

MSc Mathematics

Master's thesis

Fire Engine Travel Times

Analysis of Travel Times and Optimal Dispatching Decisions

by

Anne Frances Appelman

August 4, 2016

Supervisor: Dr. ir. Peter van de Ven

First examiner: Prof. dr. Rob van der Mei

Second examiner: Dr. Sandjai Bhulai

Department of Mathematics

Faculty of Sciences



Abstract

This master's thesis looks into the problems concerning travel times of fire engines. First, we start with the problem of finding a distribution of the travel times of fire engines, conditioned on the distance. Here, we fully describe and clarify a previously proposed model, and apply it to GPS-data of the fire department in the area of Amsterdam, the Netherlands. Second, we look at decision processes. We start with considering non-concurrent incidents, where we look at the minimum of different travel time distributions, to decide which two out of three fire engines we should send to an incident. We find that it is optimal to send fire engines with independent travel times. We then use this information when we consider concurrent incidents. In this case we use a discounted rewards infinite horizon Markov Decision Process to model how many fire engines should be sent to an incident when sending more fire engines is better for the current incident, since it reduces the expected travel time, but fire engines take time to return, so at the next incident, there will be less fire engines available. We prove that it is optimal to send more fire engines, when there is more than one fire engine available.

Title: Fire Engine Travel Times

Subtitle: Analysis of Travel Times and Optimal Dispatching Decisions

Author: Anne Frances Appelman, a.f.appelman@vu.nl, 1847341

Supervisor: Dr. ir. Peter van de Ven

First examiner: Prof. dr. Rob van der Mei

Second examiner: Dr. Sandjai Bhulai

Date: August 4, 2016

Department of Mathematics

VU University Amsterdam

de Boelelaan 1081, 1081 HV Amsterdam

<http://www.math.vu.nl/>

Contents

1. Introduction	5
1.1. Background	5
1.2. Literature	5
1.3. Overview thesis	6
2. Stochastic travel times at BWAA	9
2.1. Data Description	9
2.2. Data Analysis	10
2.3. Theory	13
2.3.1. Nonparametric approach	16
2.3.2. Parametric approach	17
2.3.3. Fit of the model	18
2.4. Filling the model	19
3. Dispatching: Non-concurrent incidents	25
3.1. T_i and T_j exponential and independent	25
3.2. T_i and T_j hyperexponential	26
3.3. T_i and T_j hypoexponential	31
3.3.1. T_i and T_j Erlang	36
3.3.2. T_i and T_j sums of different exponential random variables	38
4. Dispatching: Concurrent incidents	43
4.1. Problem description	43
4.2. Model	43
4.3. Policy iteration	45
4.3.1. T_i exponentially distributed	46
4.3.2. T_i uniformly distributed	48
4.3.3. $\log(T_i)$ Student's t distributed	49
4.4. Optimal decision	50
5. Conclusion	55
Bibliography	57
A. Matlab code	59

1. Introduction

1.1. Background

A central issue in emergency services, such as medical, fire, and police, is to be at the location of an incident fast. In case of life threatening emergencies, every minute counts. For a city like Amsterdam, the Netherlands, the fire department needs to be at the location of the fire within 6 to 10 minutes after a call has been placed, depending on the type of building (fireproof or not) and the function of the building (e.g. store versus dwelling house). The travel time of a fire engine is the largest component of the response time, and an accurate prediction of this travel time could make improvements possible.

Different kinds of improvements are possible. One could, for example, think of finding faster routes from the fire station to the fire. Also, relocating the fire stations, so that the area is better covered, can lead to an improvement (Van den Berg, 2016). And, when considering to send more than one fire engine to a fire, knowing how to optimally dispatch the fire engines could be an improvement. It is clear that for all of these improvements to be made possible, it is necessary to know as much as possible about the travel times of the fire engines.

1.2. Literature

Much research is done about analyzing travel times and dispatching vehicles in emergency services. The New York City RAND fire project, lasting from 1968 to 1975, was a successful research project, which made use of analytical and statistical modelling, which led to key changes at the New York City Fire Department (Larson, 1972; Archibald et al., 1979; Kolesar, 2012). A part of this research project was about the travel times of the fire engines. Kolesar et al. (1975) divide the travel time into an acceleration phase, a cruising speed phase, and a deceleration phase. They use these different phases to model the travel time depending on the distance, and they take the mean of the travel times as the expectation of the travel time. A similar research project on ambulance planning, is the Dutch REPRO project. This project, lasting from 2012 to 2016, considers new and efficient planning methods for ambulance services, and advanced decision models for relocating ambulances (Jagtenberg et al., 2015; Van Barneveld et al., 2015, 2016).

Budge et al. (2010) improve the model of Kolesar et al. (1975), by taking the median rather than the mean as the expectation of the travel time. Budge et al. (2010) argue that since travel times are non-negative, the distribution of the travel times is probably

skewed to the right, and that therefore it is better to take the median rather than the mean as the expectation of the travel time.

Later, Westgate et al. (2013) introduce a different model in which they look at the different street segments separately, and take differences in travel speed into account instead of looking at the total route as one stochastic variable. They use a Bayesian method to estimate the travel time distribution, and they compare that model to the model of Budge et al. (2010). They conclude that the Bayesian method seems to give more realistic results than the method of Budge et al. (2010), since the method of Budge et al. (2010) does not take into account the different speeds of different roads. For their case study, Westgate et al. (2013) use GPS-data and apply Markov Chain Monte Carlo simulation to find the true route, since their GPS-data is not accurate enough.

One could, of course, refine this model by not only looking at different roads separately, but also taking the transitions of one road to another into account. Jenelius and Koutsopoulos (2013), for example, see a trip as a combination of running travel times along links, and delays at intersections and traffic signals. They treat this delay as a deterministic penalty. They do not consider emergency services in particular, but give a model for general services.

In all the above studies, travel times are considered stochastic. Often, researchers assume travel times to be deterministic (Church and Davis, 1974; Ingolfsson et al., 2003; Maxwell et al., 2009). In this case, fire engines are either always on time or always too late, which is not realistic, since there is always some variability, because of weather, the time of day (heavy traffic), etcetera. A big part of the variability in travel times is hard or even impossible to parameterize completely, so stochastic travel times are the better choice. Ingolfsson et al. (2008) show that, by considering ambulance travel times between a particular station and demand point pair, for a total of 352 trips, stochastic travel times lead to a more realistic model.

Besides analyzing and modelling the travel times, a lot of research is done investigating the optimal dispatching of emergency vehicles. Swersey (1982) looks at the problem of how many fire companies need to be dispatched, in order to make sure that there are just enough fire engines sent to a fire, not too few and not too many. He gives a Markovian decision model to solve this problem. Ignall et al. (1982) give an initial dispatch algorithm, that deals with the uncertainty about the true nature of the incident when the dispatch decision is made, as well as the possible competition between different incidents for the same companies that come with rising alarm rates. It considers both losses at incidents and the costs of needless company responses.

1.3. Overview thesis

In this master's thesis, we will first look at the distribution of travel times of a specific kind of fire engine: the water tender. In Chapter 2, we look at GPS-data of the fire department in the area of Amsterdam, the Netherlands. We use an already existing model, which we explain in detail, to obtain a stochastic distribution which fits our

data well. In Chapter 3, we look at minima of different random variables. Using these minima, we show that it is always optimal to send two water tenders whose travel times are independent rather than dependent, when the expectation of the individual travel times are the same in the dependent case and in the independent case. We also show that in some situations it is smarter to send a water tender with a larger expected travel time, than one with a shorter expected travel time if it yields independence. In Chapter 4, we combine the results of Chapters 2 and 3 by modelling a discounted rewards infinite horizon Markov Decision Process. In this chapter, we look at a dispatching problem. We have N water tenders, and we want to know how many we should send to a fire, knowing that sending more water tenders at this moment is better for the current incident, but results in having fewer water tenders for future incidents. We conclude this chapter with proving that the more water tenders we have, the more we should send to a fire.

2. Stochastic travel times at BWAA

In this chapter, we take a closer look at the travel times of emergency travel to a serious incident. We will look at some GPS-data, and apply it to the model Budge et al. (2010) which fits a distribution to the travel time data conditioned on the distance. We choose this model, because it is very clear and insightful. We will describe this model in detail, since in the original paper by Budge et al. (2010) much detail in the derivation of the model is missing, making it difficult to assess the correctness of the model. We conclude this chapter with comparing our results to the results of Budge et al. (2010).

2.1. Data Description

The fire department in the area of Amsterdam, ‘Brandweer Amsterdam-Amstelland’ (BWAA), provided a large dataset of GPS-data of fire engines from November 4, 2015 until February 24, 2016. This dataset contains the following information: Licence plate; Date and time; Description event; Mileage; Speed; Latitude; Longitude; Street; Address; Zip code; Town. BWAA has a variety of types of cars, and of every type there are multiple cars. Examples of these different types are water tenders, turntable ladders, and service cars. The licence plate is necessary to know the type of the car. ‘Description event’ could say the following things: Start driving; Driving; End driving; Start flashing lights; End flashing lights; Sirens. Date and time is measured every time there is a different event, and when the description is ‘driving’, it gives the date and time every 10 or every 30 seconds (this differs per car). So, we only get a data point when the motor of a car is actually running. We could thus see this data set as a list of different locations, each belonging to a specific car with at a specific time doing something specific at that time (driving, turning sirens or flashing lights on or off).

BWAA distinguishes three levels of emergency:

- Level 1.* This is a fire to which there has to be sent a fire engine as soon as possible;
- Level 2.* This is a small fire which doesn’t spread out, like a trash can in the street, such that there has to be sent a fire engine, but not with a great hurry;
- Level 3.* This is a small incident for which there is not a great hurry, like a cat in a tree.

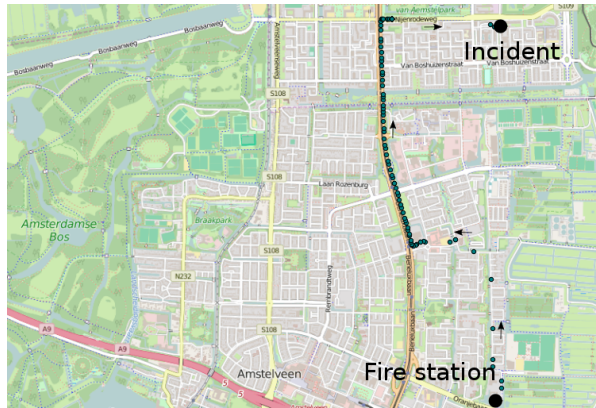
We hope that we can find a model, which gives us the expected travel time based on the distance of the ride. We are interested in the travel times of fire engines to incidents, at which a fire engine has priority to all the other traffic. So, we only want rides of emergency level 1. These rides need to start at the fire station, and end at the incident.

2.2. Data Analysis

We start with GPS-data of fire engines from November 4, 2015 until February 24, 2016. We will analyze the rides to fires. As sketched above, the data contains a great deal of information which is not of interest to us, so we throw away much of the data and only keep what we need:

- *Water tenders.* The dataset contains travel data of all different cars from the fire department. We are only interested in the rides to fires, so that means that we only want the water tenders, since water tenders drive to incidents of emergency level 1. We discard the travel data of the other cars of the fire department. Turntable ladders could of course also drive to fires, but they would also drive to a cat in a tree, and since we are not interested in that kind of emergency, we do not consider these cars.
- *Flashing lights.* Since water tenders can also ride when there is no emergency (to tank gas or to go to a grocery store), we only keep those rides where the flashing lights were on.
- *Rides.* Since the dataset is basically a list of coordinates of the different cars, we do not have the rides. So we cut the list of coordinates in shorter lists, which start at ‘start driving’, and end at ‘end driving’.
- *Actual rides.* Every water tender is tested every morning. When a water tender is tested, the engine starts running, and the sirens will go on and off for a short period. Sometimes the water tenders even make a short test ride through the immediate neighbourhood. We find that one water tender (located at Osdorp) gets tested every day, every two hours. We delete all these test rides from our data set.
- *One way rides.* Often, a water tender does not turn of the engine when he arrives at a fire. Sometimes the water tender will stand at the location of the fire for a couple minutes, sometimes hours, before turning of the engine. And sometimes the engine turns of when the water tender is already back at the fire station, creating a round trip. We thus have to shorten a great deal of the rides, such that we only consider the part starting at the fire station until arriving at the fire. We also have rides where the water tender was called back before arriving at the fire. Since these water tenders never arrived at the fires, we delete these rides from our dataset. To do all this, we actually have to plot all the rides, since staying at the same location for a couple minutes could also mean that the water tender was stuck in traffic. We plot the rides in QGIS (QGIS Development Team, 2009), which is an Open Source Geographic Information System (GIS) licensed under the GNU General Public License. An example of a one way ride and a round trip are shown in Figure 2.1. Note that every dot represents a data point, and that a data point is generated every 10 or 30 seconds, or, when some event has happened, like turning on the flashing lights or sirens. So, a lot of dots close to each other indicate

here the fast ride from a fire station to a fire, and not the slower ride from the fire returning to the fire station.



(a) One way ride from fire station Amstelveen at 13/11/2015.



(b) Round trip from fire station Hendrik at 15/01/2015.

Figure 2.1.: Different rides from different fire stations plotted in QGIS.

- *Abnormalities.* Among the remaining rides are some abnormalities. Examples of such abnormalities are a ride from one fire station to another fire station, a ride to a gas station, a ride where the starting point is far away from the second point, while the time spent between these two points is only a few seconds, and a collection of points which looks nowhere near a ride. We delete all these abnormalities from our dataset.

It takes much time to clean the data, so we are not able to adjust all the data we got from BWAA. Instead, we only adjust the data of the November 2015 and eight days of January 2016. We are left with a dataset of 426 rides.

For the initial analysis, we consider a subset of our data: GPS data of fire engines from January 11, 2016 until January 18, 2016. This subset contains 105 rides. Figure 2.2 shows a plot of the travel times against the distances together with the logarithm of the travel times against the distances. We see that there seems to be a positive correlation between the distance and the travel times: the longer the distance, the

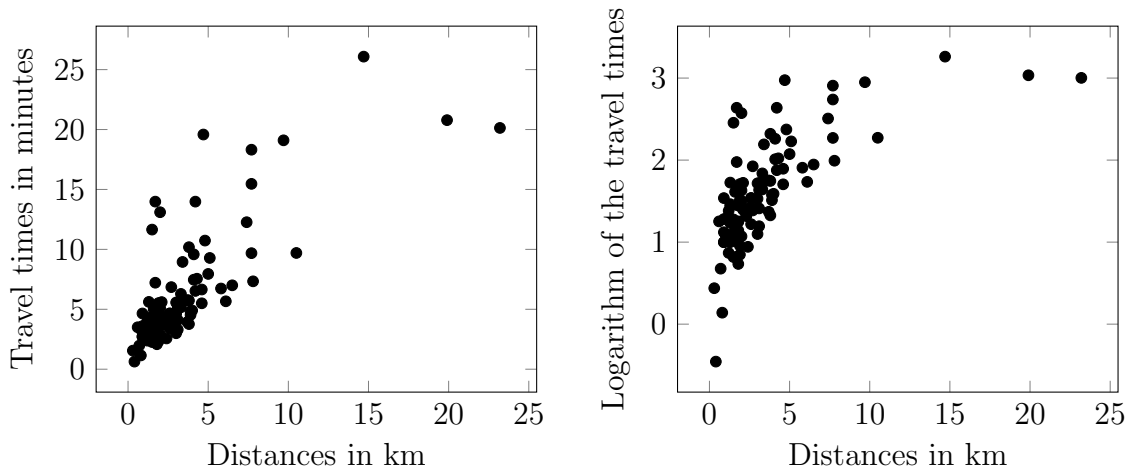


Figure 2.2.: Scatterplots of travel times against the distances and the logarithm of the travel times against the distances.

longer the travel times. Most of the distances and travel times are short, which is to be expected, since there are 19 stations in a region of 282 km^2 . There seems to be much variation, but we do not know whether or not this is to be expected. We will now look at the logarithm of the travel times. We do this because of the following. We know that, because of the Central Limit Theorem, averages of random variables, independently drawn from independent distributions, converge to a normal distribution. When these random variables are all positive, it is likely that they do not converge to a normal, but to a log-normal distribution, since it is likely that they are skewed. So, if we look at the travel time as a random variable T_d , dependent on distance d , T_d would be log-normally distributed, and therefore $\log(T_d)$ would be normally distributed. We also know that a small random sample from a normal distribution is Student's t distributed - only when the sample size is sufficiently large, it will still follow a normal distribution. Now we look again at our data. We call our set of travel times \tilde{T}_d , where we let them be dependent on the distance. Since $|\tilde{T}_d|$ is not large, but only 105, we assume that $\log(\tilde{T}_d)$ follows a Student's t distribution, rather than a normal distribution. In R (RStudio Team, 2015) we fit a normal distribution and a Student's t distribution to the logarithm of our travel times \tilde{T}_d . For the normal distribution it gives $\text{fit}_N(\log(\tilde{T}_d)) \sim N(\mu, \sigma) = N(1.59, 0.62)$ and for the Student's t distribution it gives $\text{fit}_t(\log(\tilde{T}_d)) \sim t(\mu, \sigma, \tau) = t(1.54, 0.49, 4.77)$. A picture of the fit of these distributions to the empirical distribution function of $\log(\tilde{T}_d)$ is shown in Figure 2.3. We use the Kolmogorov-Smirnov test to evaluate the fit of these distributions. The results are shown in the table below.

