE method

Hypothesis Testing

For statistical verification of the correctness of the derivation, we will use Pearson's χ^2 test of goodness of fit. As the null hypothesis H_0, we denote that the observed clearances of the kth order come from the distribution (2.10) with known parameters. We test H_0 against the alternative hypothesis H_1, that their distribution is different, thus preserving the notation of the density $f_{D^{(k)}}$ from section 1.3 :

$$H_0 : f_{A_k} = f_{D^{(k)}} \quad \text{vs.} \quad H_1 : f_{A_k} \neq f_{D^{(k)}} \,, \tag{2.12}$$

where we have labeled f_{A_k} as the density of simulated clearances accepted by exactly k vehicles.

For 4 different sets of parameters, see Table 2.1, 10^5 values of time clearances with acceptance orders $k \in \{0, 1, 2, 3, 4\}$ were simulated. For each such k and each set of parameters, the data were then divided into parts with 100 observations each, and with each such part, Pearson's χ^2 test of the

Set of parameters	α	β	λ	μ
S_1	2	5	4	6
S_2	5	12	13	12
S_3	6	4	5	2
S_4	1	4	2	2

Tabulka 2.1: Sets of parameters S_1, S_2, S_3, S_4 used for Pearson and gamma goodness of fit tests

	Acceptance order k				
	$k = 0$	$k = 1$	$k = 2$	$k = 3$	$k = 4$
Set S_1	0.050	0.046	0.041	0.041	0.042
Set S_2	0.051	0.043	0.045	0.055	0.040
Set S_3	0.060	0.046	0.047	0.055	0.047
Set S_4	0.055	0.065	0.044	0.053	0.050

Tabulka 2.2: Ratio of rejection out of a total of 1000 conducted Pearson's tests of the hypothesis (2.12) for different sets of parameters, see Table 2.1, and various acceptance orders.

hypothesis (2.12) was conducted. Within the Pearson test, the range of observed variable values was always divided into 10 bins according to the 10quantiles of the theoretical distribution. After calculating the expected realizations of the random variable E_j, $j \in \{1, \ldots, 10\}$ in each bin, Pearson's statistic was used to evaluate the test

$$\chi^2 = \sum_{i=1}^{10} \frac{(O_i - E_i)^2}{E_i}, \tag{2.13}$$

where we have denoted O_i as the number of observed simulated values in the jth bin for $j \in \{1, \ldots, 10\}$. The hypothesis was rejected exactly when

$$\chi^2 \geq \chi^2_{0.95}(9), \tag{2.14}$$

where on the left side of the inequality is Pearson's statistic (2.13) and on the right side is the 95quantile of the χ^2 distribution with nine degrees of freedom. Simultaneously with Pearson's tests, hypothesis tests were conducted to determine whether the corresponding distribution f_{A_k} belongs to the family of gamma distributions. The general form of the gamma density (1.5) is indeed very simple compared to the density (2.10), hence it is desirable to evaluate how good an approximation of the derived distribution actually is. For this purpose, in the **R** programming environment, the function *gamma_test* from the *goft* library was used, which is further described in [**Rdocumentation**]. The hypothesis with the gamma distribution was rejected precisely when the pvalue determined by the function *gamma_test* was less than 0.05.

The ratios of rejections of the individual Pearson's χ^2, resp. gamma tests at the 5% level of statistical significance are presented in Table 2.2, resp. 2.3. In the case of Pearson's χ^2 tests, the rejection ratio of the hypothesis ranges within $[0.04, 0.065]$, which corresponds very well to the set level of significance of the test, which due to the rejection rule (2.14) is equal to 5%. This numerically confirms the correctness of the derived density (2.10). In the case of the presented results of tests for the hypothesis with the gamma distribution, the rejection share is slightly lower than the desired 5%, however, it seems that the approximation by the gamma distribution is satisfactory, see further analysis of empirical data in section 3.4.

2.3 Siegloch's Function

In this section, we will derive Siegloch's function $s(t)$ for gammadistributed critical clearances. Knowing the traffic intensity and the distribution of clearance in the main flow, we can then determine the capacity of the intersection, see relationship (1.2). In accordance with the verbal definition of $s(t)$

Acceptance order k					
	$k=0$	$k=1$	$k=2$	$k=3$	$k=4$
Set S_1	0.017	0.023	0.020	0.018	0.023
Set S_2	0.022	0.015	0.019	0.025	0.018
Set S_3	0.036	0.022	0.022	0.020	0.026
Set S_4	0.028	0.027	0.027	0.029	0.019

Tabulka 2.3: Ratio of rejection out of a total of 1000 conducted tests of the hypothesis, whether simulated clearances of the kth order from the derived distribution (2.10) with parameters from various sets $S_1, \ldots, S_4$ are gamma distributed with estimated parameters.

stated in section 1.2, we have the defining relationship

$$s(t) = \sum_{k=0}^{+\infty} k p_k(t)\,, \tag{2.15}$$

where $p_k(t)$ is the probability of accepting a clearance of length t by exactly k vehicles, thus using the variable Z introduced in (2.1) it holds for $k \in \mathbb{N}$

$$p_k(t) = \Pr\left[Z < t < Z + Y_{k+1}\right]\,.$$

Utilizing the regular transformation $g : \mathbb{R}_+^2 \to \mathbb{R}_+^2$ given by:

$$g : \begin{pmatrix} Z \\ Y_{k+1} \end{pmatrix} \mapsto \begin{pmatrix} Q \\ R \end{pmatrix} = \begin{pmatrix} Z \\ Z + Y_{k+1} \end{pmatrix},$$

we obtain the formula for the joint density of variables Q, R in the form

$$f_{Q,R}(q,r) = f_{Z,Y}(q, r-q) = f_Z(q) f_Y(r-q)$$

for $q > 0$ and $r - q > 0$. The probability $p_k(t)$ can then be expressed as

$$p_k(t) = \Pr\left[Q < t < R\right] = \Pr\left[(Q,R) \in A_t\right] = \iint_{A_t} f_{Q,R}(q,r)\,\mathrm{d}(q,r)\,,$$

where we have defined the set $A_t := \{(q,r) \in \mathbb{R}^2 : q < t < r\}$. By applying Fubini's theorem, we finally get

$$p_k(t) = \int_0^t f_Z(q) \left(\int_t^{+\infty} f_Y(r-q)\,\mathrm{d}r \right) \mathrm{d}q\,, \tag{2.16}$$

where recall that $f_Z(z) = \bigstar_{i=1}^{k} f_Y(z)$.