Distribution	D/W-statistic	p -value
$N(1.59, 0.62)$	0.96203	0.0042
$t(1.54, 0.49, 4.77)$	0.06651	0.7417

We conclude that we reject the hypothesis that $\log(\tilde{T}_d)$ follows a normal distribution, and we do not reject the hypothesis that $\log(\tilde{T}_d)$ follows a Student's t distribution.

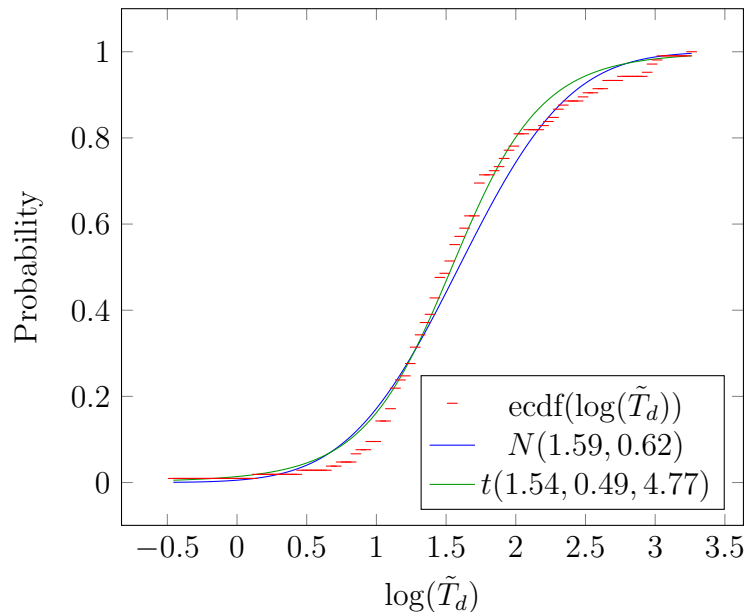


Figure 2.3.: Empirical distribution function of $\log(\tilde{T}_d)$ with the fitted normal distribution and the fitted Student's t distribution.

2.3. Theory

In our upcoming theoretic analysis, we follow the paper of Budge et al. (2010) which is an example of a study on the distribution of travel times of emergency vehicles. The reason that we choose this model out of all the different models mentioned in Chapter 1, is that this model is very clear and insightful. Budge et al. (2010) also assume that the travel times follow a Student's t distribution, and that these travel times depend on the distances travelled. Since the mathematics used in the model of Budge et al. (2010) is not that clear - they skip a lot of steps we think are necessary to understand what is done - we will compute every calculation in detail. Budge et al. (2010) use data from the Emergency Medical Service (EMS) system in Calgary, Alberta. This data consists, after removing outliers, of 6,886 travel time observations. They found that a shifted and scaled Student's t distribution provides a good fit to the log of the travel time distributions. They propose a regression model in which the log of the travel time T_d , where T_d is a random variable, follows a Student's t distribution that is shifted by $\log(m_d)$ and scaled by c_d :

$$\log(T_d) = \log(m_d) + c_d\varepsilon,$$

where m_d and c_d are unknown parameters depending on the distance, and ε follows a centered Student's t distribution with τ degrees of freedom, so $\mathbb{E}[\varepsilon] = 0$, and $\text{var}[\varepsilon] =$

$\tau/(\tau - 2)$, where we assume $\tau > 2$. This gives us

$$\begin{aligned}
\mathbb{E}[\log(T_d)] &= \mathbb{E}[\log(m_d) + c_d\varepsilon] \\
&= \mathbb{E}[\log(m_d)] + \mathbb{E}[c_d\varepsilon] \\
&= \log(m_d) + c_d\mathbb{E}[\varepsilon] \\
&= \log(m_d),
\end{aligned}$$

and

$$\begin{aligned}
\text{var}[\log(T_d)] &= \mathbb{E}[\log(T_d)^2] - \mathbb{E}[\log(T_d)]^2 \\
&= \mathbb{E}[(\log(m_d) + c_d\varepsilon)^2] - \log(m_d)^2 \\
&= \mathbb{E}[(\log(m_d)^2 + 2\log(m_d)c_d\varepsilon + c_d^2\varepsilon^2)] - \log(m_d)^2 \\
&= \mathbb{E}[(\log(m_d)^2) + \mathbb{E}[2\log(m_d)c_d\varepsilon] + \mathbb{E}[c_d^2\varepsilon^2]] - \log(m_d)^2 \\
&= \log(m_d)^2 + c_d^2\mathbb{E}[\varepsilon^2] - \log(m_d)^2 \\
&= c_d^2\tau/(\tau - 2),
\end{aligned}$$

since $\tau/(\tau - 2) = \text{var}[\varepsilon] = \mathbb{E}[\varepsilon^2] - (\mathbb{E}[\varepsilon])^2 = \mathbb{E}[\varepsilon^2]$. If we reverse the log transformation, we obtain

$$\begin{aligned}
T_d &= \exp(\log(m_d) + c_d\varepsilon) \\
&= \exp(\log(m_d)) \exp(c_d\varepsilon) \\
&= m_d \exp(c_d\varepsilon).
\end{aligned}$$

Now we want to find the median of T_d , i.e. we want to find the x for which $\mathbb{P}(T_d \leq x) \geq \frac{1}{2}$ and $\mathbb{P}(T_d \geq x) \geq \frac{1}{2}$:

$$\begin{aligned}
\mathbb{P}(T_d \leq x) \geq \frac{1}{2} &\iff \mathbb{P}(m_d \exp(c_d\varepsilon) \leq x) \geq \frac{1}{2} \\
&\iff \mathbb{P}\left(\exp(c_d\varepsilon) \leq \frac{x}{m_d}\right) \geq \frac{1}{2} \\
&\iff \mathbb{P}(c_d\varepsilon \leq \log(x) - \log(m_d)) \geq \frac{1}{2} \\
&\iff \frac{\log(x) - \log(m_d)}{c_d} = 0 \\
&\iff x = m_d,
\end{aligned}$$

where in the second-last step we use that the median of a centered Student's t distribution is zero. And

$$\begin{aligned}
\mathbb{P}(T_d \geq x) \geq \frac{1}{2} &\iff \mathbb{P}(m_d \exp(c_d \varepsilon) \geq x) \geq \frac{1}{2} \\
&\iff \mathbb{P}\left(\exp(c_d \varepsilon) \geq \frac{x}{m_d}\right) \geq \frac{1}{2} \\
&\iff \mathbb{P}(c_d \varepsilon \geq \log(x) - \log(m_d)) \geq \frac{1}{2} \\
&\iff \frac{\log(x) - \log(m_d)}{c_d} = 0 \\
&\iff x = m_d,
\end{aligned}$$

So, m_d is the median of T_d . Hence, m_d is both is the median and the expected value of T_d . Finding the coefficient of variation of T_d , CV_{T_d} , is more work. According to Rigby and Stasinopoulos (2006) we have, for T_d Student's t distributed,

$$\text{CV}_{T_d} = \frac{3(F_{T_d}^{-1}(0.75) - F_{T_d}^{-1}(0.25))}{4m_d},$$

where F_{T_d} is the cumulative distribution function for T_d . Furthermore,

$$F_{T_d}^{-1}(\alpha) = m_d \exp(c_d t_{\tau, \alpha}),$$

where $t_{\tau, \alpha}$ is the 100α centile of a t distribution with τ degrees of freedom. We assume $c_d t_{\tau, \alpha}$ to be small enough, such that the Taylor expansion of $\exp(c_d t_{\tau, \alpha})$ can be approximated by $1 + c_d t_{\tau, \alpha} + \frac{1}{2}(c_d t_{\tau, \alpha})^2$. This gives us

$$\begin{aligned}
F_{T_d}^{-1}(0.75) - F_{T_d}^{-1}(0.25) &= m_d(\exp(c_d t_{\tau, 0.75}) - \exp(c_d t_{\tau, 0.25})) \\
&= m_d\left(1 + c_d t_{\tau, 0.75} + \frac{1}{2}(c_d t_{\tau, 0.75})^2\right. \\
&\quad \left. - 1 - c_d t_{\tau, 0.25} - \frac{1}{2}(c_d t_{\tau, 0.25})^2\right) \\
&\approx m_d\left(c_d t_{\tau, 0.75} + \frac{1}{2}(c_d t_{\tau, 0.75})^2\right. \\
&\quad \left. - c_d t_{\tau, 0.25} - \frac{1}{2}(c_d t_{\tau, 0.25})^2\right).
\end{aligned}$$

According to Johnson et al. (1994) we have

$$t_{\tau, \alpha} \approx \Phi^{-1}(\alpha) + \Phi^{-1}(\alpha)(\Phi^{-1}(\alpha)^2 + 1)/(4\tau),$$

so

$$\begin{aligned}
t_{\tau, 0.75} &\approx 0.674 + 0.674((0.674)^2 + 1)/(4\tau) = 0.674 + 0.245/\tau, \\
t_{\tau, 0.25} &\approx -0.674 - 0.674((-0.674)^2 + 1)/(4\tau) = -0.674 - 0.245/\tau.
\end{aligned}$$

Hence,

$$\begin{aligned} c_d t_{\tau,0.75} &= 0.674 + 0.245/\tau \\ &= -c_d t_{\tau,0.25}, \end{aligned}$$

and

$$\begin{aligned} (c_d t_{\tau,0.75})^2 &= (0.674 + 0.245/\tau)^2 \\ &= (-0.674 - 0.245/\tau)^2 \\ &= (c_d t_{\tau,0.25})^2. \end{aligned}$$

Thus

$$F_{T_d}^{-1}(0.75) - F_{T_d}^{-1}(0.25) \approx 2m_d c_d (0.674 + 0.245/\tau).$$

This gives us

$$\begin{aligned} \text{CV}_{T_d} &= \frac{3(F_{T_d}^{-1}(0.75) - F_{T_d}^{-1}(0.25))}{4m_d} \\ &\approx \frac{3 \cdot 2m_d c_d (0.674 + 0.245/\tau)}{4m_d} \\ &= \frac{3}{2} c_d (0.674 + 0.245/\tau) \\ &= c_d (1.012 + 0.368/\tau). \end{aligned}$$

So, we have that c_d is approximately equal to a coefficient of variation of T_d .

Concluding, we have seen that when $\log(T_d)$ is Student's t distributed, we can write T_d as $T_d = m_d \exp(c_d \varepsilon)$, where m_d is the median of T_d , c_d is approximately equal to the coefficient of variation of T_d , and ε is a Student's t distributed error with τ degrees of freedom.

To find the functions m_d and c_d , Budge et al. (2010) uses two different approaches: a parametric and a nonparametric approach.

2.3.1. Nonparametric approach

In the nonparametric approach, Budge et al. (2010) use the regression model

$$t_i = m_{d_i} \exp(c_{d_i} \varepsilon_i), \quad i = 1, \dots, n,$$

where n is the number of observations, t_i and d_i are the travel time and distance for observation i respectively, and ε_i follows a Student's t distribution with τ degrees of freedom. They assume that the functions m and c are twice continuously differentiable, but otherwise arbitrary. m should be seen as a function for which we have $m(d) = m_d$. They estimate the functions m and c by maximizing the penalized log likelihood function,

$$l_p = l - \frac{1}{2} \left(\lambda_m \int_{-\infty}^{\infty} m''(u)^2 du + \lambda_c \int_{-\infty}^{\infty} (\log c)''(u)^2 du \right),$$

where m'' and $(\log c)''$ are the second derivatives of m and $\log c$ respectively, $l = \sum_{i=1}^n l_i$ is the log likelihood for the data, and l_i is the log likelihood for observation i .

2.3.2. Parametric approach

For the parametric approach, Budge et al. (2010) take the model of Kolesar et al. (1975) as a basis. Kolesar et al. (1975) assume that a fire engine accelerates from the origin at a rate a , until it reaches cruising velocity v_c . Then, the vehicle starts to decelerate with rate a coming to a stop at the destination. The cruising velocity is reached at distance $d_c = v_c^2/(2a)$ and time $t_c = v_c/a$. The median of the travel time T , as a function of the distance d , is

$$\text{median}[T_d] \equiv m_d = \begin{cases} 2\sqrt{d/a} & d \leq 2d_c, \\ v_c/a + d/v_c & d > 2d_c. \end{cases}$$

Kolesar et al. (1975) use this relationship to model the mean travel time, Budge et al. (2010) use this relationship to model the median travel time, because of the skewness of the data. This model thus assumes that for short trips, the travel time increases with the square root of the distance, and for long trips, the travel time increases linearly with the distance. Budge et al. (2010) correspond the acceleration and deceleration phases to travel on residential or arterial roads, and the cruising speed phase to highway travel. They parametrize c_d as follows:

$$c_d = \frac{\sqrt{b_0(b_2 + 1) + b_1(b_2 + 1)m_d + b_2m_d^2}}{m_d},$$

where b_0 represents the variability at the beginning and end of a trip, which is independent of the distance of the trip, b_1 represents short-term variation in speed within a trip, and b_2 represents a measure of variability due to factors that are not included in the model, such as weather or traffic conditions. For a motivation of this exact parametrization of c_d , we refer to Budge et al. (2010). They then fit this parametric model by maximizing the log likelihood $l = \sum_{i=1}^n l_i$. Rigby and Stasinopoulos (2006) obtained this log likelihood for observation i as follows. The log likelihood for observation i is defined as $\log(f_{T_{d_i}}(t_i))$, where $f_{T_{d_i}}$ is the probability density function of T_{d_i} . We can derive $f_{T_{d_i}}$ from the probability density function of ε_i , since $T_{d_i} = m_{d_i} \exp(c_{d_i}\varepsilon_i)$. So we have

$$f_{\varepsilon_i}(x) = \left| \frac{dt_i}{dx} \right| f_{T_{d_i}}(t_i), \quad (2.1)$$

where

$$f_{\varepsilon_i}(x) = \frac{\Gamma((\tau + 1)/2)}{\Gamma(\tau/2)\sqrt{\tau\pi}} \left(1 + \frac{x^2}{\tau} \right)^{-(\tau+1)/2}$$

by definition, since ε_i follows a centered Student's t distribution with τ degrees of freedom. When we work out (2.1) we obtain

$$\begin{aligned} f_{\varepsilon_i}(x) &= \left| \frac{dt_i}{dx} \right| f_{T_{d_i}}(t_i) \\ &= m_{d_i} c_{d_i} \exp(c_{d_i} x) f_{T_{d_i}}(t_i) \\ &= t_i \exp(-c_{d_i} x) c_{d_i} \exp(c_{d_i} x) f_{T_{d_i}}(t_i) \\ &= t_i c_{d_i} f_{T_{d_i}}(t_i). \end{aligned}$$

Now we can calculate the log likelihood:

$$\begin{aligned} l_i &= \log(f_{T_{d_i}}(t_i)) \\ &= \log\left(\frac{1}{t_i c_{d_i}} f_{\varepsilon_i}(x)\right) \\ &= -\log(t_i) - \log(c_{d_i}) + \log(f_{\varepsilon_i}(x)) \\ &= -\log(t_i) - \log(c_{d_i}) + \log\left(f_{\varepsilon_i}\left(\frac{1}{c_{d_i}} \log\left(\frac{t_i}{m_{d_i}}\right)\right)\right). \end{aligned}$$

Budge et al. (2010) then use Microsoft Excel's Solver to maximize l over a, v_c, b_0, b_1 and b_2 , for a fixed integer value of τ , and then perform a grid search over τ .

2.3.3. Fit of the model

To measure the fit of the model, Budge et al. (2010) use a measure which is similar to a coefficient of determination. It is defined as follows. They generate N random samples (they take $N = 50$) each containing 10% of the entire sample. These samples are training samples. The remaining 90% are the holdout samples. For each sample (which consists of a training sample with a corresponding holdout sample) they calculate

$$R^2 = 1 - \frac{\sum_{i=1}^{n_{\text{holdout}}} |T_{d_i} - \hat{m}_{d_i}|}{\sum_{i=1}^{n_{\text{holdout}}} |T_{d_i} - \bar{m}|}, \quad (2.2)$$

with \hat{m}_{d_i} the fitted median estimate based on the training sample, \bar{m} the empirical median of the holdout sample and where the summations are over the holdout sample. From the 50 R^2 values Budge et al. (2010) generate this way, they choose the R^2 value which is the median to be the coefficient of determination, and they use the training sample belonging to this R^2 value to estimate the parameters, which they then compare to the parameters estimated using the entire sample. Note that if \hat{m}_{d_i} is a good fit, we have that $|T_{d_i} - \hat{m}_{d_i}| < |T_{d_i} - \bar{m}|$, so a higher value of R^2 indicates a better fit of \hat{m}_{d_i} . For their data they found $R^2 = 35.77\%$ for the nonparametric model, and $R^2 = 35.53\%$ for the parametric model. They conclude that the parametric model captures the important features of the travel time distribution almost as well as the nonparametric model, where the last model does not restrict the forms of the median and coefficient of variation functions. The benefit of the parametric model is that it is more insightful, and since it is almost as good as the nonparametric model, Budge et al. (2010) claims that this parametric model is a good model.

2.4. Filling the model

In this section we apply the model, discussed in the section above, to our GPS-data, in order to fit a distribution to our travel times, dependent on the distance. Since Budge et al. (2010) found that their parametric approach is almost as good as their nonparametric approach, we choose to work with their parametric model, since we think that it is more insightful. We will apply the model to our first data sample (the eight days in January) and then use the parameters we will find for this data sample, for all of our data.

In Section 2.2 we concluded that we do not reject the hypothesis that $\log(\tilde{T}_d)$ is Student's t distributed. Now we want to see if the logarithm of our data is Student's t distributed, with a location and scaling that both depend on the distance, i.e.,

$$\log(\tilde{T}_d) = \log(m_d) + c_d \varepsilon,$$

where

$$m_d = \begin{cases} 2\sqrt{d/a} & d \leq 2d_c, \\ v_c/a + d/v_c & d > 2d_c, \end{cases}$$

$$c_d = \frac{\sqrt{b_0(b_2 + 1) + b_1(b_2 + 1)m_d + b_2m_d^2}}{m_d},$$

and ε follows a centered Student's t distribution with τ degrees of freedom.

We start with modelling m_d . For this, we use the least squares method. We choose the least squares method, since it is easier to compute by hand and easier to program than the maximum likelihood method Budge et al. (2010) use. The least squares estimator $\hat{\theta}$ minimizes the sum of squares

$$S(\theta) = \sum_{i=1}^n (\log(T_i) - f(d_i, \theta))^2$$

with respect to θ , where $\theta = \{v_c, a\}$, $f(d, \theta) = \log(m_d)$, and n is the number of rides. We minimize $S(\theta)$ by considering only the rides for which $d > 2d_c$ (since then we have all the parameters) and then solving the set of normal equations

$$\sum_{i=1}^n \frac{\partial f(d, \theta)}{\partial v} (\log(T_i) - f(d_i, \theta)) = 0,$$

$$\sum_{i=1}^n \frac{\partial f(d, \theta)}{\partial a} (\log(T_i) - f(d_i, \theta)) = 0.$$

For this, we must guess a d_c . Based on the road network of the city of Amsterdam, and the fact that we only consider rides within Amsterdam, we think that the cruising

velocity v_c is not very high. The highways are used rarely, so we consider v_c as corresponding with high street travel. Based on this assumption of the meaning of v_c , the roads in Amsterdam, and the location of the fire stations, we assume that $d_c < 0.9$ km. Of course, we do not want to take a d_c which is too small or too large. We think that choosing d_c too small has more negative consequences than choosing it too large, so we let $d_c = 1$ km just to be on the safe side (and thus only consider the T_i 's for which $d \geq 2$, which are 60 rides).

For this subset, we get

$$\begin{aligned} f(d, \theta) &= \log(v_c/a + d/v_c), \\ \frac{\partial f(d, \theta)}{\partial v} &= \frac{1/a - d/v_c^2}{v_c/a + d/v_c}, \\ \frac{\partial f(d, \theta)}{\partial a} &= \frac{-v_c/a^2}{v_c/a + d/v_c}. \end{aligned}$$

Solving the set of normal equations explicitly is not possible, since f is nonlinear in θ . We therefore use an iterative method to approximate the solution. Instead of using Microsoft Excel Solver (as Budge et al. (2010) do), we use the Gauss-Newton method with initial guess $\theta_0 = \{v_c, a\} = \{0.86, 0.50\}$, where v_c is in measured in kilometers per minute, and a in kilometers per (minutes)². This initial guess is partly based on the final parameters of Budge et al. (2010), and partly on our assumptions of maximal travel speed in Amsterdam. The Gauss-Newton method is an iterative procedure, which stops when the estimate does not change anymore. Sometimes, there is more than one equilibrium to which the Gauss-Newton method can converge, so a judicious choice of θ_0 is important. Our choice of θ_0 gives us the following results:

$$\begin{aligned} v_c &= 0.86, \\ a &= 0.49, \\ d_c &= 0.79, \\ t_c &= 1.76. \end{aligned}$$

Now, we can estimate m_d . If our estimation of m_d is good, then $\log(T_d) \approx \log(m_d)$, since for $\log(T_d) = \log(m_d) + c_d\varepsilon$, we have $\mathbb{E}[\log(T_d)] = \log(m_d)$, because $\mathbb{E}[\varepsilon] = 0$. To see if our estimation of m_d is good, we look at the QQ-plot of $\log(\tilde{T}_d)$ against $\log(m_d)$, see Figure 2.4, and the fit of $\log(m_d)$ to $\log(\tilde{T}_d)$, shown in Figure 2.5. The QQ-plot is right-tailed, which indicates that $\log(\tilde{T}_d)$ has more extreme values than $\log(m_d)$. This is exactly what we would expect, since $\log(\tilde{T}_d)$ has variability while $\log(m_d)$ has no variability. The fit of $\log(m_d)$ to $\log(\tilde{T}_d)$ in Figure 2.5 looks good. We conclude that $\log(m_d)$ is a good estimator for our data $\log(\tilde{T}_d)$. The next thing we should do, is find the right parameters b_0 , b_1 and b_2 , so we can estimate c_d . This is rather difficult and takes much time, so we leave it with just a first guess of these parameters and look how good the results are. We choose $(b_0, b_1, b_2) = (0.03, 0.0001, 0.04)$, based on the values Budge et al. (2010) obtained for these parameters, and the differences between the values of v_c and a that we found and that Budge et al. (2010) found. If $\log(\tilde{T}_d)$

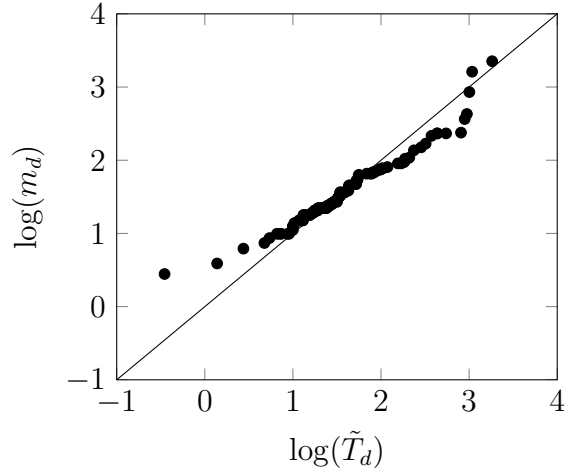


Figure 2.4.: QQ-plot of $\log(\tilde{T}_d)$ against $\log(m_d)$.

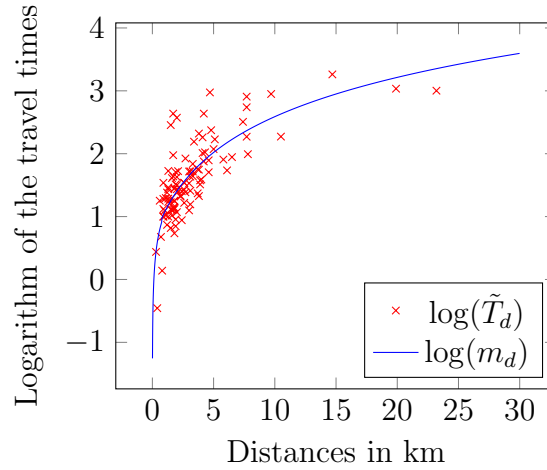


Figure 2.5.: The fit of $\log(m_d)$ to $\log(\tilde{T}_d)$.

has a Student's t distribution with location m_d , scaling c_d , and τ degrees of freedom, then $\frac{\log(\tilde{T}_d) - \log(m_d)}{c_d}$ has a centered Student's t distribution with τ degrees of freedom. So, we test if $\frac{\log(\tilde{T}_d) - \log(m_d)}{c_d}$ has a centered Student's t distribution with τ degrees of freedom. Since we can not derive τ , we have to choose τ , and based on the results of Budge et al. (2010), and on our own evaluation of the data, we choose $\tau = 4$. The Kolmogorov-Smirnov test gives a D -statistic of 0.081107 and a p -value = 0.4945. The fit of the centered Student's t distribution to $\frac{\log(\tilde{T}_d) - \log(m_d)}{c_d}$ is shown in Figure 2.6. From this figure and the Kolmogorov-Smirnov test, we conclude that the parametric model describes the data well.

Now we will look at the rest of our data. We consider the month November of the year 2015, November 4 until November 30 to be precise, which contains 321 rides. A plot of the travel times against the distances together with the logarithm of the travel times

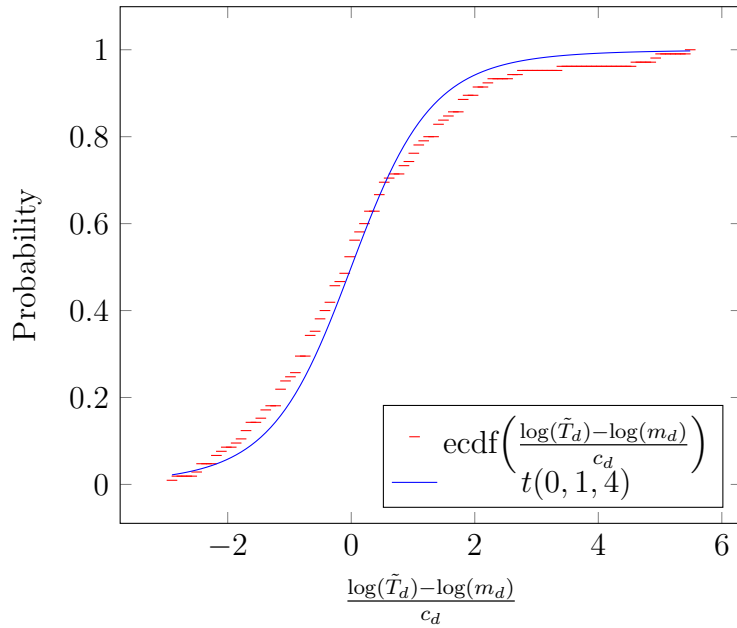


Figure 2.6.: Empirical distribution function of $\frac{\log(\tilde{T}_d) - \log(m_d)}{c_d}$ with the fitted centered t distribution with $\tau = 4$ degrees of freedom.

against the distances, is shown in Figure 2.7. We will now apply the parametric model,

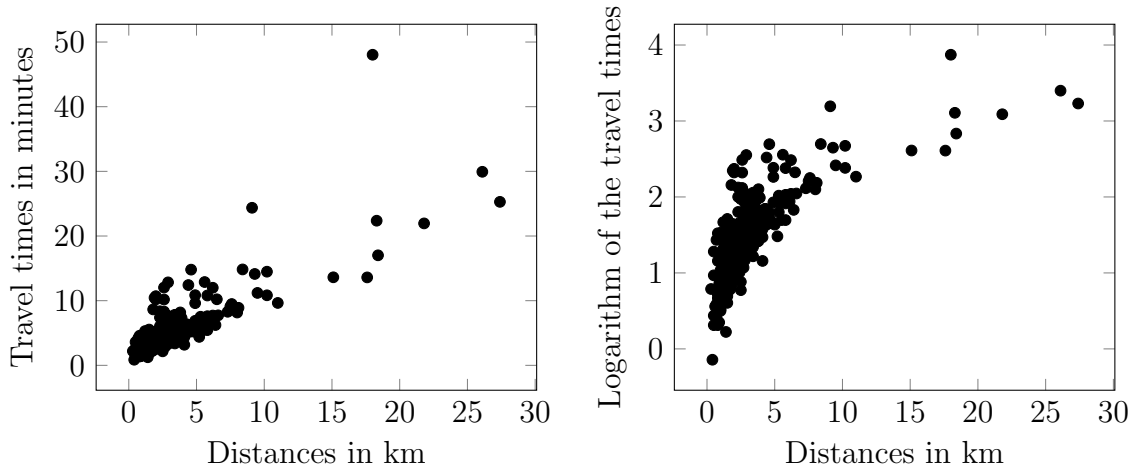


Figure 2.7.: Scatterplots of travel times against the distances and the logarithm of the travel times against the distances for 04/11/2015 until 30/11/2015.

and use the parameters obtained for the data of January. A fit of $\log(m_d)$ to $\log(\tilde{T}_d)$ for the data of November is shown in Figure 2.8. This looks fine, but when we take a look at the empirical distribution function of $\frac{\log(\tilde{T}_d) - \log(m_d)}{c_d}$, we see that a the centered Student's t distribution with $\tau = 4$ degrees of freedom is not at all a good fit. The location seems

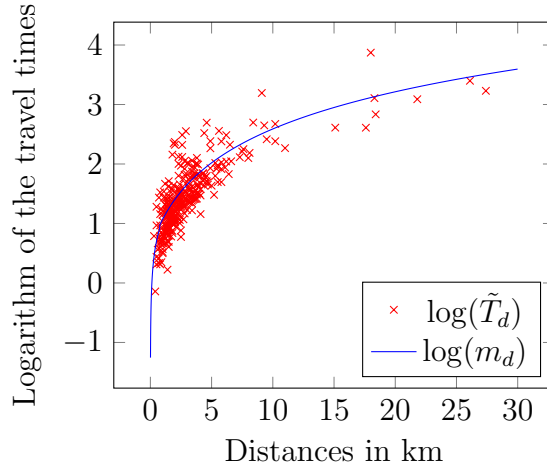


Figure 2.8.: The fit of $\log(m_d)$ to $\log(\tilde{T}_d)$ for the data of November.

to be wrong, so $\log(m_d)$ is not a good estimator for $\log(\tilde{T}_d)$. The Kolmogorov-Smirnov test confirms this. It gives a D -statistic of 0.26168 and a p -value which is smaller than $2.2e-16$. We conclude with the R^2 measure defined in (2.2). For the training sample we

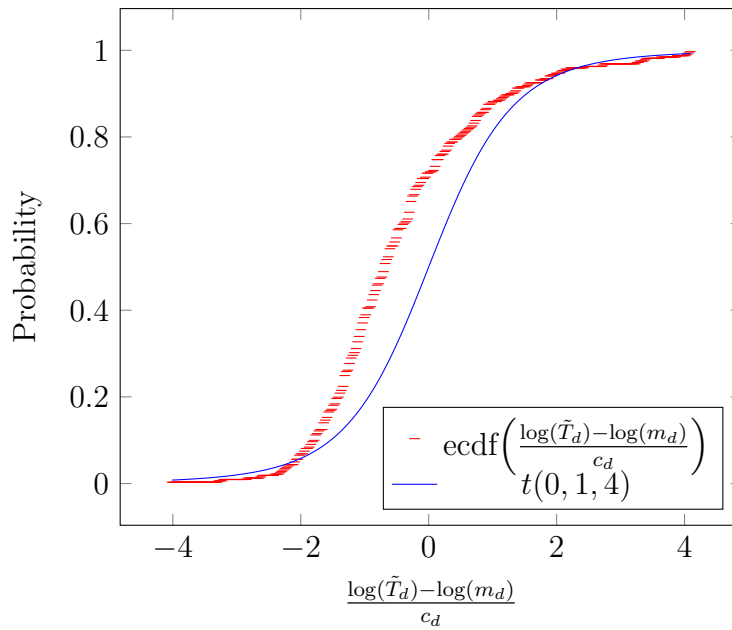


Figure 2.9.: Empirical distribution function of $\frac{\log(\tilde{T}_d) - \log(m_d)}{c_d}$ with the fitted centered t distribution with $\tau = 4$ degrees of freedom for the data of November.

take the data from January, and for the hold-out sample we take the data of November. For our data and choice of parameters we have that $R^2 = 0.3708937$. Even though we do not think that the parametric model works well here, we find an R^2 value similar

to the R^2 value of Budge et al. (2010) ($R^2 = 37.09\%$ compared to $R^2 = 35.53\%$). So, maybe this is actually quite good, considering that we work with actual data and not only with a theoretic model.

Summarizing, in this chapter we filtered all unnecessary information out of our GPS-data. We described a model on the distribution of travel times by Budge et al. (2010) in detail and applied it to our GPS-data. We found that it is likely that our travel times follow a Student's t distribution with $\tau = 4$ degrees of freedom, where the location and scale parameters are based on the distance.

3. Dispatching: Non-concurrent incidents

In Chapter 1, we said that one of the possible improvements for the fire department, when having insight in the behavior of the travel times, is optimal dispatching of water tenders. In this chapter, we will look at decision making concerning the dispatching of water tenders. We consider non-concurrent incidents, meaning that only one incident happens at a time, and that we always have all our water tenders to our disposal, which we consider to be equal to two or three. In Chapter 4, we combine the results of this chapter with the results of Chapter 2. So, we start with a basic model in this chapter, and we amplify it in the next chapter, where we consider concurrent incidents.

We let T_i , $i = 1, 2, 3$, be the distribution of the travel time of water tender i . We send two fire engines to the incident, such that the minimum of the travel times, $T_{ij}^* = \min\{T_i, T_j\}$, is minimized. We look at different situations, such as water tenders having a possibility of sharing the same routes, or, water tenders sharing a part of their route. We model this by considering different distributions of T_i and T_j . Note that the case that T_i and T_j have a deterministic distribution is trivial, since then it is always directly clear which water tender will arrive first at the incident.

3.1. T_i and T_j exponential and independent

We start with the situation where T_i and T_j are independent and exponentially distributed. The results shown here are well-known, but we state them nevertheless, since they will be used further in this chapter.