Siegloch's Function $s(t)$ for Gamma Distribution

From the previous section, we know from the reproductive property of the gamma distribution that $Z \sim \mathrm{Gamma}(k\alpha, \beta)$. By substituting the appropriate densities into relationship (2.16) we obtain

$$p_k(t) = \frac{\beta^{k\alpha}\beta^{\alpha}}{\Gamma(k\alpha)\Gamma(\alpha)} \int_0^t q^{k\alpha-1} e^{-\beta q} \underbrace{\left(\int_t^{+\infty} (r-q)^{\alpha-1} e^{-\beta(r-q)} \, \mathrm{d}r \right)}_{\text{denoted } \mathcal{K}} \mathrm{d}q \tag{2.17}$$

and by substituting the variables $\beta(r-q) = s$ in the integral $\mathcal{K}$ we arrive at the expression

$$\mathcal{K} = \frac{1}{\beta^{\alpha}} \int_{\beta(t-q)}^{+\infty} s^{\alpha-1} e^{-s} \, \mathrm{d}s = \frac{\Gamma(\alpha, \beta(t-q))}{\beta^{\alpha}} \,,$$

where we denoted the incomplete upper gamma function as Γ with two input values[1]. For the adjustment of the incomplete upper gamma function, we use the known recursive formula resulting from the method of integration by parts, namely

$$\Gamma(s+1, x) = s\,\Gamma(s, x) + x^s e^{-x} \,, \quad \forall s > 0,\ x > 0 \,. \tag{2.18}$$

For $s \in \mathbb{N}_0$, the formula (2.18) can be further adjusted to the form

$$\Gamma(s+1, x) = \Gamma(s+1)\, e^{-x} \sum_{k=0}^{s} \frac{x^k}{k!} \,, \quad \forall s \in \mathbb{N}_0,\ x > 0 \,. \tag{2.19}$$

Therefore, we reintroduce the restriction of the parameter α to natural numbers, thus consider the variable Y with Erlang distribution. Then we can use the formula (2.19) to adjust $\mathcal{K}$:

$$\mathcal{K} = \frac{\Gamma(\alpha)}{\beta^{\alpha}} e^{-\beta(t-q)} \sum_{l=0}^{\alpha-1} \frac{\beta^l (t-q)^l}{l!} \,, \tag{2.20}$$

where we chose the same notation as in (2.19)

$$s + 1 = \alpha, \quad x = \beta(t-q).$$

By substituting (2.20) into (2.17), we further resolve the emerging Beta integral in a manner analogous to (2.6):

$$\begin{aligned} p_k(t) &= \frac{\beta^{k\alpha}}{\Gamma(k\alpha)} e^{-\beta t} \sum_{l=0}^{\alpha-1} \frac{\beta^l}{l!} \int_0^t q^{k\alpha-1} (t-q)^l \, \mathrm{d}q \\ &= \frac{\beta^{k\alpha}}{\Gamma(k\alpha)} e^{-\beta t} \sum_{l=0}^{\alpha-1} \frac{\beta^l}{l!} t^{k\alpha+l} \, \mathrm{B}(k\alpha, l+1) \\ &= e^{-\beta t} \sum_{l=0}^{\alpha-1} \frac{(\beta t)^{k\alpha+l}}{(k\alpha+l)!} \,, \quad \forall k \in \mathbb{N} \,. \end{aligned} \tag{2.21}$$

[1]Here occurs a collision with the notation of Γ for the gamma distribution. However, from the context, it is always clear which interpretation of the symbol Γ is currently being used.

The validity of relationship (2.21) can easily be extended to include the value $k = 0$, realizing that

$$\begin{aligned} p_0(t) &= \mathsf{Pr}\,[Y > t] = 1 - F_Y(t) = 1 - \frac{1}{\Gamma(\alpha)}\gamma(\alpha,\beta t) \\ &= \frac{\Gamma(\alpha) - \gamma(\alpha,\beta t)}{\Gamma(\alpha)} = \frac{\Gamma(\alpha,\beta t)}{\Gamma(\alpha)} \\ &= e^{-\beta t}\sum_{l=0}^{\alpha-1}\frac{(\beta t)^l}{l!}\,, \end{aligned}$$

where $F_Y(t)$ we have denoted the cumulative distribution function of the variable Y, γ is the incomplete lower gamma function and the last equality uses the relationship (2.19).

Overall, we have achieved the expression $p_k(t)\,, \forall k \in \mathbb{N}_0,\ t > 0$. We can therefore now substitute our partial results into the definitional relationship (2.15), i.e.,

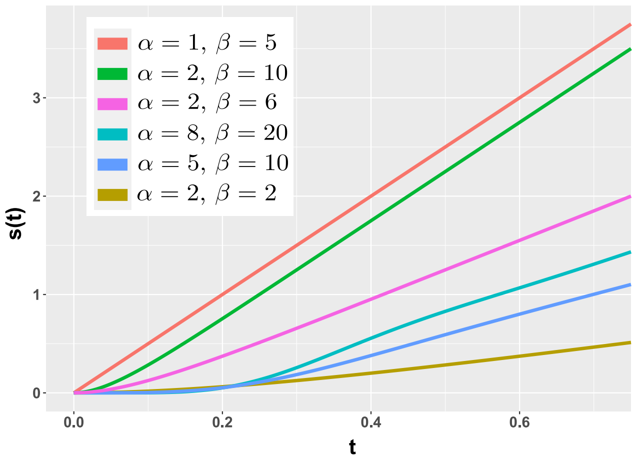
$$s(t) = e^{-\beta t}\sum_{l=0}^{\alpha-1}\sum_{k=0}^{+\infty} k\frac{(\beta t)^{k\alpha+l}}{(k\alpha+l)!}\,, \quad \forall t > 0\,.$$

Especially for $\alpha = 1$, at which the gamma density (1.5) degenerates to the density of an exponential distribution with parameter β, Siegloch's function is a linear function with a slope of β:

$$\begin{aligned} s(t) &= e^{-\beta t}\sum_{k=0}^{+\infty} k\frac{(\beta t)^k}{k!} = e^{-\beta t}\sum_{k=1}^{+\infty}\frac{(\beta t)^k}{(k-1)!} \\ &= e^{-\beta t}\beta t\sum_{k=1}^{+\infty}\frac{(\beta t)^{k-1}}{(k-1)!} = e^{-\beta t}\beta t\, e^{\beta t} \\ &= \beta t\,, \quad \forall t > 0\,. \end{aligned} \tag{2.22}$$

In practice, Siegloch's function is approximated by a linear function as follows. For each observed clearance $x > 0$, its acceptance order $k \in \mathbb{N}_0$ is recorded. The set of all such ordered pairs (x, k) is then fitted by the method of linear regression with a line, which corresponds to the approximation of Siegloch's function $s(t)$.

The development of Siegloch's function, and especially its linear asymptote, for different parameters of the distribution $Y \sim \text{Erlang}(\alpha,\beta)$ is shown in Figure 2.4.



Obrázek 2.4: Siegloch's function $s(t)$ for different values of parameters of critical clearances $Y \sim \text{Erlang}(\alpha,\beta)$

Kapitola 3

Traffic Data Analysis

In this chapter, an analysis of empirical data from three different uncontrolled Tintersections will be conducted. For this analysis, we will use the mathematical model of the intersection and the variable notation[1], which were introduced in the previous chapters. The main subject of interest is the comparison of the shape of the empirical density of the partial distribution of order $k \in \mathbb{N}_0$ with the derived density (2.10). For this purpose, Pearson's χ^2 test of goodness of fit, which was introduced in detail and whose functionality was verified by simulation means in section 2.2, will be used.