Theorem 3.1. *Let $T_i \sim \text{Exp}(\lambda_i)$ and $T_j \sim \text{Exp}(\lambda_j)$, $\lambda_i, \lambda_j > 0$, be independent. Then*

$$T_{ij}^* \sim \text{Exp}(\lambda_i + \lambda_j).$$

Proof. Since $T_i \sim \text{Exp}(\lambda_i)$ we have that

$$\mathbb{P}(T_i > t) = 1 - F_{T_i}(t) = \exp(-t\lambda_i).$$

So we have

$$\begin{aligned}
\mathbb{P}(T_{ij}^* > t) &= \mathbb{P}(\min\{T_i, T_j\} > t) \\
&= \mathbb{P}(T_i > t, T_j > t) \\
&= \mathbb{P}(T_i > t)\mathbb{P}(T_j > t) \\
&= \exp(-t\lambda_i)\exp(-t\lambda_j) \\
&= \exp(-t(\lambda_i + \lambda_j)).
\end{aligned}$$

Hence, $T_{ij}^* \sim \text{Exp}(\lambda_i + \lambda_j)$. □

What we can conclude from this, is that when our travel times are exponentially distributed, it is always a good idea to send two cars instead of one, since the expectation of the minimum of two is always smaller than the expectation of the minimum of one: $\mathbb{E}[T_{ij}^*] = \frac{1}{\lambda_i + \lambda_j} < \frac{1}{\lambda_i} = \mathbb{E}[T_i]$. This is a general result, which we will prove in Subsection 3.3.1. So, the current policy of the BWAA to always send two water tenders instead of one is a good policy. The question arises which two water tenders we should send. We will look into this in the next sections, where we consider hyperexponential and hypoexponential travel times.

3.2. T_i and T_j hyperexponential

Now, we consider T_i and T_j hyperexponentially distributed with two phases. A property of the hyperexponential distribution is that it has a decreasing failure rate. This means, in the case of travel times, that when a water tender is on its way to an incident and the ride already takes a long time, it will be even a lot longer with high probability. The idea behind hyperexponential travel times is the following: there is more than one way to travel from one point to another, so you have multiple routes to choose from. Say, there are two different routes to go from the fire station to the fire, and that the travel times of both different routes are exponentially distributed with different parameters. We choose the first route with probability α and the second route with probability $1 - \alpha$. Then

$$T_i = \begin{cases} X_i & \text{with probability } \alpha \\ Y_i & \text{with probability } 1 - \alpha, \end{cases}$$

where X_i and Y_i are independent and both exponentially distributed with parameters λ_i and μ_i respectively. We denote this as $T_i \sim H_2(\alpha, 1 - \alpha; \lambda_i, \mu_i)$. The tail probability of T_i is

$$\mathbb{P}(T_i > t) = \alpha \exp(-\lambda_i t) + (1 - \alpha) \exp(-\mu_i t).$$

We want to obtain the optimal decision when we have two water tender at one fire station and one water tender at another fire station, see Figure 3.1. We will first look

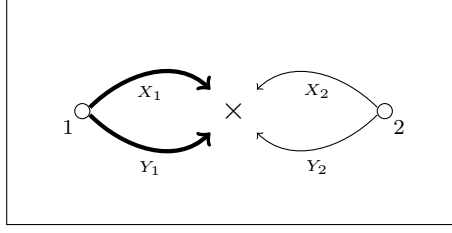


Figure 3.1.: Possibility of a shared route. \circ_1 and \circ_2 are the fire stations, \times is the location of the fire. X_1, X_2, Y_1, Y_2 are the distributions of travel times. Water tenders A and B are located at \circ_1 and water tender C is located at \circ_2 . Water tenders A and B could drive the same route.

at the situation where we have two water tenders at different fire stations, and then, we will look at the situation where we have two water tenders at the same fire station. We will compare these two situations to obtain the optimal decision.

Let us now consider two different water tenders stationed at two different fire stations, such that these water tenders do not share any routes. Then we come to the following theorem:

Theorem 3.2. *Let*

$$\begin{aligned} T_i &\sim H_2(\alpha, 1 - \alpha; \lambda_i, \mu_i) \\ T_j &\sim H_2(\beta, 1 - \beta; \lambda_j, \mu_j), \end{aligned}$$

be independent. Then the minimum T_{ij}^ of T_i and T_j is hyperexponentially distributed with four phases:*

$$T_{ij}^* \sim H_4(\alpha\beta, \alpha(1 - \beta), (1 - \alpha)\beta, (1 - \alpha)(1 - \beta); \lambda_i + \lambda_j, \lambda_i + \mu_j, \mu_i + \lambda_j, \mu_i + \mu_j).$$

Proof. The minimum of T_i and T_j can be obtained as follows:

$$\begin{aligned} \mathbb{P}(T_{ij}^* > t) &= \mathbb{P}(T_i > t, T_j > t) \\ &= \mathbb{P}(T_i > t)\mathbb{P}(T_j > t) \\ &= (\alpha \exp(-\lambda_i t) + (1 - \alpha) \exp(-\mu_i t)) \\ &\quad \times (\beta \exp(-\lambda_j t) + (1 - \beta) \exp(-\mu_j t)) \\ &= \alpha\beta \exp(-(\lambda_i + \lambda_j)t) + \alpha(1 - \beta) \exp(-(\lambda_i + \mu_j)t) \\ &\quad + \beta(1 - \alpha) \exp(-(\mu_i + \lambda_j)t) + (1 - \alpha)(1 - \beta) \exp(-(\mu_i + \mu_j)t). \end{aligned}$$

Hence, T_{ij}^* is hyperexponentially distributed with four phases:

$$T_{ij}^* \sim H_4(\alpha\beta, \alpha(1 - \beta), (1 - \alpha)\beta, (1 - \alpha)(1 - \beta); \lambda_i + \lambda_j, \lambda_i + \mu_j, \mu_i + \lambda_j, \mu_i + \mu_j).$$

□

The expectation of T_{ij}^* is

$$\begin{aligned}\mathbb{E}[T_{ij}^*] &= \int_0^\infty \mathbb{P}(T_{ij}^* > t) dt \\ &= \frac{\alpha\beta}{\lambda_i + \lambda_j} + \frac{\alpha(1-\beta)}{\lambda_i + \mu_j} + \frac{\beta(1-\alpha)}{\mu_i + \lambda_j} + \frac{(1-\alpha)(1-\beta)}{\mu_i + \mu_j}.\end{aligned}$$

Now, we will look at a different situation. Say, we have two water tenders at the same station. The travel times are still hyperexponentially distributed with two phases. Since they are stationed at the same fire station, they share both routes. The probabilities with which each route is chosen, can differ. This leads to the following theorem:

Theorem 3.3. *Let*

$$\begin{aligned}X &\sim \text{Exp}(\lambda) \\ Y &\sim \text{Exp}(\mu),\end{aligned}$$

be independent and define

$$\begin{aligned}T_i &= \begin{cases} X & \text{with probability } \alpha \\ Y & \text{with probability } 1 - \alpha, \end{cases} \\ T_j &= \begin{cases} X & \text{with probability } \beta \\ Y & \text{with probability } 1 - \beta, \end{cases},\end{aligned}$$

for $0 \leq \alpha, \beta \leq 1$. Then the minimum T_{ij}^ of T_i and T_j is hyperexponentially distributed with three phases:*

$$T_{ij}^* \sim H_3\left(\alpha\beta, \alpha + 2\alpha\beta + \beta, (1-\alpha)(1-\beta); \lambda, \lambda + \mu, \mu\right).$$

Proof. For the minimum T_{ij}^* of T_i and T_j we have

$$\begin{aligned}T_{ij}^* &= \min\{T_i, T_j\} \\ &= \begin{cases} X & \text{with probability } \alpha\beta \\ \min\{X, Y\} & \text{with probability } \alpha(1-\beta) + (1-\alpha)\beta \\ Y & \text{with probability } (1-\alpha)(1-\beta). \end{cases}\end{aligned}$$

From Section 3.1 we know that $\min\{X, Y\} \sim \text{Exp}(\lambda + \mu)$. Hence,

$$T_{ij}^* \sim H_3\left(\alpha\beta, \alpha + 2\alpha\beta + \beta, (1-\alpha)(1-\beta); \lambda, \lambda + \mu, \mu\right).$$

□

We consider a situation where we have three water tenders. Two of these water tenders are stationed at the same fire station, the last one is stationed at a different fire station. All three water tenders have hyperexponentially distributed travel times with two phases. We want to know which two of these three water tenders should be sent to an incident such that the expected travel is minimized. This leads to the following theorem:

Theorem 3.4. *Let*

$$\begin{aligned} T_A &= \begin{cases} X_1 & \text{with probability } \alpha \\ Y_1 & \text{with probability } 1 - \alpha, \end{cases} \\ T_B &= \begin{cases} X_1 & \text{with probability } \beta \\ Y_1 & \text{with probability } 1 - \beta, \end{cases} \\ T_C &= \begin{cases} X_2 & \text{with probability } \beta \\ Y_2 & \text{with probability } 1 - \beta, \end{cases}, \end{aligned}$$

where

$$\begin{aligned} X_1 &\sim \text{Exp}(\lambda) \\ Y_1 &\sim \text{Exp}(\mu) \\ X_2 &\sim \text{Exp}(\lambda) \\ Y_2 &\sim \text{Exp}(\mu). \end{aligned}$$

Then

$$\mathbb{E}[T_{AB}] \geq \mathbb{E}[T_{AC}].$$

Proof. From Theorem 3.3 we know that

$$T_{AB}^* \sim H_3\left(\alpha\beta, \alpha + 2\alpha\beta + \beta, (1 - \alpha)(1 - \beta); \lambda, \lambda + \mu, \mu\right).$$

So, the tail probability of T_{AB}^* is given by

$$\begin{aligned} \mathbb{P}(T_{AB}^* > t) &= \alpha\beta \exp(-\lambda t) + (\alpha + 2\alpha\beta + \beta) \exp(-(\lambda + \mu)t) \\ &\quad + (1 - \alpha)(1 - \beta) \exp(-\mu t). \end{aligned}$$

Thus, the expectation T_{AB}^* is given by

$$\mathbb{E}[T_{AB}^*] = \frac{\alpha\beta}{\lambda} + \frac{\alpha + 2\alpha\beta + \beta}{\lambda + \mu} + \frac{(1 - \alpha)(1 - \beta)}{\mu}.$$

From Theorem 3.2 we know that

$$\begin{aligned} T_{AC}^* &\sim H_4(\alpha\beta, \alpha(1 - \beta), (1 - \alpha)\beta, (1 - \alpha)(1 - \beta); 2\lambda, \lambda + \mu, \mu + \lambda, 2\mu) \\ &= H_3(\alpha\beta, \alpha + 2\alpha\beta + \beta, (1 - \alpha)(1 - \beta); 2\lambda, \lambda + \mu, 2\mu). \end{aligned}$$

So, the tail probability of T_{AC}^* is given by

$$\begin{aligned}\mathbb{P}(T_{AC}^* > t) &= \alpha\beta \exp(-2\lambda t) + (\alpha + 2\alpha\beta + \beta) \exp(-(\lambda + \mu)t) \\ &\quad + (1 - \alpha)(1 - \beta) \exp(-2\mu t).\end{aligned}$$

Thus, the expectation T_{AC}^* is given by

$$\mathbb{E}[T_{AC}^*] = \frac{\alpha\beta}{2\lambda} + \frac{\alpha+2\alpha\beta+\beta}{\lambda+\mu} + \frac{(1-\alpha)(1-\beta)}{2\mu}.$$

Hence, the difference between $\mathbb{E}[T_{AB}^*]$ and $\mathbb{E}[T_{AC}^*]$ is

$$\begin{aligned}\mathbb{E}[T_{AB}^*] - \mathbb{E}[T_{AC}^*] &= \frac{\alpha\beta}{\lambda} + \frac{\alpha+2\alpha\beta+\beta}{\lambda+\mu} + \frac{(1-\alpha)(1-\beta)}{\mu} - \frac{\alpha\beta}{2\lambda} - \frac{\alpha+2\alpha\beta+\beta}{\lambda+\mu} - \frac{(1-\alpha)(1-\beta)}{2\mu} \\ &= \frac{\alpha\beta}{2\lambda} + \frac{(1-\alpha)(1-\beta)}{2\mu} \\ &> 0.\end{aligned}$$

Hence, $\mathbb{E}[T_{AB}] \geq \mathbb{E}[T_{AC}]$. □

The theorem states that when the distribution of T_B and T_C is the same, we should send A and C . This seems very obvious. We will now look at an example where T_B and T_C do not have the same expectation, such that $\mathbb{E}[T_B] < \mathbb{E}[T_C]$, but that the expected minimum travel time of water tenders A and C is smaller than the expected minimum travel time of water tenders A and B .

Example 3.5. Let

$$\begin{aligned}T_A &= \begin{cases} X_1 & \text{with probability } \frac{1}{2} \\ Y_1 & \text{with probability } \frac{1}{2}, \end{cases} \\ T_B &= \begin{cases} X_1 & \text{with probability } \frac{1}{3} \\ Y_1 & \text{with probability } \frac{2}{3}, \end{cases} \\ T_C &= \begin{cases} X_2 & \text{with probability } \frac{3}{5} \\ Y_2 & \text{with probability } \frac{2}{5}, \end{cases},\end{aligned}$$

where

$$X_1 \sim \text{Exp}(4)$$

$$Y_1 \sim \text{Exp}(3)$$

$$X_2 \sim \text{Exp}(3)$$

$$Y_2 \sim \text{Exp}(2).$$

Then

$$\mathbb{E}[T_A] = \frac{1}{2} \cdot \frac{1}{4} + \frac{1}{2} \cdot \frac{1}{3} = \frac{7}{24}$$

$$\mathbb{E}[T_B] = \frac{1}{3} \cdot \frac{1}{4} + \frac{2}{3} \cdot \frac{1}{3} = \frac{5}{12}$$

$$\mathbb{E}[T_C] = \frac{3}{5} \cdot \frac{1}{3} + \frac{2}{5} \cdot \frac{1}{2} = \frac{2}{5}.$$

Hence, $\mathbb{E}[T_A] < \mathbb{E}[T_B] < \mathbb{E}[T_C]$. According to Theorem 3.3 we have

$$\mathbb{E}[T_{AB}^*] = \frac{1}{2} \cdot \frac{1}{3} \cdot \frac{1}{4} + \left(\frac{1}{2} \cdot \frac{2}{3} + \frac{1}{2} \cdot \frac{1}{3}\right) + \frac{1}{2} \cdot \frac{2}{3} \cdot \frac{1}{3} = \frac{113}{504}.$$

And according to Theorem 3.2 we have

$$\mathbb{E}[T_{AC}^*] = \frac{1}{2} \cdot \frac{3}{5} \cdot \frac{1}{7} + \frac{1}{2} \cdot \frac{2}{5} \cdot \frac{1}{6} + \frac{1}{2} \cdot \frac{3}{5} \cdot \frac{1}{6} + \frac{1}{2} \cdot \frac{2}{5} \cdot \frac{1}{5} = \frac{349}{2100}.$$

Hence,

$$\mathbb{E}[T_{AB}^*] - \mathbb{E}[T_{AC}^*] = \frac{113}{504} - \frac{349}{2100} = \frac{731}{12600} > 0.$$

Thus, although $\mathbb{E}[T_B] < \mathbb{E}[T_C]$, $\mathbb{E}[T_{AB}^*] > \mathbb{E}[T_{AC}^*]$.

One could argue how realistic hyperexponentially distributed travel times are. First, if one sends two water tenders from the same base and one can choose which route they take, one would let them take different routes. So we need the drivers of the water tenders to choose their routes independently. Second, we consider the travel time per route to be exponentially distributed, which does not agree with our results from Chapter 2. Despite this, we continue to look at travel time distributions based on the exponential distribution, since the exponential distribution is insightful and gives clearer results than other, more realistic distributions.

3.3. T_i and T_j hypoexponential

In this section we look at travel times which are hypoexponentially distributed. The idea behind this is that we have different routes from a fire station to an incident, and that these routes intersect with each other. At every intersection, the driver of the water tender can choose which route he will follow. Figure 3.2 shows a graph of this situation. This seems not realistic, but later in this section we give two examples of hypoexponential distributions, the Erlang distribution and a sum of different exponential distributions, which have a better connection to reality. The hypoexponential distribution has an increasing failure rate. Thus, when a water tender is on its way to an incident and the ride already takes a long time, it will be with high probability that the water tender will soon arrive at the location of the incident. We start with some results for a general hypoexponential distribution, where the exponential components are multiplied by a factor α and $1 - \alpha$, $0 \leq \alpha \leq 1$, respectively.

Theorem 3.6. *Let*

$$X \sim \text{Exp}(\lambda)$$

$$Y \sim \text{Exp}(\mu),$$

be independent and define

$$T_i = \alpha_i X + (1 - \alpha_i) Y$$

$$T_j = \alpha_j X + (1 - \alpha_j) Y,$$

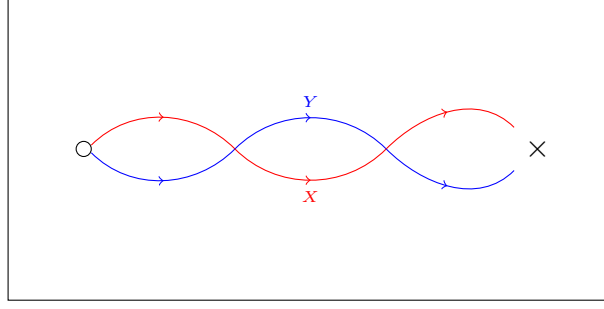


Figure 3.2.: Possibility of a partly shared route. \circ is the location of the fire station and \times is the location of the fire. X and Y are the distributions of the travel times of the red and blue route respectively. At each intersection of the red and the blue route, a water tender can choose which route he takes from that point.

for $0 \leq \alpha_i, \alpha_j \leq 1$. Then T_i and T_j are hyperexponentially distributed and dependent and

$$T_{ij}^* \sim \text{Hypo}_3 \left(\frac{\mu\alpha_j}{\mu\alpha_j - \lambda(1-\alpha_j)}, \frac{\lambda(1-\alpha_i)}{\lambda(1-\alpha_i) - \mu\alpha_i}, \frac{\lambda\mu(\alpha_i - \alpha_j)}{(\mu\alpha_i - \lambda(1-\alpha_i))(\lambda(1-\alpha_j) - \mu\alpha_j)}; \frac{\lambda}{\alpha_j}, \frac{\mu}{1-\alpha_i}, \lambda + \mu \right).$$

Proof. The minimum of T_i and T_j can be obtained as follows:

$$\begin{aligned} T_{ij}^* &= \min\{T_i, T_j\} \\ &= \min\{\alpha_i X + (1 - \alpha_i)Y, \alpha_j X + (1 - \alpha_j)Y\} \\ &= (1 - \alpha_i)Y + \min\{\alpha_i X, \alpha_j X + (1 - \alpha_j - 1 + \alpha_i)Y\} \\ &= \alpha_j X + (1 - \alpha_i)Y + \min\{(\alpha_i - \alpha_j)X, (\alpha_i - \alpha_j)Y\} \\ &= \alpha_j X + (1 - \alpha_i)Y + (\alpha_i - \alpha_j) \min\{X, Y\}. \end{aligned}$$

If we look at the tail probability of T_{ij}^* , we get, by the law of total probability,

$$\begin{aligned} \mathbb{P}(T_{ij}^* > t) &= \mathbb{P}(\alpha_j X + (1 - \alpha_i)Y + (\alpha_i - \alpha_j) \min\{X, Y\} > t) \\ &= \mathbb{P}(\alpha_j X + (1 - \alpha_i)Y + (\alpha_i - \alpha_j)X > t | X \leq Y) \mathbb{P}(X \leq Y) \\ &\quad + \mathbb{P}(\alpha_j X + (1 - \alpha_i)Y + (\alpha_i - \alpha_j)Y > t | X > Y) \mathbb{P}(X > Y). \end{aligned}$$

We will work out these products of probabilities separately. By the definition of conditional probability we get,

$$\begin{aligned} \mathbb{P}(\alpha_j X + (1 - \alpha_i)Y + (\alpha_i - \alpha_j)X > t | X \leq Y) \mathbb{P}(X \leq Y) \\ = \mathbb{P}(\alpha_j X + (1 - \alpha_i)Y + (\alpha_i - \alpha_j)X > t, X \leq Y). \end{aligned}$$

If we now again apply the law of total probability we get

$$\begin{aligned}
& \mathbb{P}(\alpha_j X + (1 - \alpha_i)Y + (\alpha_i - \alpha_j)X > t, X \leq Y) \\
&= \mathbb{P}(\alpha_i X + (1 - \alpha_i)Y > t, X \leq Y) \\
&= \mathbb{P}(\alpha_i X + (1 - \alpha_i)Y > t, X \leq Y | Y > t) \mathbb{P}(Y > t) \\
&\quad + \mathbb{P}(\alpha_i X + (1 - \alpha_i)Y > t, X \leq Y | Y \leq t) \mathbb{P}(Y \leq t) \\
&= \mathbb{P}(\alpha_i X + (1 - \alpha_i)Y > t, X \leq Y | Y > t) \mathbb{P}(Y > t) + 0 \\
&= \mathbb{P}(\alpha_i X + (1 - \alpha_i)Y > t, X \leq Y, Y > t),
\end{aligned}$$

where we use that

$$\begin{aligned}
\mathbb{P}(\alpha_i X + (1 - \alpha_i)Y > t, X \leq Y | Y \leq t) &\leq \mathbb{P}(\alpha_i Y + (1 - \alpha_i)Y > t | Y \leq t) \\
&= \mathbb{P}(Y > t | Y \leq t) \\
&= 0.
\end{aligned}$$

We will now integrate over the densities of X and Y , so we can express the probability in terms of λ , μ , α_i and α_j .

$$\begin{aligned}
& \mathbb{P}(\alpha_i X + (1 - \alpha_i)Y > t, X \leq Y, Y > t) \\
&= \int_0^\infty f_X(x) \mathbb{P}(Y > \frac{t - \alpha_i x}{1 - \alpha_i}, Y \geq x, Y > t) dx.
\end{aligned}$$

Note that when $x > t$ we have that $\frac{t - \alpha_i x}{1 - \alpha_i} < x, t$ and that when $x \leq t$ we have that $\frac{t - \alpha_i x}{1 - \alpha_i} \geq x, t$. So we end up with

$$\begin{aligned}
& \int_0^\infty f_X(x) \mathbb{P}(Y > \frac{t - \alpha_i x}{1 - \alpha_i}, Y \geq x, Y > t) dx \\
&= \int_0^t f_X(x) \int_{\frac{t - \alpha_i x}{1 - \alpha_i}}^\infty f_Y(y) dy dx + \int_t^\infty f_X(x) \int_x^\infty f_Y(y) dy dx \\
&= \int_0^t \lambda \exp(-\lambda x) \exp(-\mu \frac{t - \alpha_i x}{1 - \alpha_i}) dx + \int_t^\infty \lambda \exp(-\lambda x) \exp(-\mu x) dx \\
&= \exp(\frac{-\mu t}{1 - \alpha_i}) \lambda \int_0^t \exp(-\frac{x(\lambda(1 - \alpha_i) - \mu \alpha_i)}{1 - \alpha_i}) dx + \frac{\lambda}{\lambda + \mu} \exp(-t(\lambda + \mu)) \\
&= \frac{\lambda(1 - \alpha_i)}{\lambda(1 - \alpha_i) - \mu \alpha_i} \left(\exp(\frac{-\mu t}{1 - \alpha_i}) - \exp(-t(\lambda + \mu)) \right) + \frac{\lambda}{\lambda + \mu} \exp(-t(\lambda + \mu)) \\
&= \frac{\lambda(1 - \alpha_i)}{\lambda(1 - \alpha_i) - \mu \alpha_i} \exp(\frac{-\mu t}{1 - \alpha_i}) + \frac{\lambda \mu}{(\lambda + \mu)(\mu \alpha_i - \lambda(1 - \alpha_i))} \exp(-t(\lambda + \mu)).
\end{aligned}$$

We can evaluate $\mathbb{P}(\alpha_j X + (1 - \alpha_i)Y + (\alpha_i - \alpha_j)Y > t | X > Y) \mathbb{P}(X > Y)$ in a similar way, which gives as a result:

$$\begin{aligned}
& \mathbb{P}(\alpha_j X + (1 - \alpha_i)Y + (\alpha_i - \alpha_j)Y > t | X > Y) \mathbb{P}(X > Y) \\
&= \frac{\mu \alpha_j}{\mu \alpha_j - \lambda(1 - \alpha_j)} \exp(\frac{-\lambda t}{\alpha_j}) + \frac{\lambda \mu}{(\lambda + \mu)(\lambda(1 - \alpha_j) - \mu \alpha_j)} \exp(-t(\lambda + \mu)).
\end{aligned}$$

Hence, our final result is

$$\begin{aligned}
\mathbb{P}(T_{ij}^* > t) &= \frac{\lambda(1-\alpha_i)}{\lambda(1-\alpha_i)-\mu\alpha_i} \exp\left(\frac{-\mu t}{1-\alpha_i}\right) + \frac{\lambda\mu}{(\lambda+\mu)(\mu\alpha_i-\lambda(1-\alpha_i))} \exp(-t(\lambda+\mu)) \\
&\quad + \frac{\mu\alpha_j}{\mu\alpha_j-\lambda(1-\alpha_j)} \exp\left(\frac{-\lambda t}{\alpha_j}\right) + \frac{\lambda\mu}{(\lambda+\mu)(\lambda(1-\alpha_j)-\mu\alpha_j)} \exp(-t(\lambda+\mu)) \\
&= \frac{\mu\alpha_j}{\mu\alpha_j-\lambda(1-\alpha_j)} \exp\left(\frac{-\lambda t}{\alpha_j}\right) + \frac{\lambda(1-\alpha_i)}{\lambda(1-\alpha_i)-\mu\alpha_i} \exp\left(\frac{-\mu t}{1-\alpha_i}\right) \\
&\quad + \frac{\lambda\mu(\alpha_i-\alpha_j)}{(\mu\alpha_i-\lambda(1-\alpha_i))(\lambda(1-\alpha_j)-\mu\alpha_j)} \exp(-t(\lambda+\mu)).
\end{aligned}$$

We conclude that

$$T_{ij}^* \sim Hypo_3\left(\frac{\mu\alpha_j}{\mu\alpha_j-\lambda(1-\alpha_j)}, \frac{\lambda(1-\alpha_i)}{\lambda(1-\alpha_i)-\mu\alpha_i}, \frac{\lambda\mu(\alpha_i-\alpha_j)}{(\mu\alpha_i-\lambda(1-\alpha_i))(\lambda(1-\alpha_j)-\mu\alpha_j)}; \frac{\lambda}{\alpha_j}, \frac{\mu}{1-\alpha_i}, \lambda+\mu\right).$$

□

The expression for T_{ij}^* is not very intuitive. Therefore, we look at the situations where α_i and α_j are either equal to 1 or 0, since for these situations we know the results at forehand. When $\alpha_i = 1$ and $\alpha_j = 1$, we have that

$$\mathbb{P}(T_{ij}^* > t) = \mathbb{P}(\min\{X, X\} > t) = \exp(-t\lambda).$$

When $\alpha_i = 0$ and $\alpha_j = 0$, we have that

$$\mathbb{P}(T_{ij}^* > t) = \mathbb{P}(\min\{Y, Y\} > t) = \exp(-t\mu).$$

And, finally, when $\alpha_i = 1$ and $\alpha_j = 0$, we have that

$$\mathbb{P}(T_{ij}^* > t) = \mathbb{P}(\min\{X, Y\} > t) = \exp(-t(\lambda+\mu)).$$

The expectation of T_{ij}^* is

$$\begin{aligned}
\mathbb{E}[T_{ij}^*] &= \int_0^\infty \mathbb{P}(T_{ij}^* > t) dt \\
&= \frac{\mu\alpha_j}{\mu\alpha_j-\lambda(1-\alpha_j)} \frac{\alpha_j}{\lambda} + \frac{\lambda(1-\alpha_i)}{\lambda(1-\alpha_i)-\mu\alpha_i} \frac{1-\alpha_i}{\mu} + \frac{\lambda\mu(\alpha_i-\alpha_j)}{(\mu\alpha_i-\lambda(1-\alpha_i))(\lambda(1-\alpha_j)-\mu\alpha_j)} \frac{1}{\lambda+\mu} \\
&= \frac{\alpha_j}{\lambda} + \frac{1-\alpha_i}{\mu} + \frac{\alpha_i-\alpha_j}{\lambda+\mu}.
\end{aligned}$$

We will now compare the dependent to the independent situation. This gives us the following result:

Theorem 3.7. *Let*

$$\begin{aligned}
T_A &= \alpha X_1 + (1-\alpha)Y_1 \\
T_B &= \beta X_1 + (1-\beta)Y_1 \\
T_C &= \beta X_2 + (1-\beta)Y_2,
\end{aligned}$$

where

$$\begin{aligned} X_1 &\sim \text{Exp}(\lambda) \\ X_2 &\sim \text{Exp}(\lambda) \\ Y_1 &\sim \text{Exp}(\mu) \\ Y_2 &\sim \text{Exp}(\mu), \end{aligned}$$

and $0 < \alpha, \beta < 1$. Then

$$\mathbb{E}[T_{AB}^*] > \mathbb{E}[T_{AC}^*].$$

Proof. The expectations of the minimum of T_A and T_B , and T_A and T_C are, respectively,

$$\begin{aligned} \mathbb{E}[T_{AB}^*] &= \frac{\beta}{\lambda} + \frac{1-\alpha}{\mu} + \frac{\alpha-\beta}{\lambda+\mu}, \\ \mathbb{E}[T_{AC}^*] &= \frac{\alpha\beta}{2\lambda} + \frac{(1-\alpha)(1-\beta)}{2\mu} + \frac{\alpha-2\alpha\beta+\beta}{\lambda+\mu}. \end{aligned}$$

Thus, we have

$$\begin{aligned} \mathbb{E}[T_{AB}^*] - \mathbb{E}[T_{AC}^*] &= \frac{\beta(2-\alpha)}{2\lambda} + \frac{(1-\alpha)(1+\beta)}{2\mu} + \frac{2\beta(\alpha-1)}{\lambda+\mu} \\ &= \frac{\beta(2-\alpha)\mu(\lambda+\mu)}{2\lambda\mu(\lambda+\mu)} + \frac{(1-\alpha)(1+\beta)\lambda(\lambda+\mu)}{2\lambda\mu(\lambda+\mu)} + \frac{2\beta(\alpha-1)2\lambda\mu}{2\lambda\mu(\lambda+\mu)} \\ &= \frac{\lambda^2(1-\alpha+\beta-\alpha\beta)}{2\lambda\mu(\lambda+\mu)} + \frac{\lambda\mu(1-\alpha-\beta+2\alpha\beta)}{2\lambda\mu(\lambda+\mu)} + \frac{\mu^2\beta(2-\alpha)}{2\lambda\mu(\lambda+\mu)} \\ &= \frac{\lambda^2(1-\alpha)(1+\beta)}{2\lambda\mu(\lambda+\mu)} + \frac{\lambda\mu(1-\alpha)(1-\beta)}{2\lambda\mu(\lambda+\mu)} + \frac{\lambda\mu\alpha\beta}{2\lambda\mu(\lambda+\mu)} + \frac{\mu^2\beta(2-\alpha)}{2\lambda\mu(\lambda+\mu)} \\ &> 0, \end{aligned}$$

since $0 < \alpha, \beta < 1$. Hence, $\mathbb{E}[T_{AB}^*] > \mathbb{E}[T_{AC}^*]$. \square

The theorem states that when the expectation of T_B and T_C is the same, we should send A and C . This seems obvious. We will now look at an example where T_B and T_C do not have the same expectation, such that $\mathbb{E}[T_B] < \mathbb{E}[T_C]$, but it is still better to send water tenders A and C instead of A and B . Note that this is not always true, but that there are cases in which it is true.

Example 3.8. We consider three water tenders, A , B and C , whose travel times are hyperexponentially distributed. Water tenders A and B are stationed at the same fire station. For a particular fire, the distribution of the travel times is as follows:

$$\begin{aligned} T_A &\sim \frac{1}{2}X_1 + \frac{1}{2}Y_1 \\ T_B &\sim \frac{1}{3}X_1 + \frac{2}{3}Y_1 \\ T_C &\sim \frac{1}{5}X_2 + \frac{4}{5}Y_2, \end{aligned}$$

where

$$\begin{aligned} X_1 &\sim \text{Exp}(4) \\ X_2 &\sim \text{Exp}(2) \\ Y_1 &\sim \text{Exp}(3) \\ Y_2 &\sim \text{Exp}(3). \end{aligned}$$

Then

$$\begin{aligned} \mathbb{E}[T_A] &= \frac{7}{24} \\ \mathbb{E}[T_B] &= \frac{11}{36} \\ \mathbb{E}[T_C] &= \frac{11}{30}. \end{aligned}$$

Hence, $\mathbb{E}[T_A] < \mathbb{E}[T_B] < \mathbb{E}[T_C]$. If we could send only one water tender, we would send water tender A . But, if we could send two water tenders, which would be the second water tender to send. $\mathbb{E}[T_B]$ is smaller than $\mathbb{E}[T_C]$, but A and B are dependent since they share the same routes, where A and C do not. To make a good decision, we will look at $\mathbb{E}[T_{AB}^*]$ and $\mathbb{E}[T_{AC}^*]$:

$$\begin{aligned} \mathbb{E}[T_{AB}^*] &= \frac{1}{3} \frac{1}{4} + \frac{1}{2} \frac{1}{3} + \frac{1}{6} \frac{1}{12} = \frac{19}{72} \\ \mathbb{E}[T_{AC}^*] &= \frac{1}{2} \frac{1}{5} \frac{1}{6} + \frac{1}{2} \frac{4}{5} \frac{1}{7} + \frac{1}{5} \frac{1}{2} \frac{1}{5} + \frac{1}{2} \frac{4}{5} \frac{1}{6} = \frac{337}{2100}. \end{aligned}$$

Hence, $\mathbb{E}[T_{AB}^*] > \mathbb{E}[T_{AC}^*]$, so we should send A and C even though $\mathbb{E}[T_B] < \mathbb{E}[T_C]$.

We will now look at two examples of hypoexponential distributions. First, we look at the Erlang(2) distribution, which can be seen as the hypoexponential(2) distribution which we just illustrated, with $\alpha = \frac{1}{2}$, $X \stackrel{d}{=} Y$ and the sum of X and Y multiplied by two. Second, we look at a sum of different exponential random variables. Here, the factor α is equal to $\frac{1}{2}$ and the sum of X and Y is multiplied by two.