3.1 Empirical Data Used

For this work, data from measurements conducted at the following three intersections[2] in Germany, which we denote consecutively as intersection no. 1, no. 2, and no. 3:

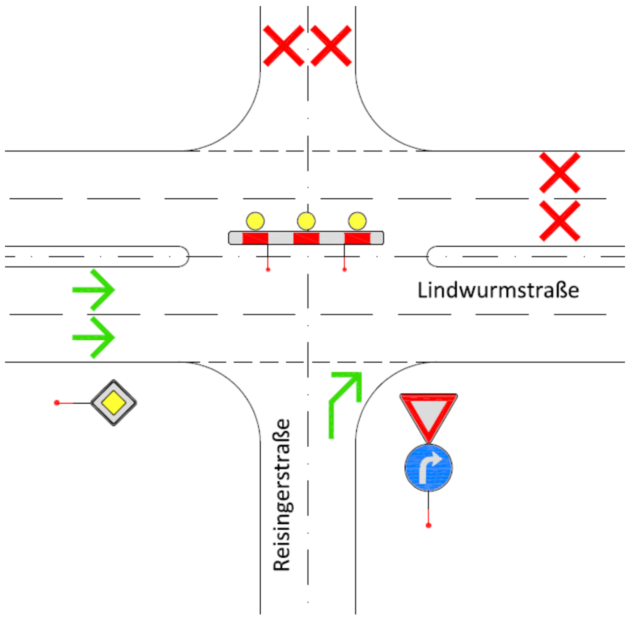
LindwurmstraßeReisingerstraße
: is an intersecting Tintersection of the directionally divided fourlane local road Lindwurmstraße with the twolane undivided road Reisingerstraße in Munich. At the time of measurement, one lane of Lindwurmstraße was closed, see Figure 3.1, and the local traffic lights were turned off during the closure. This temporarily turned the cross intersection into a Tintersection.

LindwurmstraßeFliegenstraße
: is a Tintersection with marked driving priority using traffic signs, see Figure 3.2.

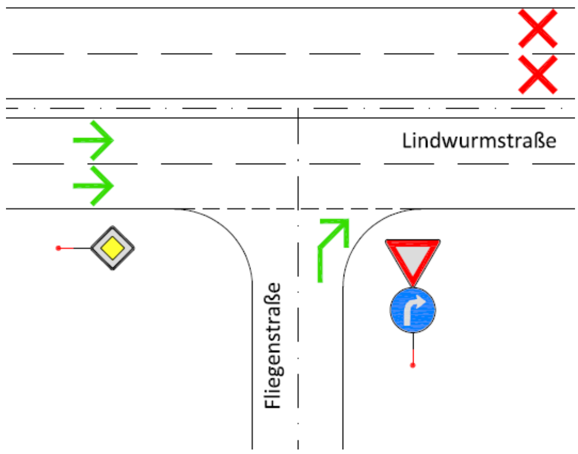
LeisnigerstraßeRobertMatzkeStraße
: is an intersecting Tintersection of the undivided twolane local road Leisnigerstraße with the two-lane undivided trafficcalmed local road RobertMatzkeStraße in Dresden. At the time of measurement, the street opposite the entrance to RMStraße was closed, and Leisnigerstraße was made oneway, see Figure 3.3.

[1] especially the clearance X and the critical clearance Y

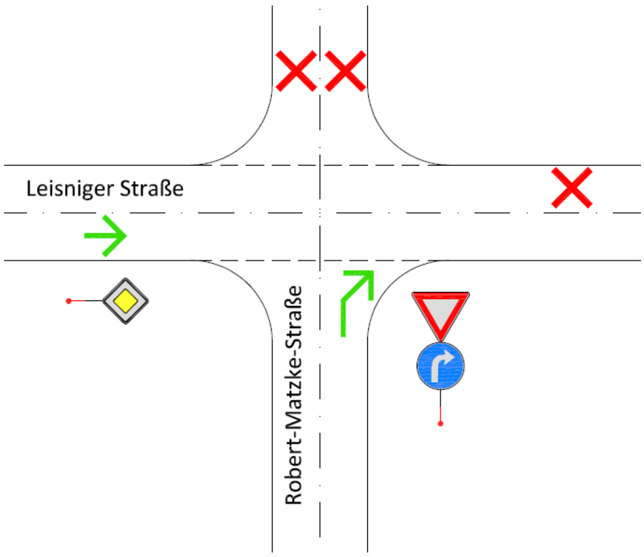
[2] the first part of the name indicates the main road, the second part then the side road



Obrázek 3.1: Scheme of the Lindwurmstraße Reisingerstraße intersection



Obrázek 3.2: Scheme of the Lindwurmstraße Fliegenstraße intersection



Obrázek 3.3: Scheme of the Leisnigerstraße RobertMatzkeStraße intersection

For each intersection, approximately 30 000 time gaps on the main road were recorded during the measurement period. These values are further supplemented with information about

(a) the acceptance order of the gaps,

(b) vehicle speeds,

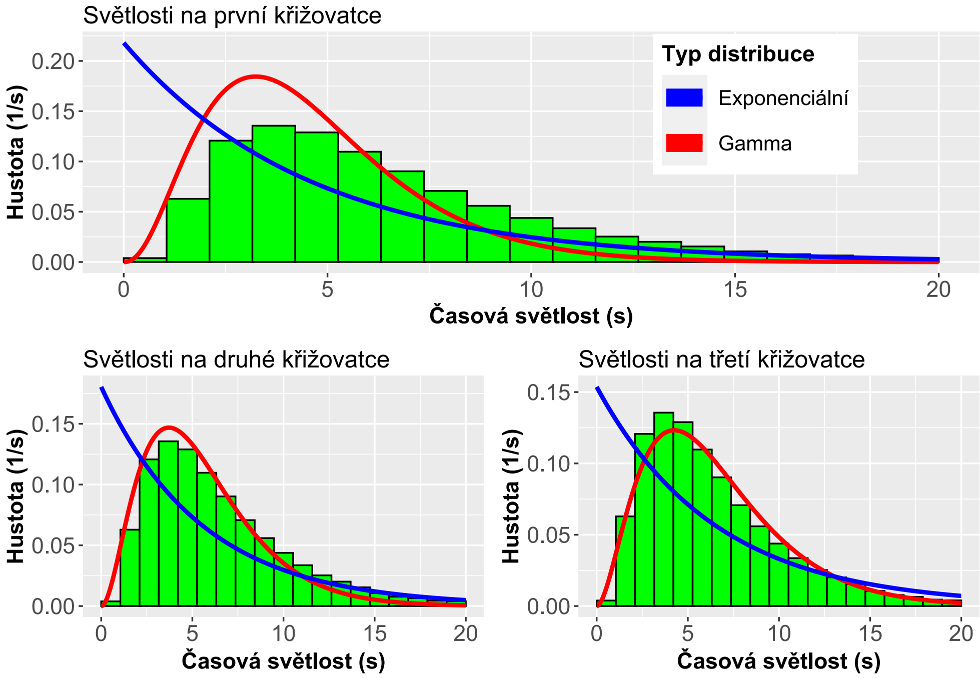
(c) local traffic density[3] calculated from fixed sections after every 50 vehicles.

Figure 3.4 shows histograms of all observed time gaps for all intersections. From the comparison with estimated gamma and exponential densities, it is evident that the hitherto used assumption of the exponential distribution of time gaps is significantly insufficient. The gamma density appears as a more friendly approximation, however, as we previously stated in section 1.3, there is still room for improvement of the current model.

Table 3.1 categorizes the time gaps at all intersections according to the acceptance order k. Gaps at intersection no. 3 are significantly more often accepted by three or more vehicles compared to intersection no. 1. Adding to this fact the information about the average traffic density on the main roads, which are 15.4, 12.5 and 11.1 vehicles per kilometer for each of the intersections, it is not yet possible to unequivocally decide whether the critical gap for drivers at intersection no. 1 is greater than at intersections no. 2 and 3. It is possible that the increased traffic density was the cause of the rarely occurring higher order gaps.

In the picture 3.5, kernel density estimates of light intensities for different acceptance orders are displayed at all intersections. Due to the small number of measured values for acceptance orders $k \geq 4$, their respective graphs were omitted. Goodnessoffit tests were further conducted for the displayed light intensities using the gamma_test function, which was already used in section 2.2. From the pvalues of individual tests, as shown in table 3.2, it follows that for the acceptance order $k = 3$, it is not possible to reject at a significance level of 5% that the distributions of these light intensities are from the gamma distribution family. On the other hand, for the unaccepted light intensities, the hypothesis was always rejected at the 5% significance level.

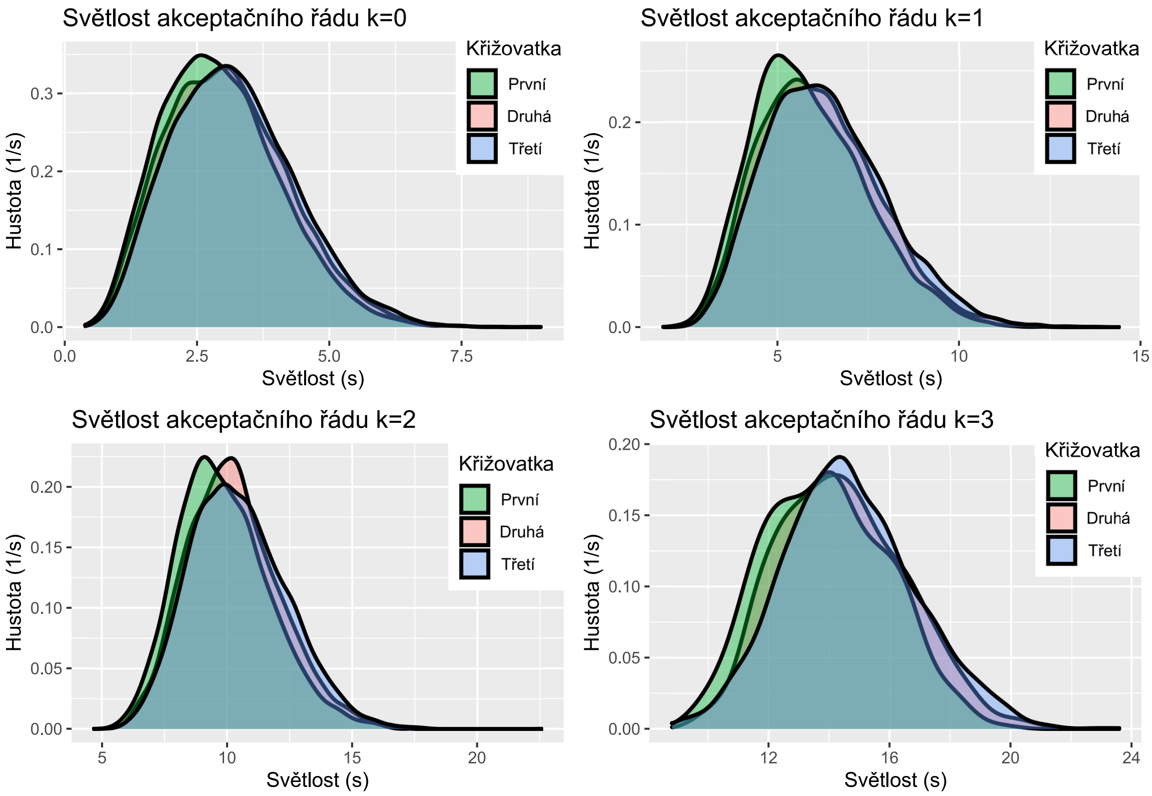
[3] traffic density is commonly expressed in units of vehicles per kilometer



Obrázek 3.4: Histograms of observed time gaps at different intersections, fitted with densities of the exponential and gamma distribution with estimated parameters.

	Intersection 1			Intersection 2			Intersection 3		
	count	%	EX (s)	count	%	EX (s)	count	%	EX (s)
$k = 0$	16196	56.7 %	2.95	10799	46.2 %	3.08	12340	37.5 %	3.20
$k = 1$	10155	35.6 %	5.89	9115	39.0 %	6.16	13206	40.1 %	6.34
$k = 2$	1853	6.5 %	9.88	2645	11.3 %	10.27	5145	15.6 %	10.48
$k = 3$	305	1.1 %	13.91	653	2.8 %	14.43	1554	4.7 %	14.67
$k = 4$	38	0.1 %	18.32	139	0.6 %	18.53	485	1.5 %	18.85
$k = 5$	2	0.0 %	21.19	36	0.2 %	22.56	153	0.5 %	23.08
$k = 6$	1	0.0 %	23.19	8	0.0 %	26.73	43	0.1 %	27.45
$k = 7$	0	0.0 %	0.00	4	0.0 %	31.80	19	0.0 %	32.22
Total	28550			23399			32945		

Tabulka 3.1: Categorization of gaps at all intersections according to their acceptance order, their percentage share of the total number of measurements, and the average value of gaps EX for the given acceptance order in seconds



Obrázek 3.5: Kernel density estimates of probability densities for light intensities of acceptance orders $k \in \{0, 1, 2, 3\}$ at all intersections

	Intersection 1	Intersection 2	Intersection 3
$k = 0$	< 0.001	< 0.001	< 0.001
$k = 1$	< 0.001	0.21	0.07
$k = 2$	0.01	0.08	0.03
$k = 3$	0.74	0.54	0.85

Tabulka 3.2: Pvalues of goodnessoffit tests for the distributions of light intensities for the given acceptance orders and gamma distributions with estimated parameters, conducted using the gamma_test function.

3.2 Empirical clearance Hypothesis

Let us now test a similar hypothesis as in (2.12), namely whether empirical clearances of acceptance order k have a distribution given by the derived probability density (2.10). To do this, we will use Pearson's χ^2 test in a similar way to its implementation in section 2.2. Unlike simulations where the distribution parameters of clearances X and critical clearances Y are known, in the analysis of empirical clearances, these parameters must be estimated from measurements. For this, we will use the MLE method proven in [5], whose specific application is described in the following section 3.3.

3.3 Use of the MLE Method

To obtain the most credible estimate of the parameter $\theta := (\alpha, \beta, \lambda, \mu) \in \mathbb{N} \times \mathbb{R}_+^3$, let us maximize the likelihood function L of the random sample $D_1^{(k)}, \ldots, D_n^{(k)}$ of size $n \in \mathbb{N}$ from the partial distribution $D^{(k)}$ of order k, which as a function of variable θ is in the form

$$L(\theta \,|\, \mathbf{d}) := \prod_{i=1}^{n} f_{D^{(k)}}(d_j)\,, \quad \forall\, d_j > 0,\ j \in \{1, \ldots, n\}\,, \tag{3.1}$$

where we denoted $\mathbf{d} = (d_1, \ldots, d_n)$ the realization of the random sample.