3.3.1. T_i and T_j Erlang

We now consider every route as a combination of different subroutes, for which the travel times are all independent and identically exponentially distributed. So, we have

$$T_i = X_i + Y_i,$$

where X_i and Y_i are both $\text{Exp}(\lambda)$ distributed. Hence, T_i is Erlang(2) distributed.

We will now prove a theorem about the difference between the expected value of the minimum of dependent and independent Erlang distributed travel times.

Theorem 3.9. *Let*

$$\begin{aligned} T_A &= X_1 + Y_1 \\ T_B &= X_1 + Y_1 \\ T_C &= X_2 + Y_2, \end{aligned}$$

where X_1, X_2, Y_1 and Y_2 are all $\text{Exp}(\lambda)$ distributed and independent. Then,

$$\mathbb{E}[T_{AB}^*] > \mathbb{E}[T_{AC}^*].$$

Proof. We start with evaluating T_{AB}^* , so we can calculate $\mathbb{E}[T_{AB}^*]$. After that, we do the same for T_{AC}^* .

$$T_{AB}^* = \min\{T_A, T_B\} = \min\{X_1 + Y_1, X_1 + Y_1\} = X_1 + Y_1 = T_A,$$

and

$$\mathbb{E}[T_{AB}^*] = \mathbb{E}[T_A] = \frac{2}{\lambda}.$$

Furthermore,

$$\begin{aligned} \mathbb{P}(T_{AC}^* > t) &= \mathbb{P}(T_A > t)\mathbb{P}(T_C > t) \\ &= (1 + \lambda t) \exp(-\lambda t)(1 + \lambda t) \exp(-\lambda t) \\ &= (1 + \lambda t)^2 \exp(-2\lambda t), \end{aligned}$$

which gives

$$\begin{aligned} \mathbb{E}[T_{AC}^*] &= \int_0^\infty (1 + \lambda t)^2 \exp(-2\lambda t) dt \\ &= \int_0^\infty \exp(-2\lambda t) dt + 2 \int_0^\infty \lambda t \exp(-2\lambda t) dt + \int_0^\infty (\lambda t)^2 \exp(-2\lambda t) dt \\ &= \frac{1}{2\lambda} + \int_0^\infty \exp(-2\lambda t) dt + \frac{\lambda}{2} \int_0^\infty t \exp(-2\lambda t) dt \\ &= \frac{1}{2\lambda} + \frac{1}{2\lambda} + \frac{1}{4} \int_0^\infty \exp(-2\lambda t) dt \\ &= \frac{1}{\lambda} + \frac{1}{8\lambda} \\ &= \frac{9}{8\lambda}. \end{aligned}$$

When we look at the difference between $\mathbb{E}[T_{AB}^*]$ and $\mathbb{E}[T_{AC}^*]$, we see

$$\mathbb{E}[T_{AB}^*] - \mathbb{E}[T_{AC}^*] = \frac{2}{\lambda} - \frac{9}{8\lambda} = \frac{7}{8\lambda} > 0.$$

Hence, $\mathbb{E}[T_{AB}^*] > \mathbb{E}[T_{AC}^*]$. □

This result looks rather general. Intuitively we would expect that we do not need T_A, T_B and T_C to be Erlang distributed for it to be true that the minimum of two random variables is smaller than one of the two random variables. We will prove this in the next theorem.

Theorem 3.10. *Let T_A, T_B and T_C be general distributions, where $T_A = T_B \stackrel{d}{=} T_C$, and T_C is independent of T_A and T_B . Then*

$$\mathbb{E}[T_{AB}^*] \geq \mathbb{E}[T_{AC}^*].$$

Proof. Since $T_A = T_B$, we have

$$T_{AB}^* = \min\{T_A, T_B\} = T_A.$$

For the minimum of T_A and T_C we have the following:

$$\begin{aligned} \mathbb{P}(T_{AC}^* > t) &= \mathbb{P}(T_A > t)\mathbb{P}(T_C > t) \\ F_{T_{AC}^*}(t) &= 1 - (1 - F_{T_A}(t))(1 - F_{T_C}(t)) \\ &= F_{T_A}(t) + F_{T_C}(t) + F_{T_A}(t)F_{T_C}(t). \end{aligned}$$

We want to prove that $F_{T_{AB}^*}(t) \leq F_{T_{AC}^*}(t)$ for $t \geq 0$, since this implies $\mathbb{E}[T_{AB}^*] \geq \mathbb{E}[T_{AC}^*]$. We will prove this by contradiction. Say $F_{T_{AB}^*}(t) > F_{T_{AC}^*}(t)$. Then we have

$$\begin{aligned} F_{T_{AB}^*}(t) &> F_{T_{AC}^*}(t) \\ F_{T_A} &> F_{T_A}(t) + F_{T_C}(t) + F_{T_A}(t)F_{T_C}(t) \\ 0 &> F_{T_C}(t)(1 - F_{T_A}(t)), \end{aligned}$$

which is a contradiction. Hence, $F_{T_{AB}^*}(t) \leq F_{T_{AC}^*}(t)$, and thus $\mathbb{E}[T_{AB}^*] \geq \mathbb{E}[T_{AC}^*]$. \square

Note that this comparison is similar to the comparison between one general distribution and the minimum of two distributions, as is Section 3.1.

3.3.2. T_i and T_j sums of different exponential random variables

We consider a similar situation as in the previous section, only now the travel times of the subroutes are exponentially distributed with different parameters. Figure 3.3 shows a graph of this situation. We have

$$T_i = X_i + Y_i,$$

where

$$\begin{aligned} X_i &\sim \text{Exp}(\lambda) \\ Y_i &\sim \text{Exp}(\mu), \end{aligned}$$

$\lambda \neq \mu$, and the distribution of T_i is

$$\begin{aligned} f_{T_i}(x) &= \int_0^x f_{X_i}(x-y)f_{Y_i}(y)dy \\ &= \int_0^x \lambda \exp(-\lambda(x-y))\mu \exp(-\mu y)dy \\ &= \lambda\mu \exp(-\lambda x) \int_0^x \exp(-(\mu-\lambda)y)dy \\ &= \frac{\lambda\mu}{\mu-\lambda} (\exp(-\lambda x) - \exp(-\mu x)). \end{aligned}$$

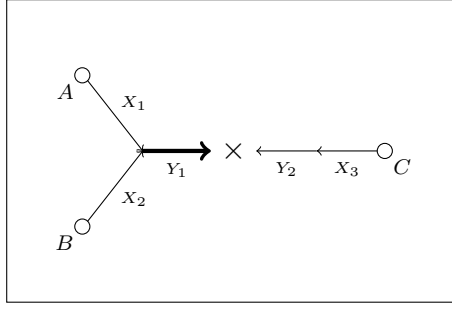


Figure 3.3.: Partly shared route. A, B, C , are the fire stations, \times is the location of the fire. X_1, X_2, X_3, Y_1, Y_2 are the distributions of travel times. Water tenders located at A and B share a part of their route.

The tail probability then becomes

$$\begin{aligned}
 \mathbb{P}(T_i > t) &= \int_t^\infty f_{T_i}(x) dx \\
 &= \int_t^\infty \frac{\lambda\mu}{\mu-\lambda} (\exp(-\lambda x) - \exp(-\mu x)) dx \\
 &= \frac{\mu}{\mu-\lambda} \exp(-\lambda t) + \frac{\lambda}{\lambda-\mu} \exp(-\mu t).
 \end{aligned}$$

Theorem 3.11. *Let*

$$\begin{aligned}
 T_A &= X_1 + Y_1 \\
 T_B &= X_2 + Y_1 \\
 T_C &= X_3 + Y_2,
 \end{aligned}$$

where

$$\begin{aligned}
 X_1 &\sim \text{Exp}(\lambda) \\
 X_2 &\sim \text{Exp}(\lambda) \\
 X_3 &\sim \text{Exp}(\lambda) \\
 Y_1 &\sim \text{Exp}(\mu) \\
 Y_2 &\sim \text{Exp}(\mu).
 \end{aligned}$$

Then

$$\mathbb{E}[T_{AB}^*] > \mathbb{E}[T_{AC}^*].$$

Proof. First, we evaluate T_{AB}^* , so we can compute $\mathbb{E}[T_{AB}^*]$. After that, we compute $\mathbb{P}(T_{AC} > t)$ with which we can compute $\mathbb{E}[T_{AC}^*]$.

$$\begin{aligned} T_{AB}^* &= \min\{T_A, T_B\} \\ &= \min\{X_1 + Y_1, X_2 + Y_1\} \\ &= Y_1 + \min\{X_1, X_2\}, \end{aligned}$$

and hence,

$$\mathbb{E}[T_{AB}^*] = \frac{1}{\mu} + \frac{1}{2\lambda} = \frac{2\lambda + \mu}{2\lambda\mu}.$$

Further

$$\begin{aligned} \mathbb{P}(T_{AC} > t) &= \mathbb{P}(T_A > t)\mathbb{P}(T_C > t) \\ &= \left(\frac{\mu}{\mu - \lambda} \exp(-\lambda t) + \frac{\lambda}{\lambda - \mu} \exp(-\mu t) \right)^2 \\ &= \frac{\mu^2}{(\mu - \lambda)^2} \exp(-2\lambda t) + \frac{2\lambda\mu}{(\mu - \lambda)(\lambda - \mu)} \exp(-(\lambda + \mu)t) + \frac{\lambda^2}{(\lambda - \mu)^2} \exp(-2\mu t), \end{aligned}$$

and hence,

$$\begin{aligned} \mathbb{E}[T_{AC}^*] &= \frac{\mu^2}{2\lambda(\mu - \lambda)^2} + \frac{2\lambda\mu}{(\mu - \lambda)(\lambda - \mu)(\lambda + \mu)} + \frac{\lambda^2}{2\mu(\lambda - \mu)^2} \\ &= \frac{\mu^3(\lambda - \mu)^2(\lambda + \mu) + 4\lambda^2\mu^2(\lambda - \mu)(\lambda + \mu) + \lambda^3(\mu - \lambda)^2(\lambda + \mu)}{2\lambda\mu(\mu - \lambda)^2(\lambda - \mu)^2(\lambda + \mu)} \\ &= \frac{(\lambda - \mu)^4(\lambda^2 + 3\lambda\mu + \mu^2)}{2\lambda\mu(\mu - \lambda)^2(\lambda - \mu)^2(\lambda + \mu)} \\ &= \frac{\lambda^2 + 3\lambda\mu + \mu^2}{2\lambda\mu(\lambda + \mu)}. \end{aligned}$$

The difference then is

$$\begin{aligned} \mathbb{E}[T_{AB}^*] - \mathbb{E}[T_{AC}^*] &= \frac{2\lambda + \mu}{2\lambda\mu} - \frac{\lambda^2 + 3\lambda\mu + \mu^2}{2\lambda\mu(\lambda + \mu)} \\ &= \frac{(2\lambda + \mu)(\lambda + \mu) - (\lambda^2 + 3\lambda\mu + \mu^2)}{2\lambda\mu(\lambda + \mu)} \\ &= \frac{\lambda^2}{2\lambda\mu(\lambda + \mu)} \\ &> 0, \end{aligned}$$

for $\lambda \neq \mu$, $\lambda, \mu > 0$. Hence, $\mathbb{E}[T_{AB}^*] > \mathbb{E}[T_{AC}^*]$. \square

Just as in Theorem 3.7, the result of Theorem 3.11 seems very obvious. Therefore, we will now look at an example where we have that $\mathbb{E}[T_B] < \mathbb{E}[T_C]$, but it is still better to send water tenders A and B instead of A and C .

Example 3.12. We consider three water tenders, A , B and C , where water tenders A and B share a part of their route to the fire. The travel times are distributed as follows:

$$\begin{aligned} T_A &\sim X_1 + Y_1 \\ T_B &\sim X_2 + Y_1 \\ T_C &\sim X_3 + Y_2, \end{aligned}$$

where

$$\begin{aligned} X_1 &\sim \text{Exp}(4) \\ X_2 &\sim \text{Exp}(4) \\ X_3 &\sim \text{Exp}(3) \\ Y_1 &\sim \text{Exp}(2) \\ Y_2 &\sim \text{Exp}(2). \end{aligned}$$

Then

$$\begin{aligned} \mathbb{E}[T_A] &= \frac{3}{4} \\ \mathbb{E}[T_B] &= \frac{3}{4} \\ \mathbb{E}[T_C] &= \frac{5}{6} \end{aligned}$$

and hence, $\mathbb{E}[T_A] = \mathbb{E}[T_B] < \mathbb{E}[T_C]$. If we would send only one water tender, we would send water tender A or B . But if we send two, the choice of sending A and B is not necessarily the better choice, since A and B are dependent. We take a look at the minimum of T_A and T_B and the minimum of T_A and T_C to see which two water tenders we should send. The minimum of T_A and T_B is

$$\mathbb{E}[T_{AB}^*] = \frac{1}{2} + \frac{1}{4+4} = \frac{5}{8}.$$

For the minimum of T_A and T_C we first look at the tail probability of T_{AC}^* :

$$\begin{aligned} \mathbb{P}(T_{AC}^* > t) &= \left(\frac{2}{2-4} \exp(-4t) + \frac{4}{4-2} \exp(-2t) \right) \left(\frac{2}{2-3} \exp(-3t) + \frac{3}{3-2} \exp(-2t) \right) \\ &= \left(-\exp(-4t) + 2 \exp(-2t) \right) \left(-2 \exp(-3t) + 3 \exp(-2t) \right) \\ &= 2 \exp(-7t) - 3 \exp(-6t) - 4 \exp(-5t) + 6 \exp(-4t), \end{aligned}$$

and hence,

$$\mathbb{E}[T_{AC}^*] = \frac{2}{7} - \frac{1}{2} - \frac{4}{5} + \frac{3}{2} = \frac{17}{35}.$$

So

$$\mathbb{E}[T_{AB}^*] = \frac{5}{8} > \frac{17}{35} = \mathbb{E}[T_{AC}^*].$$

Hence, it is better to send water tender A and C instead of A and B , even though the travel time of B is shorter than the travel time of C .

In this section we have seen different types of the hypoexponential distribution. We have seen that the expectation of the minimum of independent random variables is smaller than the expectation of the minimum of dependent random variables. In practice

this means that it is optimal to send the two out of three water tenders who do not share a (part of a) route then two who do share a (part of a) route. So, the current policy of the BWAA, to send two water tenders to each incident, is a good policy. If the BWAA would get more than one water tender per fire station in the future, they should not send two of those water tenders to an incident, instead of two water tenders from different fire stations. Also, it is of importance for the BWAA to take into account the probability of taking the same route, when sending water tenders from two fire stations, which are located closely to each other. Just as in Section 3.2 one could argue how realistic travel times are which are based on exponential travel times, since these do not agree with our results in Chapter 2. Again, just as in Section 3.2 we choose the exponential distribution, since it is insightful and gives clearer results than other, more realistic distributions.

4. Dispatching: Concurrent incidents

In this chapter, we consider multiple incidents happening at the same time. As a consequence, not all water tenders are available, since some of them are at another incident. And, if we send a water tender to a fire, we can not use it for a certain period of time. So, the dispatching decision no longer depends on the current incident only. We will model this situation as a Markov Decision Process. Further, we will look at a specific example to get insight in the behaviour of the optimal policy. We end this chapter with a theorem about the behaviour of the optimal policy.

4.1. Problem description

We consider an infinite horizon, meaning that our time t goes to infinity. In our model there is only one water tender at each fire station, and we let N be the number of fire stations. We believe that being in time at an incident which is happening now, is more important than being in time at an incident happening in the future, so we take in a fixed pre-determined discount factor $\alpha \in (0, 1)$. Every time epoch t , an incident occurs, and we need to decide how many and which water tenders we should send. A water tender which is sent to an incident, stays at the incident for at least one time epoch, but it can be longer. Sending more water tenders means an equal or shorter expected travel time. But, on the other hand, sending more water tenders leads to less water tenders when the next incident occurs, at time epoch $t + 1$. Also, the expected minimum travel time depends on the choice of the water tenders sent. The distribution of the individual travel times, and whether or not the travel times are independent, does matter, see Chapter 3.