For practical computation, it is advantageous to convert the task of maximizing the likelihood function (3.1) over the parameter space $\Theta := \mathbb{N} \times \mathbb{R}_+^3$ to minimizing the negatively taken loglikelihood function ℓ, thus we seek

$$\arg\min_{\theta \in \Theta} \left(-\ell(\theta \,|\, \mathbf{d}) \right) = \arg\min_{\theta \in \Theta} \left(-\ln L(\theta \,|\, \mathbf{d}) \right). \tag{3.2}$$

For solving this optimization task, the function[4] *nlm* from the *stats* library was used, which searches for a minimum using the line search method. To evaluate the quality of the estimates that nlm provides, clearances of orders $k \in \{0, 1, 3\}$ were simulated for different (known) sets of parameters of the distributions X, Y, further see subsection 3.3. Unfortunately, the complex form of the loglikelihood function

$$\begin{aligned} \ell(\theta|\mathbf{d}) = (k\alpha + \lambda)n \ln(\beta + \mu) + \sum_{i=1}^{n} (k\alpha + \lambda - 1) \ln(d_i) - (\beta + \mu) \sum_{i=1}^{n} d_i \\ + \sum_{i=1}^{n} \ln\left(\sum_{j=0}^{\alpha-1} \frac{(d_i \beta)^j}{(k\alpha + j)!} \right) - n \ln\left(\sum_{j=0}^{\alpha-1} \frac{\beta^j}{(\beta + \mu)^j} \cdot \frac{\Gamma(k\alpha + \lambda + j)}{(k\alpha + j)!} \right) \end{aligned} \tag{3.3}$$

resulted in nlm failing to estimate the parameter $\alpha \in \mathbb{N}$ and only returning the initial, manually entered estimate. Therefore, we were forced to fix the parameter α at a fixed value and search for the solution of task (3.2) in the space $\mathbb{R}_+^3$. After this adjustment, the outputs of nlm were more favorable, but still grossly insufficient. Therefore, we further scaled the mean value of the clearance $\mathrm{E}X = \lambda/\mu = 1$, which reduced the parameter space Θ to only $\mathbb{R}_+^2$ and the resulting estimates were significantly refined. According to [4], scaling the mean value of clearances is a common practice in traffic measurement analyses that use the so called 3*s unification procedure*.

[4] from the programming environment **R**

Parameter Set	α	β	$\lambda = \mu$
T_1	4	10	4
T_2	2	6	2
T_3	1	2	10

Tabulka 3.3: Parameter sets T_1, T_2, T_3 used to test the quality of estimates of the distribution $D^{(k)}$ for $k = 0, 1, 3$, provided by the nlm function.

Let us now summarize for clarity the assumptions introduced above for numerical verification of the quality of nlm function estimates:

- $\alpha \in \mathbb{N}$ is a fixed parameter that is not estimated,
- the distribution of time clearances is scaled so that $\mathrm{E}X = \lambda/\mu = 1$, i.e., equivalently $\lambda = \mu$.

This has reduced the original parameter space $\mathbb{N} \times \mathbb{R}_+^3$ for task (3.2) to

$$\Theta := \mathbb{R}_+^2 = \{\, (\beta, \mu) \,|\, \beta > 0,\ \mu > 0 \,\}.$$

Quality of Estimates from the nlm Function

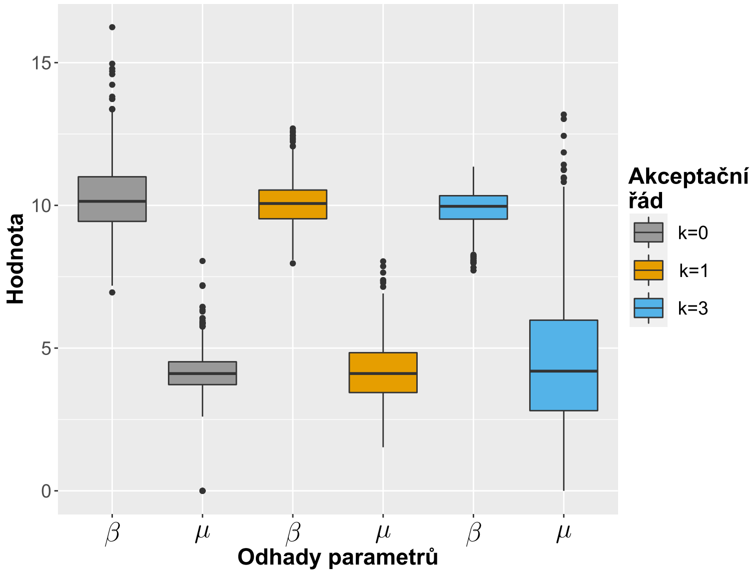
For the parameter sets T_1, T_2, T_3, see Table 3.3, 10^5 clearances of order $k \in \{0, 1, 3\}$ were simulated. These were divided into parts of 100 observations each and an estimate was made for each part using the nlm function. For the initial parameter estimate, which is required for using nlm, the actual (known) values of the parameters were used, which were deviated by random realizations from a uniform distribution on the interval $(-1, 1)$. For the rare cases when any of the parameters was estimated as a negative value, this estimate was overwritten with the value 0. This situation only occurred for the parameter $\mu = \lambda$ and was usually accompanied by an aboveaverage value of the parameter β estimate. The resulting boxplots in Figures 3.6, 3.7, and 3.8 correspond to the unbiasedness of the estimate and overall also the insensitivity of nlm to the choice of parameters α, β, $\lambda = \mu$.

3.4 Pearson's χ^2 Test

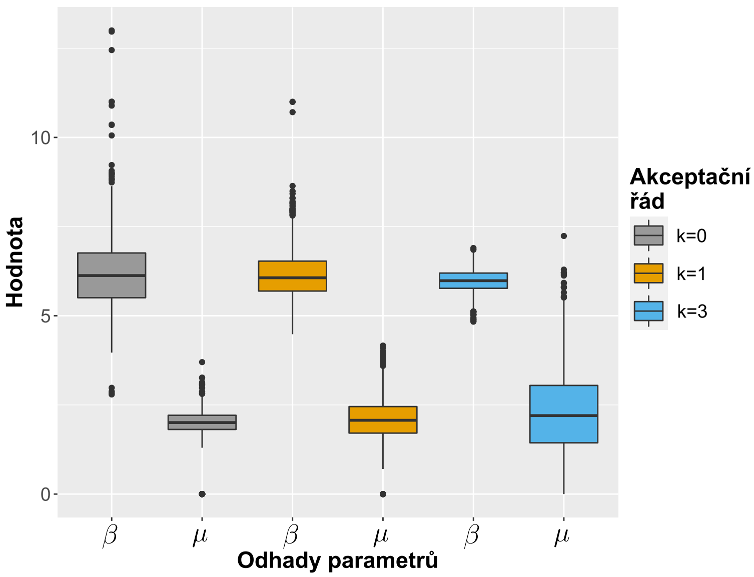
In this section, we will conduct Pearson's χ^2 test of the empirical clearance hypothesis[5] from measurements at the intersections introduced in section 3.1. In accordance with the established restrictions for the originally fourdimensional parameter space Θ in the optimization task (3.2), we cannot generally use the procedure for estimating the parameters β and μ as in section 3.3. In the case of the simulations, we knew for which value of $\alpha \in \mathbb{N}$ we had to minimize the negatively taken loglikelihood function ℓ, unfortunately, we do not have this information available now. Implementing a more complex algorithm to find the solution to the task (3.2), which could also estimate the parameter α, might be a suitable solution, but in this work, we will suffice with a simpler following procedure.

As we have already mentioned before, for the special case $\alpha = 1$, the gamma density (1.5) coincides with the density of the exponential distribution. Therefore, if we still consider the gamma distribution of critical clearances as a generalization of the exponential distribution previously used in the literature,

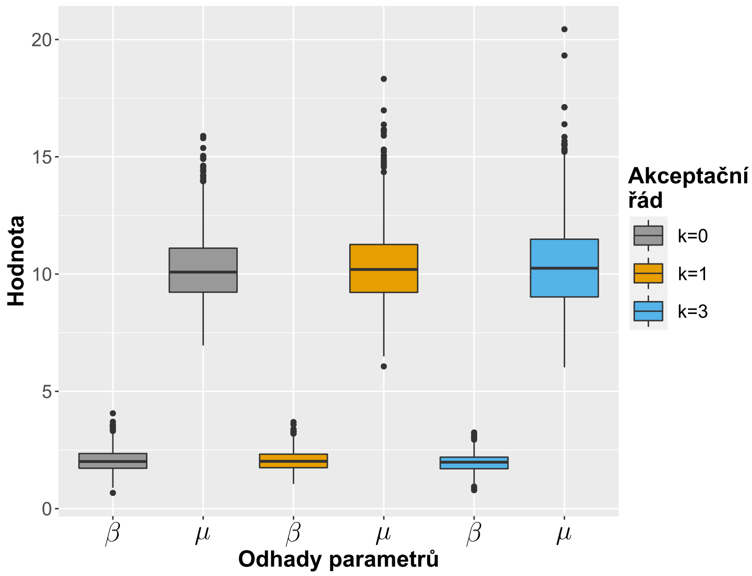
[5] see section 3.2



Obrázek 3.6: Boxplots of parameter estimates β and $\mu = \lambda$ for the distribution $D^{(k)}$ for $k = 0, 1, 3$ using the nlm function with parameter set T_1, see Table 3.3.



Obrázek 3.7: Boxplots of parameter estimates β and $\mu = \lambda$ for the distribution $D^{(k)}$ for $k = 0, 1, 3$ using the nlm function with parameter set T_2, see Table 3.3.



Obrázek 3.8: Boxplots of parameter estimates β and $\mu = \lambda$ for the distribution $D^{(k)}$ for $k = 0, 1, 3$ using the nlm function with parameter set T_3, see Table 3.3.

it is justifiable to consider in the first approximation the values of the parameter α limited by a certain constant, which is *relatively close* to one. Therefore, for each considered data set, see below, Pearson's χ^2 test was performed several times, namely with a variable value of the parameter $\alpha \in \{1, \ldots, 15\}$. That is, for a given data segment, we chose $\alpha \leq 15$ and then estimated the remaining parameters β and μ using the nlm function. With these parameter estimates, we then carried out Pearson's test of goodness of fit. The obtained triples of parameters $(\alpha^i, \beta^i, \mu^i)$, $i \in \{1, \ldots, 15\}$ were compared according to the corresponding values of the Pearson statistic (2.13) and as the estimate of the parameters α, β, μ, we determined the triple $(\alpha^*, \beta^*, \mu^*)$ that had its lowest value. As a result of Pearson's test for a given segment, we then stored the result for these parameter values.

Implementation of the χ^2 Test

For each intersection, we first excluded clearances that belonged to *extreme states* at the intersection from the following analysis. By extreme states, we mean very high or low vehicle speeds and values of local traffic density. Specifically, we excluded clearances that corresponded to vehicles with speed and local traffic density lower, or higher than the tenth and ninetieth empirical percentile of speed and local density for the respective intersection. This gave us almost homogeneous data sets, which, according to the author [4], is often desirable in traffic data analysis.

Furthermore, for each intersection, we scaled the clearance values so that their average value was 1 second and for the acceptance order $k \in \{0, 1, 2, 3\}$ we then divided them into parts of 100 observations, if possible[6]. Following the procedure mentioned above, we then estimated the parameters α, β, μ for each part and conducted Pearson's χ^2 test analogously to section 2.2. The range of clearance values of order k was divided into 10 bins with edges at the 10quantiles of the distribution given by the density (2.10) with the estimated parameters.

Unlike the rule (2.14) determining the rejection of the test during simulations, we now need to consider the number of estimated parameters. Since we knew the value of α for each test, the number of

[6]if the number of clearances was not a multiple of 100, the last part contained fewer observations than the others

Acceptance Order				
	k=0	k=1	k=2	k=3
Pearson	0.029	0.047	0.000	0.500
Gamma	0.038	0.016	0.000	0.000
Number of tests	104	64	11	2

Tabulka 3.4: Ratios of rejection of Pearson's tests, respectively gamma tests of hypotheses considered in section 3.4 and the number of tests conducted for intersection No. 1 (both at a significance level of 5 %)

estimated parameters is 2. Therefore, we reject the hypothesis exactly when for the Pearson statistic χ^2 it holds that

$$\chi^2 \geq \chi^2_{0.95}(7), \tag{3.4}$$

where on the right side of the inequality is the 95quantile of the χ^2 distribution with seven degrees of freedom.

Similarly to section 2.2, along with Pearson's test, a test of the hypothesis was conducted to determine if the observed scaled clearances of order k have a distribution from the gamma family. The results of the tests were evaluated according to the pvalues returned by the gamma_test function.

Evaluation of the Tests

For individual intersections, the ratios of rejection of Pearson's tests and gamma tests are presented in Tables 3.4, and 3.5. Because a different number of tests were conducted for each order of acceptance, the total number of tests from which the rejection ratio was determined is also listed in the tables for clarity.

From Pearson's tests, the highest ratio of rejection belonged to clearances of order $k = 3$ at intersection No. 1, namely 50 %. However, this figure is almost irrelevant given the two tests conducted. Aside from this outlier value, the overall ratios of rejection by Pearson's test were in the interval $[0, 0.064]$. Moreover, only for order $k = 2$, respectively $k = 1$ at intersection No. 2, respectively No. 3, were the rejection ratios higher than 5 %. This fact corresponds to a good fit of the data with the model given by density (2.10) within the set significance level of the tests 5 %. Therefore, we can state that the hypothesis of empirical properties with our derived distribution remains unrefuted.

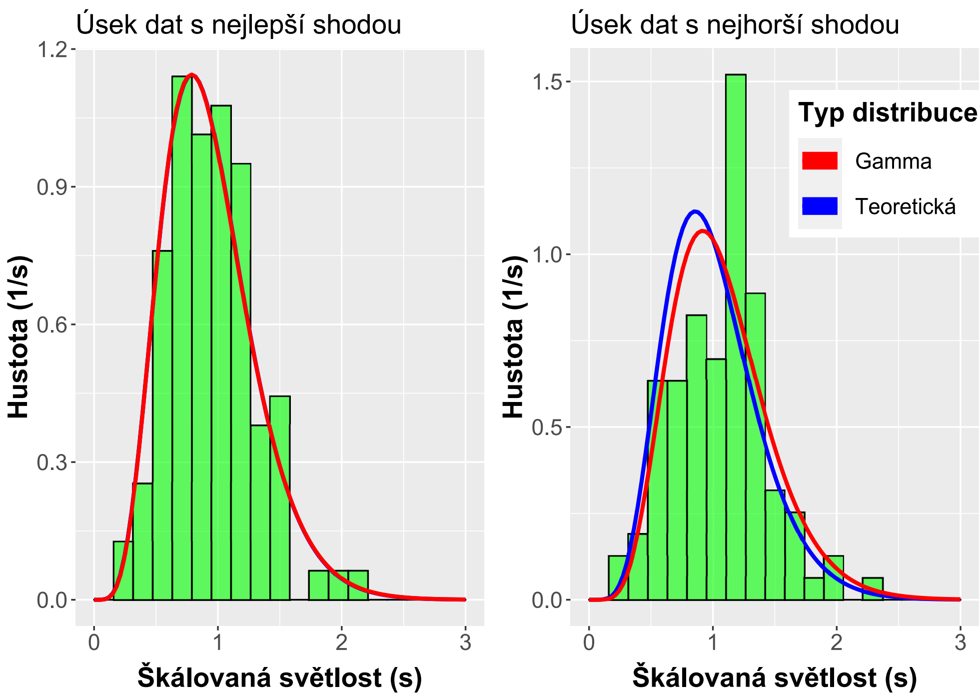
The ratios of rejection using gamma tests for zeroorder clearances are higher at each intersection than with Pearson's tests. On the other hand, for clearances of higher orders, they have a lower ratio, often even zero, with one exception. The gamma distribution is therefore a very accurate approximation of the distribution of clearances of orders $k = 1, 2, 3$. For zeroorder clearances, this approximation is worse but still very satisfactory.