4.2. Model

A water tender can only be sent to a fire when it is at the fire station. We let $\{s_t\}$ be the state process, indicating whether or not a water tender is at the fire station at time t . We have $s_t = (s_t(1), \dots, s_t(N))$,

$$s_t(i) = \begin{cases} 1 & \text{if water tender } i \text{ is at fire station } i \text{ at time } t, \\ 0 & \text{otherwise,} \end{cases}$$

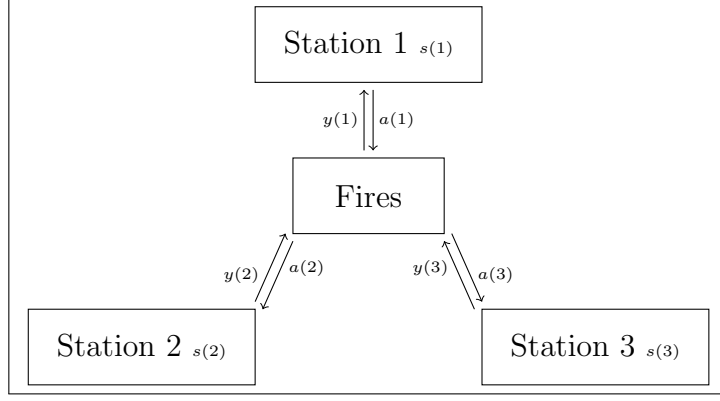


Figure 4.1.: The Markov Decision Process for $N = 3$. The variables $s(i)$, $a(i)$ and $y(i)$, $i = 1, 2, 3$, indicate if the water tender is at the fire station, goes to a fire or returns from a fire, respectively.

for $i \in \{1, \dots, N\}$ and $s_t \in \mathcal{S}_1 \times \dots \times \mathcal{S}_N = \mathcal{S}$, $\mathcal{S} = \{0, 1\}^N$. We consider an action space $\mathcal{A}(s) = \mathcal{A}_1(s) \times \dots \times \mathcal{A}_N(s)$, such that for $a_t = (a_t(1), \dots, a_t(N)) \in \mathcal{A}(s)$ we have

$$a_t(i) = \begin{cases} 1 & \text{if we send water tender } i \text{ to the fire at time } t, \\ 0 & \text{otherwise,} \end{cases}$$

$$\mathcal{A}(s) = \{a_t \in \{0, 1\}^N : a_t(i) \leq s_t(i)\},$$

since we can not send a water tender, when it is not at the fire station. We let $y_t = (y_t(1), \dots, y_t(N))$ indicate whether or not a water tender is back at the station:

$$y_t(i) = \begin{cases} 1 & \text{if } s_{t-1}(i) = 0 \text{ or } a_{t-1}(i) = 1, \\ & \text{and the water tender returns to the fire station,} \\ 0 & \text{otherwise,} \end{cases}$$

and let y_t happen according to a Bernoulli process with parameter p . So, if $s_{t-1}(i) = 0$ or $a_{t-1}(i) = 1$ then $y_t(i) = 1$ with probability p .

We let events happen in the following order. First some or none of the water tenders which were sent away, come back, so y_t happens. Then we measure the state s_t , the number of water tenders we have, and after that we make a decision a_t , how many water tenders we send away. This gives us

$$s_{t+1} = s_t - a_t + y_{t+1}.$$

Figure 4.1 shows a graph of this situation for $N = 3$. We let T_i be the time in minutes it takes water tender i to arrive at an incident, and define

$$T(a_t) = \min\{T_i : a_t(i) = 1\}.$$

We consider certain costs, if it takes a water tender more than M minutes to arrive at a fire:

$$c(a_t) = \mathbb{1}(T(a_t) \geq M),$$

so, sending more water tenders leads to less costs $c(a, t)$. This leads to the total discounted expected costs for time t :

$$V_t(s_t) = \min_{a_t \in \mathcal{A}(s_t)} \{c(a_t) + \alpha \sum_{s' \in \mathcal{S}} p_{t+1}^s(s'|s_t, a_t) V_{t+1}(s')\}, \quad (4.1)$$

where $p_{t+1}^s(s_{t+1}|s_t, a_t)$ is the probability that at time epoch $t + 1$, the state is s_{t+1} , given that the state variable and the decision variable at epoch t , are s_t and a_t respectively. So, the costs we make in state s_t is the sum of the direct costs, $c(a_t)$, according to our decision a_t , and the costs we make in the future, based on our current decision, where we discount our future costs with a factor α . Of all the possible options a_t , we choose that one that minimizes our total discounted expected costs $V_t(s_t)$. We assume that V_t converges to some V for $t \rightarrow \infty$, which leads to the optimal total discounted expected costs:

$$V(s) = \min_{a \in \mathcal{A}(s)} \{c(a) + \alpha \sum_{s' \in \mathcal{S}} p^s(s'|s, a) V(s')\}, \quad (4.2)$$

4.3. Policy iteration

In this section, we look at a special case of (4.2), where we only consider how many water tenders we should send, independent of which one it is. We thus assume that the travel times of all water tenders are identical and independently distributed (i.i.d). This is different from what we did in Chapter 3, but, since this situation is easier to evaluate, we use this model to get some insight in the behaviour of the optimal policy. The optimal policy gives us for every situation s the optimal decision a . We will determine this optimal policy by using policy iteration. We let $s = \#$ water tenders at the fire stations and $a = \#$ water tenders sent from the fire stations to the incident. We assume that the cost function $c(a)$ is decreasing in a , which means that it is cheaper to send more water tenders. This makes sense, since sending more cars reduces the expected travel time.

For the policy iteration we consider

$$V(s) = \min_{a < s} \{c(a) + \alpha \sum_{k=0}^{N-s} p^k (1-p)^{N-s-k} \binom{N-s}{k} V(s-a+k)\},$$

where p is the probability that a water tender returns. Note that this value function equals the value in (4.2) after taken the limit to infinity. For policy iteration, we need a starting value, so we take $V(0) = 0$, for some “reference” state 0. The steps of the policy iteration method are as follows:

- 1 Take some policy P .
- 2 Compute V for this policy.
- 3 Find a better P' . If non exists: stop.
- 4 Set $P := P'$ and go to step 2.

Here we compute the new P' by

$$P'(s) = \arg \min_{a < s} \{c(a) + \alpha \sum_{s' \in \mathcal{S}} p_{t+1}^s(s'|s_t, a_t) V_{t+1}(s')\}.$$

The Matlab code (MATLAB, 2015) for this policy iteration can be found in the appendix.

We will now consider policy iteration for different distributions of the travel time.

4.3.1. T_i exponentially distributed

We consider exponential travel times, so $T_i \sim \text{Exp}(\lambda)$ for all $i = 1, \dots, N$. From Chapter 3, we know that the minimum of a i.i.d. exponentially distributed variables is $\text{Exp}(a\lambda)$ distributed. This, combined with the fact that the tail probability of T_i is $\mathbb{P}(T_i > M) = \exp(-\lambda M)$, gives us an expression for the cost function:

$$c(a) = \exp(-a\lambda M).$$

A graph of the cost function for $a = 1, \dots, 100$, $\lambda = 0.1$, and $M = 6$, is shown in Figure 4.2. We see that the cost function decreases fast, and is almost equal to zero at $a = 10$, which means that there is not much extra benefit if we send more than 10 water tenders. This seems realistic, since we don not expect the fastest water tender out of 100 water

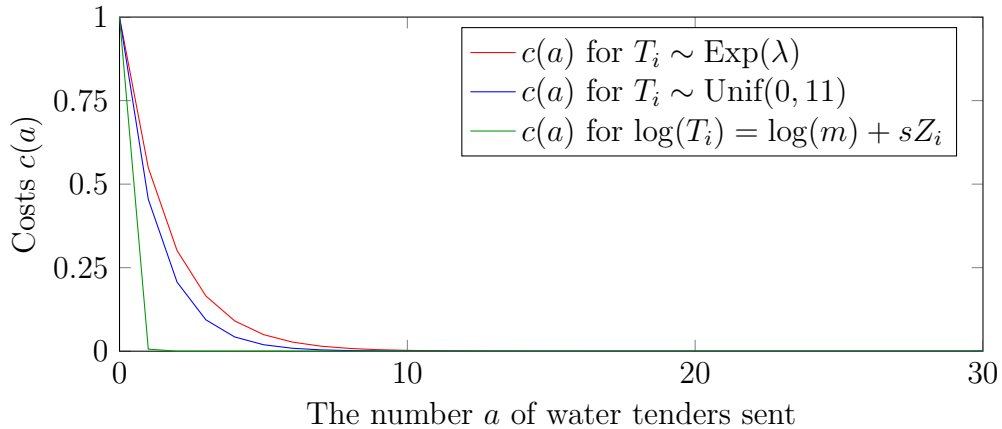


Figure 4.2.: Different cost functions for $a = 1, \dots, 100$, $M = 6$, $\lambda = 0.1$, $m = \frac{2}{5}$, $s = \frac{1}{2}$, $Z_i \sim \text{Student's } t(0, 0, \tau)$, and $\tau = 3$.

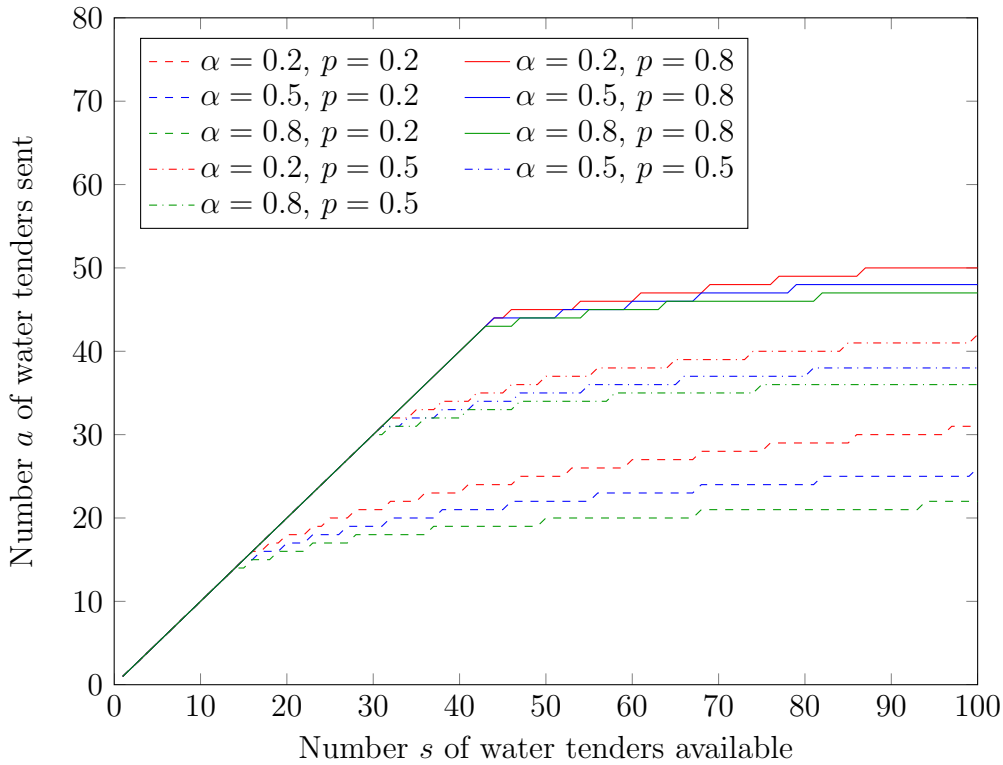


Figure 4.3.: Optimal policy for $N = 100$ water tenders, the travel time T_i is $\text{Exp}(0.1)$ distributed and the water tender needs to arrive in $M = 6$ minutes.

tenders be much faster than the fastest water tender out of, say, 20 water tenders. The optimal policy depends on our choice of p and α . A bigger p means that water tenders will come back more quickly, and a higher value of α means that we consider actions in the future more important. Note that when we have zero water tenders, there does not exist an optimal policy, since there is no choice to be made. A graph of the optimal policy for different values of p and α , if we consider $N = 100$, is shown in Figure 4.3. 100 water tenders seems like a lot, but, as we said before, we want to get some insight into the behaviour of the optimal policy P . In Figure 4.3 we see that the optimal policy is non-decreasing in s , the number water tenders available. It is remarkable that the optimal decision is to send 50 water tenders when there are 100 water tenders available, $\alpha = 0.2$, and $p = 0.8$. In reality we would never make such a decision, even if our water tenders would return fast. The reason that P increases intuitively fast probably lays in the fact that we do not bring costs into account for sending more water tenders, we only give a benefit. The reason that we, in reality, would never send so many water tenders is that for every water tender sent there need to come firemen along, if only for driving the water tender, and that costs gasoline. These are major disadvantages, which we did not take into account in this example.

4.3.2. T_i uniformly distributed

Now, we consider uniformly distributed travel times at the interval $(0, 11)$, so

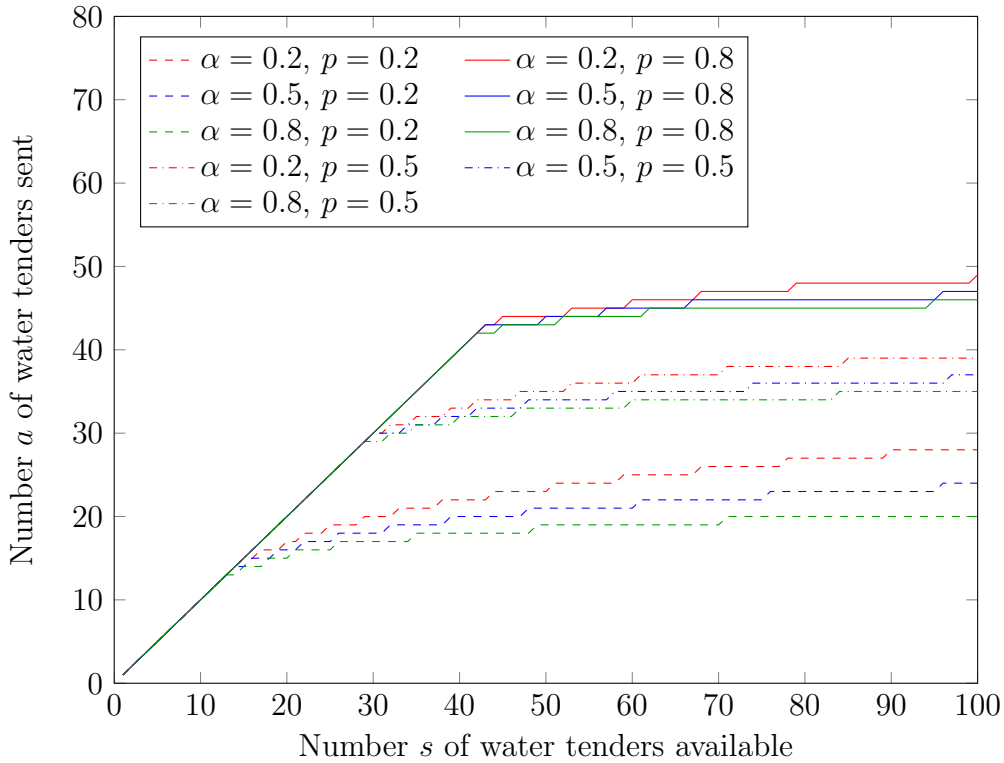


Figure 4.4.: Optimal policy for $N = 100$ water tenders, the travel time T_i is $\text{Unif}(0, 11)$ distributed and the water tender needs to arrive in $M = 6$ minutes.

$T_i \sim \text{Unif}(0, 11)$. This means that every water tender takes between 0 and 11 minutes to arrive at a fire, and that every travel time between 0 and 11 minutes is equally likely. We choose this interval, since most of our travel times considered in Chapter 2 are in this interval, and because we choose $M = 6$. The cost function then becomes

$$\begin{aligned} c(a) &= \mathbb{1}(T(a) \geq M) \\ &= \mathbb{P}(T_i \geq M)^a \\ &= \left(\frac{11-M}{11}\right)^a. \end{aligned}$$

A graph of the cost function, for $a = 1, \dots, 100$ and $M = 6$, is shown in Figure 4.2. We see that this cost is decreasing comparable fast to the cost function for $T_i \sim \text{Exp}(\lambda)$. Figure 4.4 shows a graph of the optimal policy for different values of p and α , if we consider $N = 100$. We see that the optimal policy P for $T_i \sim \text{Unif}(0, 11)$, is similar to the optimal policy for $T_i \sim \text{Exp}(\lambda)$. P increases only slightly slower, which is probably due to the fact that $c(a)$ decreases only slightly faster for $T_i \sim \text{Unif}(0, 11)$ than for $T_i \sim \text{Exp}(\lambda)$.

4.3.3. $\log(T_i)$ Student's t distributed

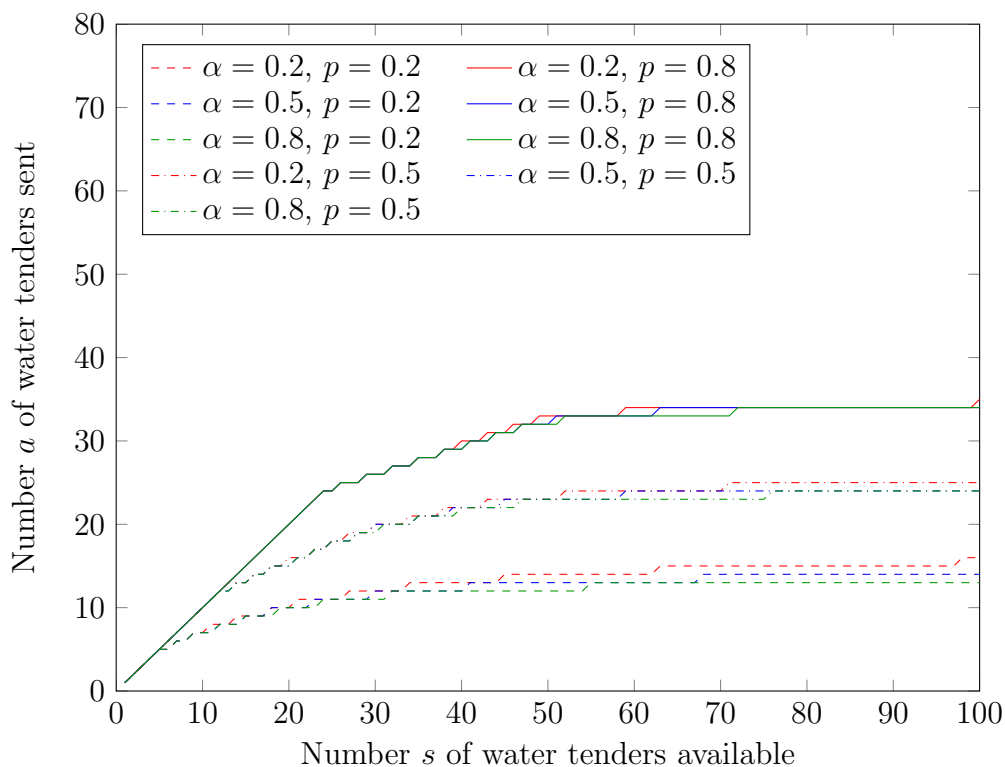


Figure 4.5.: Optimal policy for $N = 100$ water tenders, the travel time $\log(T_i)$ is Student's $t(\frac{2}{5}, \frac{1}{2}, 3)$ distributed, and the water tender needs to arrive in $M = 6$ minutes.