In Figures 3.9, 3.10, and 3.11, comparisons of some histograms, estimated gamma densities, and densities $f_{D^{(k)}}$ with estimated parameters for various clearances that were part of the above tests are displayed. The shown clearances were always for a given intersection and acceptance order in a certain segment, which had 100 observations. For a given k and intersection, we always present a comparison for two data segments (sections), namely for which the Pearson statistic was lowest and highest[7].

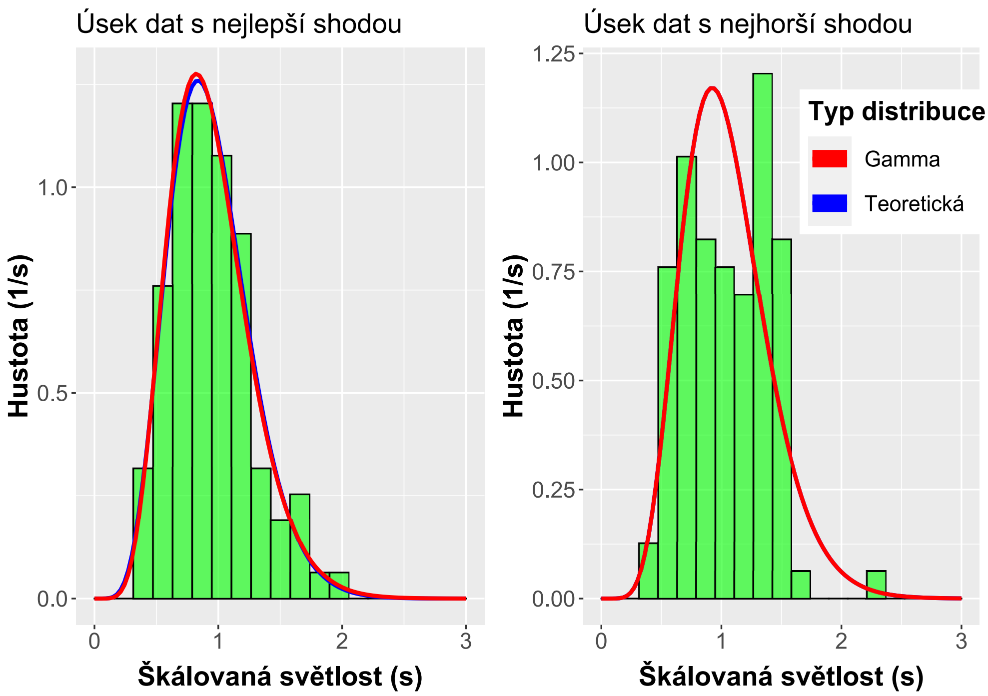
[7] referred to as the segment with the best and worst match, respectively

Acceptance Order				
	k=0	k=1	k=2	k=3
Pearson	0.064	0.047	0.030	0.000
Gamma	0.077	0.000	0.030	0.111
Number of tests	78	85	33	9

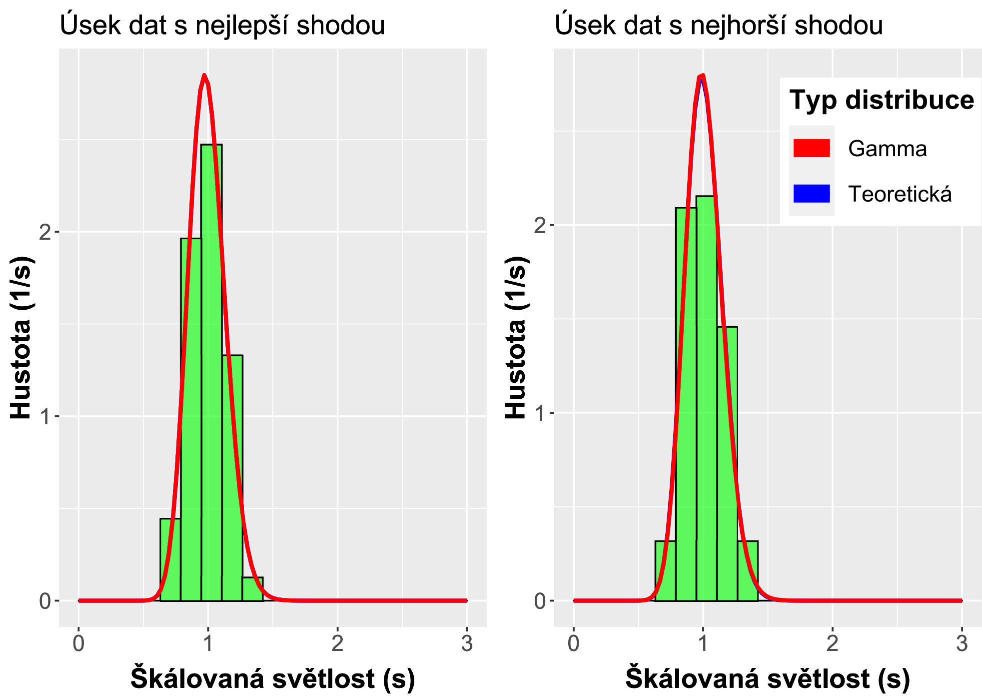
Tabulka 3.5: Ratios of rejection of Pearson's tests, respectively gamma tests of hypotheses considered in section 3.4 and the number of tests conducted for intersection No. 3 (both at a significance level of 5 %)



Obrázek 3.9: Data segments used in testing hypotheses of empirical clearances that showed the best and worst match for acceptance order $k = 0$ at intersection No. 1.



Obrázek 3.10: Data segments used in testing hypotheses of empirical clearances that showed the best and worst match for acceptance order $k = 0$ at intersection No. 3.



Obrázek 3.11: Data segments used in testing hypotheses of empirical clearances that showed the best and worst match for acceptance order $k = 3$ at intersection No. 1.

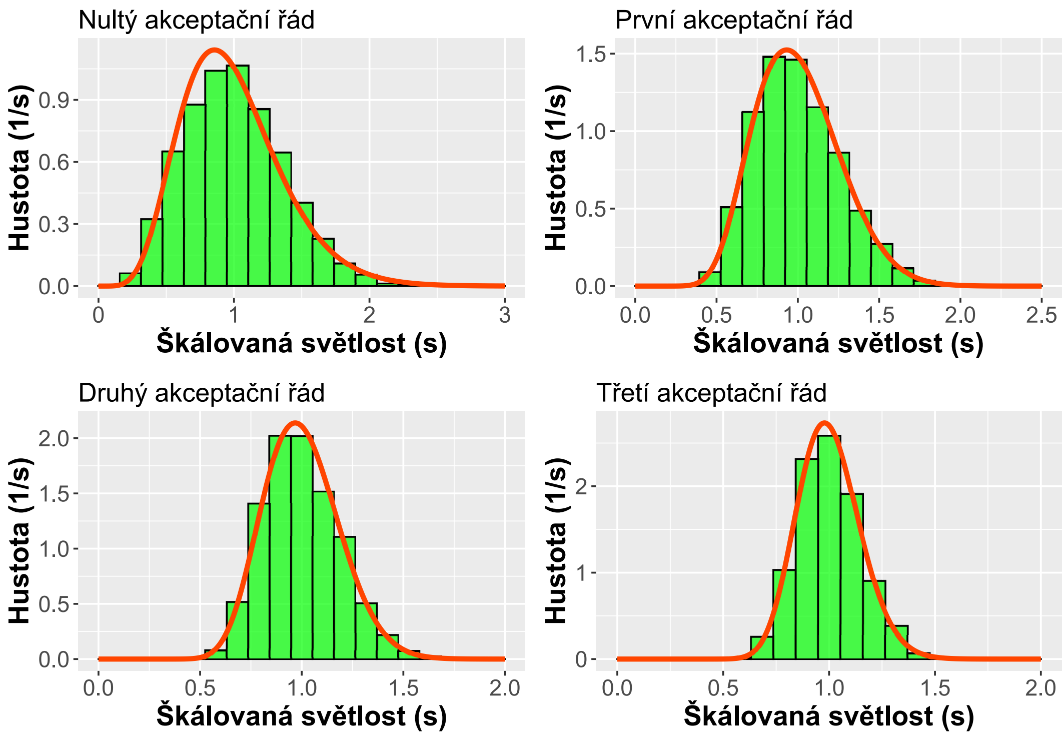
	Intersection No. 1			Intersection No. 2			Intersection No. 3		
	α	β	μ	α	β	μ	α	β	μ
k=0	1	< 0.001	6.683	15	6.056	6.353	4	0.843	7.080
k=1	15	21.130	4.066	15	21.138	3.773	15	21.151	4.161
k=2	1	2.000	28.003	4	9.413	23.440	15	36.497	5.325
k=3	13	44.730	15.958	1	3.000	47.032	1	3.000	43.071

Tabulka 3.6: Estimates of parameters of the distribution (2.10) for empirical clearances of various orders at all intersections. Values were obtained using Pearson's χ^2 test.

3.5 Parameter Estimates for Different Intersections

To get an idea of the values of the parameters α, β, μ and their dependence on the acceptance order for the examined intersections, we used Pearson's test for *unsegmented* clearances. For a given intersection and acceptance order $k \in \{0, 1, 2, 3\}$, we proceeded analogously to the testing in section 3.4, but the data sets were not divided into segments of 100 observations. The parameter values obtained from the tests are listed in Table 3.6 and the graphical comparison of the corresponding histograms and densities (2.10) with estimated parameters for intersection No. 3 is shown in Figure 3.12.

Let us recall that the parameter α was limited to a maximum value of 15 during the testing phase. It is possible that if we had allowed higher values for this parameter, we would have achieved a better fit in Pearson's test, and thus a different estimate for the triplet α, β, μ. However, the test was too unstable for a higher set limit of α.



Obrázek 3.12: Histograms of scaled clearances at intersection No. 3 for various acceptance orders and corresponding estimates of density (2.10) made using Pearson's χ^2 test

Conclusion

In the first part of this work, in accordance with the terminology of Gap Acceptance theory, we defined the basic mathematical model of an uncontrolled T intersection. Furthermore, we explained the importance of analyzing critical gaps in connection with computing intersection capacities and subsequently specified the considered intersection model by adding the assumption of gamma distributed light intervals X with parameters $\lambda, \mu > 0$ and critical light intervals Y with parameters $\lambda, \mu > 0$.

In the second part, for a slightly modified intersection model where we consider critical light intervals Y with Erlang distribution, we analytically derived the form of the density function of partial distribution of order $k \in \mathbb{N}_0$:

$$f_{D^{(k)}}(u) = (\beta + \mu)^{k\alpha+\lambda}\, u^{k\alpha+\lambda-1} e^{-(\beta+\mu)u} \left(\sum_{j=0}^{\alpha-1} \frac{(u\beta)^j}{(k\alpha + j)!} \right) \cdot \left(\sum_{l=0}^{\alpha-1} \frac{\beta^l}{(\beta + \mu)^l} \frac{\Gamma(k\alpha + l + \lambda)}{(k\alpha + l)!} \right)^{-1},$$

for $u > 0$. We then demonstrated the correctness of the derivation through graphs and simulations in the programming environment **R**, utilizing the Pearson's χ^2 goodness of fit test. Furthermore, for the modified model, we computed Siegloch's function $s(t)$:

$$s(t) = e^{-\beta t} \sum_{l=0}^{\alpha-1} \sum_{k=0}^{+\infty} k \frac{(\beta t)^{k\alpha+l}}{(k\alpha + l)!}, \quad \forall t > 0\,.$$

We plotted its behavior for various values of parameters α, β, revealing a linear asymptote that is commonly used as an approximation of Siegloch's function in practice.

In the third part, we formulated the hypothesis of empirical light intervals, which we statistically tested using simulations and the validated Pearson's goodness of fit test. In the case of empirical measurements, we had to estimate the parameters of the distribution from the null hypothesis H_0. We chose the Maximum Likelihood Estimation (MLE) method and used the function `nlm` from the `stats` library to minimize the negative log likelihood function of the random sample. After imposing constraints on the parameters $\alpha \in \mathbb{N}$ and $\lambda = \mu$ due to the nature of the `nlm` function, the provided estimates were unbiased and reliable for various parameter values. The Pearson's test was always performed for a data segment of size 100 observations. The rejection rates of the hypothesis against the total number of conducted tests for a specific intersection and acceptance order k at a 5% significance level were mostly below 0.05, except for a few exceptions. Therefore, we accepted the hypothesis of empirical light intervals as non rejected.

Simultaneously with the Pearson's test, a hypothesis test was conducted to verify if the empirical light intervals of acceptance order k follow gamma distribution. We utilized the `gamma_test` function from the `goft` library. According to the resulting p values, we could not refute that the empirical data are realizations of gamma distribution. Particularly, for higher acceptance orders k, they provide a very good approximation.

Bibliografie

[1] Transportation Research Board. *Capacity and Level of Service at Unsignalized Intersections: Final Report Volume 1 - Two-Way-Stop-Controlled Intersections*. The National Academies Press, 1996.

[2] Werner Brilon a Michael Großmann. “The New German Guideline for Capacity of Unsignalized Intersections”. In: *Intersections without Traffic Signals II*. Springer Berlin Heidelberg, 1991, s. 62–82.

[3] Werner Brilon, Ralph Koenig a Rod J. Troutbeck. “Useful estimation procedures for critical gaps”. In: *Transportation Research Part A: Policy and Practice* (1999), s. 161–186.

[4] M. Krbálek a M. Krbálková. “3s-Unification for Vehicular Headway Modeling”. In: *Proceedings of SPMS 2018* (2018).

[5] A. Miller. “Nine Estimators Of Gap-Acceptance Parameters”. In: 1971.

[6] J. C. Tanner. “A Theoretical Analysis of Delays at an Uncontrolled Intersection”. In: *Biometrika* (1962), s. 163–170.

[7] R. Troutbeck a W. Brilon. “Unsignalized Intersection Theory”. In: 1997.

[8] Josef Vladař. “Odhad kritických odstupů na neřízených ú rovňových křižovatkách metodou maximální věrohodnosti”. Diplomová práce. ZČU v Plzni, Fakulta aplikovaných věd, 2009.

[9] Helmut Wegmann. “A General Capacity Formula for Unsignalized Intersections”. In: *Intersections without Traffic Signals II*. Ed. Werner Brilon. Springer Berlin Heidelberg, 1991, s. 177–191.