Recall that in Chapter 2, we discovered that the logarithm of the travel times were Student's t distributed with a certain shifting m and scaling s :

$$\log(T_i) = \log(m) + sZ_i,$$

for $i \in \{1, \dots, N\}$, where Z_i follows a centered Student's t distribution with τ degrees of freedom. To make analytical computations more clearly, we choose $\tau = 3$. Then, we have

$$\begin{aligned} f_{Z_i}(x) &= \frac{\Gamma(\frac{\tau+1}{2})}{\sqrt{\tau\pi}\Gamma(\frac{\tau}{2})} \left(1 + \frac{x^2}{\tau}\right)^{-\frac{\tau+1}{2}} \\ &= \frac{\Gamma(2)}{\sqrt{3\pi}\Gamma(\frac{3}{2})} \left(1 + \frac{x^2}{3}\right)^{-2} \\ &= \frac{6\sqrt{3}}{\pi(3 + x^2)^2}. \end{aligned}$$

The tail probability of Z_i then is

$$\begin{aligned}\mathbb{P}(Z_i > t) &= \int_t^\infty f_{Z_i}(x)dx \\ &= \int_t^\infty \frac{6\sqrt{3}}{\pi(3+x^2)^2}dx \\ &= \frac{1}{2} - \frac{\sqrt{3}t}{\pi(3+t^2)} - \frac{\text{Arctan}\left[\frac{t}{\sqrt{3}}\right]}{\pi}.\end{aligned}$$

The tail probability of T_i then becomes

$$\begin{aligned}\mathbb{P}(T_i > t) &= \mathbb{P}(\log(T_i) > \log(t)) \\ &= \mathbb{P}(\log(m) + cZ_i > \log(t)) \\ &= \mathbb{P}\left(Z_i > \frac{\log(t)-\log(m)}{c}\right) \\ &= \frac{1}{2} - \frac{\sqrt{3}\left(\frac{\log(t)-\log(m)}{c}\right)}{\pi\left(3 + \left(\frac{\log(t)-\log(m)}{c}\right)^2\right)} - \frac{\text{Arctan}\left[\frac{\log(t)-\log(m)}{c\sqrt{3}}\right]}{\pi}.\end{aligned}$$

Hence,

$$c(a) = \left(\frac{1}{2} - \frac{\sqrt{3}\left(\frac{\log(M)-\log(m)}{c}\right)}{\pi\left(3 + \left(\frac{\log(M)-\log(m)}{c}\right)^2\right)} - \frac{\text{Arctan}\left[\frac{\log(M)-\log(m)}{c\sqrt{3}}\right]}{\pi} \right)^a.$$

A graph of the cost function for $a = 1, \dots, 100$, $M = 6$, $m = \frac{2}{5}$ and $c = \frac{1}{2}$ is shown in Figure 4.2. We see that it decreases very fast. Only sending no water tenders at all results in direct costs, the direct costs of sending 2 or 20 water tenders has no difference in direct costs. Figure 4.5 shows the optimal policy. We see that it is non-decreasing, and that it increases more slowly than when T_i is exponentially or uniformly distributed.

4.4. Optimal decision

We consider again the optimal value function of (4.1). In the previous section, we saw that that the optimal policy function is non-decreasing. In this section, we want to prove this. Before we do this, we will prove that the optimal value function is partially non-increasing. For this, we follow the paper of Papadaki and Powell (2007).

First, we introduce the function $f(s_t, a_t) = s_t - a_t$, such that we have

$$s_{t+1} = f(s_t, a_t) + y_{t+1}.$$

We also introduce the probability $q_t^y(y_t)$, being the probability that y_t water tenders become available at the beginning of decision epoch t . This gives us

$$p_{t+1}^s(s_{t+1}|s_t, a_t) = q_{t+1}^y(s_{t+1} - f(s_t, a_t)). \quad (4.3)$$

Now, we give two definitions.

Definition 4.1. The partial ordering operator \preceq or \succeq are defined, on the set \mathcal{S} of N -dimensional vectors, as follows: we denote $s \preceq r$ for $s, r \in \mathcal{S}$, if $s(i) \leq r(i)$ for all $i \in \{1, \dots, N\}$, and $s \succeq r$ for $s, r \in \mathcal{S}$, if $s(i) \geq r(i)$ for all $i \in \{1, \dots, N\}$.

Definition 4.2. We say that a real-valued function F defined on an N -dimensional set \mathcal{S} is partially non-increasing if for all $r^+, r^- \in \mathcal{S}$ such that $r^+ \succeq r^-$, we have $F(r^+) \leq F(r^-)$.

Recall that $s_{t+1} = f(s_t, a_t) + y_{t+1}$, so we have $s_{t+1} \succeq f(s_t, a_t)$. Because of this we can rewrite the expectation in the optimality equation 4.1:

$$\sum_{s' \in \mathcal{S}} p_{t+1}^s(s'|s_t, a_t) V_{t+1}(s') = \sum_{s' \in \mathcal{S}, s' \succeq f(s_t, a_t)} p_{t+1}^s(s'|s_t, a_t) V_{t+1}(s').$$

To prove that $V_t(s_t)$ is partially non-increasing for all $t \geq 0$, we need the following lemma, by Papadaki and Powell (2007):

Lemma 4.3. *Let V_{t+1} be partially non-increasing in \mathcal{S} and let $x \succeq 0$. Then*

$$\sum_{i \in \mathcal{S}, i \succeq f(s+x, a)} p_{t+1}^s(i|s+x, a) V_{t+1}(i) \leq \sum_{i \in \mathcal{S}, i \succeq f(s, a)} p_{t+1}^s(i|s, a) V_{t+1}(i). \quad (4.4)$$

Proof. Note that since $x \succeq 0$, we have $f(s+x, a) \succeq f(s, a)$ for all $a \in \mathcal{A}$. So, since V_{t+1} is partially non-increasing, we have

$$\sum_{j \in \mathcal{S}, j \succeq 0} q_{t+1}^y(j) V_{t+1}(i + f(s+x, a)) \leq \sum_{k \in \mathcal{S}, k \succeq 0} q_{t+1}^y(k) V_{t+1}(i + f(s, a)).$$

We now substitute $j = i - f(s+x, a)$ in the left-handside of the above equation and $k = i - f(s, a)$ in the right-handside of the above equation. This gives us

$$\sum_{i \in \mathcal{S}, i \succeq f(s+x, a)} q_{t+1}^y(i - f(s+x, a)) V_{t+1}(i) \leq \sum_{i \in \mathcal{S}, i \succeq f(s, a)} q_{t+1}^y(i - f(s, a)) V_{t+1}(i).$$

We can use equation (4.3) to rewrite $q_{t+1}^y(i - f(s, a))$:

$$q_{t+1}^y(i - f(s, a)) = p_{t+1}^s(i|s, a).$$

Hence, we have

$$\sum_{i \in \mathcal{S}, i \succeq f(s+x, a)} p_{t+1}^s(i|s+x, a) V_{t+1}(i) \leq \sum_{i \in \mathcal{S}, i \succeq f(s, a)} p_{t+1}^s(i|s, a) V_{t+1}(i).$$

So, indeed, when V_{t+1} is partially non-increasing in \mathcal{S} and $x \succeq 0$, equation (4.4) holds. This proves the lemma. \square

Now, we have proved this lemma, we can prove that $V_t(s_t)$ is partially non-increasing for all $t \geq 0$ (Papadaki and Powell, 2007).

Theorem 4.4. *Let $c(a_t)$ be partially non-increasing for all $a_t \in \mathcal{A}$, $t \geq 0$. Then, the value function $V_t(s_t)$ is partially non-increasing in s_t for all $t \geq 0$, and thus $V(s)$ is non-increasing for all $s \in \mathcal{S}$.*

Proof. We will prove this theorem by induction. Assume that V_n is partially non-increasing for $n = t+1, \dots, T > 0$. We want to prove that V_t is partially non-increasing. Recall from equation 4.1 that V_t is defined as

$$V_t(s_t) = \min_{a_t \in \mathcal{A}} \{c(a_t) + \alpha \sum_{i \in \mathcal{S}, i \succeq f(s_t, a_t)} p_{t+1}^s(i|s_t, a_t) V_{t+1}(i)\}.$$

Since the action space is finite (there are only N water tenders, thus for every time epoch we have at most N actions), there is an action a_t^+ which attains the minimum of $V_t(s_t^+)$ for some state s_t^+ . So we have

$$V_t(s_t^+) = c(a_t^+) + \alpha \sum_{i \in \mathcal{S}, i \succeq f(s_t^+, a_t^+)} p_{t+1}^s(i|s_t^+, a_t^+) V_{t+1}(i).$$

Now, if we consider a state s_t^- such that $s_t^+ \succeq s_t^-$, we get, according to lemma 4.3,

$$\begin{aligned} V_t(s_t^+) &= c(a_t^+) + \alpha \sum_{i \in \mathcal{S}, i \succeq f(s_t^+, a_t^+)} p_{t+1}^s(i|s_t^+, a_t^+) V_{t+1}(i) \\ &\leq c(a_t) + \alpha \sum_{i \in \mathcal{S}, i \succeq f(s_t^-, a_t)} p_{t+1}^s(i|s_t^-, a_t) V_{t+1}(i) \\ &\leq \min_{a_t \in \mathcal{A}} \{c(a_t) + \alpha \sum_{i \in \mathcal{S}, i \succeq f(s_t^-, a_t)} p_{t+1}^s(i|s_t^-, a_t) V_{t+1}(i)\} \\ &= V_t(s_t^-). \end{aligned}$$

Hence, V_t is partially non-increasing. Note that if we take the limit of $T \rightarrow \infty$, we get that V_t is partially non-increasing for all $t \geq 0$, and hence V ,

$$V(s) = \min_{a \in \mathcal{A}(s)} \{c(a) + \alpha \sum_{i \in \mathcal{S}, i \succeq f(s, a)} p^s(i|s, a) V(i)\},$$

is non-increasing in $s \in \mathcal{S}$. □

Before we arrive at the main result of this chapter, Theorem 4.8, we need Topkis's theorem. First, we give a definition.

Definition 4.5. A function $W(s, a)$ is submodular in (s, a) if

$$W(s, a+1) - W(s, a) \geq W(s+1, a+1) - W(s+1, a).$$

Theorem 4.6 (Topkis). *Let*

$$\begin{aligned} P(s) &= \arg \min_{a \in \mathcal{A}(s)} \{c(a) + \alpha \sum_{j \in \mathcal{S}, j \succeq f(s, a)} p^s(j|s, a) V(j)\} \\ &= \arg \min_{a \in \mathcal{A}(s)} \{W(s, a)\}. \end{aligned}$$

If $W(s, a)$ is submodular, then $P(s)$ is non-decreasing in $s \in \mathcal{S}$.

A proof of this theorem can be found in Krishnamurthy (2016).

Now, we come to the final result of this chapter. We will prove that $P(s)$ is non-decreasing in $s \in \mathcal{S}$. For this, we need that $V(s)$ is convex. Convexity is defined as follows:

Definition 4.7. A function $V(s)$ is convex if

$$2V(s) \leq V(s+1) + V(s-1).$$

Theorem 4.8. *If $V(s)$ is convex, then $P(s)$ is non-decreasing in $s \in \mathcal{S}$.*

Proof. For $P(s)$ we have that

$$\begin{aligned} P(s) &= \arg \min_{a \in \mathcal{A}(s)} \{c(a) + \alpha \sum_{j \in \mathcal{S}, j \geq f(s,a)} p^s(j|s, a)V(j)\} \\ &= \arg \min_{a \in \mathcal{A}(s)} \{W(s, a)\}, \end{aligned}$$

where,

$$\begin{aligned} W(s, a) &= c(a) + \alpha \sum_{j \in \mathcal{S}, j \geq f(s,a)} p^s(j|s, a)V(j) \\ &= c(a) + \alpha \sum_{j \in \mathcal{S}, j \geq f(s,a)} q^y(j - f(s, a))V(j) \\ &= c(a) + \alpha \sum_{i \in \mathcal{S}, i \geq 0} q^y(i)V(i + f(s, a)) \\ &= c(a) + \alpha \sum_{i \in \mathcal{S}, i \geq 0} q^y(i)V(i + s - a). \end{aligned}$$

According to Topkis's theorem, submodularity of $W(s, a)$ is sufficient for $P(s)$ to be non-decreasing in $s \in \mathcal{S}$. Recall that $W(s, a)$ is submodular in (s, a) if

$$W(s, a+1) - W(s, a) \geq W(s+1, a+1) - W(s+1, a).$$

We define $W(s) = W(s, a+1) - W(s, a)$ and rewrite it as follows:

$$\begin{aligned} W(s) &= W(s, a+1) - W(s, a) \\ &= c(a+1) - c(a) + \alpha \sum_{i \in \mathcal{S}, i \geq 0} q^y(i) \left(V(i + s - a - 1) - V(i + s - a) \right). \end{aligned}$$

$W(s+1)$ now becomes

$$\begin{aligned} W(s+1) &= W(s+1, a+1) - W(s+1, a) \\ &= c(a+1) - c(a) + \alpha \sum_{i \in \mathcal{S}, i \geq 0} q^y(i) \left(V(i + s - a) - V(i + s - a + 1) \right). \end{aligned}$$

Now, we can evaluate if $W(s) \geq W(s + 1)$:

$$\begin{aligned} W(s) &\geq W(s + 1) \\ V(i + s - a - 1) - (V(i + s - a) &\geq V(i + s - a) - (V(i + s - a + 1) \\ V(i + s - a - 1) + (V(i + s - a + 1) &\geq 2V(i + s - a). \end{aligned}$$

Hence, $W(s, a)$ is submodular in (s, a) if and only if $V(s)$ is convex. Thus, if $V(s)$ is convex, then $P(s)$ is non-decreasing in $s \in \mathcal{S}$. This proves the theorem. \square

In this chapter, we considered concurrent incidents. We modelled this as a Markov Decision Process, and used policy iteration to get insight in the behaviour of the optimal policy in a specific situation, namely when it only matters how many water tenders should be sent to an incident, and not which ones. We concluded that it is likely that the optimal policy is non-decreasing, meaning that if we have more water tenders available, we should not send less water tenders to an incident than in the situation where we have less water tenders. We ended this chapter with a proof of this presumption. The model proposed in this chapter is a basic model, which has to be extended further, before drawing useful concluding for the BWAA.

5. Conclusion

In this thesis, we have taken a deeper look into the model of Budge et al. (2010) on travel times depending on the distance. We worked out the details of the model and applied it to the GPS-data we received from the fire department in the area of Amsterdam, the Netherlands, the BWAA. We found that the model gives slightly better results to our data of the BWAA, than it gives to the data from Calgary, Alberta, that Budge et al. (2010) used. We think that there is still a lot of potential improvement. The dataset we used was rather small and a larger set of GPS-data over a longer period of time should give more useful results. Although we find a better fit than Budge et al. (2010), we think they should be much better if we want to draw real conclusions from it.

Further research should provide a better explanation of the variability due to weather or traffic conditions. This variability depends for a large part on the time of year and the time of day. Applying this in more detail to the model could lead to a better fit of the model to the data.

Next, we looked into the situation of non-concurrent incidents. We proposed different situations, varying from having an opportunity to take the same route to sharing a part of a route, and modelled it using a hyperexponential distribution and a hypoexponential distribution respectively, to make it insightful. One reason we choose a hyperexponential and a hypoexponential distribution is because they look rather simple. Another reason is that the hyperexponential distribution has a decreasing failure rate while the hypoexponential distribution has an increasing failure rate. This means that in the case of a hyperexponential travel time a long ride has a high probability of being even a lot longer, while in the case of a hypoexponential travel time a long ride has a high probability of being almost finished. We proved that in both cases independent travel times lead to a shorter expected travel time, when the expectation of the single travel times are the same, and that there are even situations where the expectation of the single travel times are different, but the independent travel times still lead to a shorter expected travel time. Since we only considered distributions which are based on the exponential distribution, the results might not be directly applicable to reality.

We would recommend further research on the difference between minima based on dependent and independent distributions, using distributions as the normal or the Student's t distribution.

We concluded this thesis with considering concurrent incidents. We modelled this as a discounted infinite horizon Markov Decision Process in which we considered independent travel times, and a cost function, which is decreasing in the number of water tenders sent. We also assumed that water tenders come back according to a geometric distribution.

We proved that the optimal value function is non-increasing in the number of water tenders available, and that the optimal policy is non-decreasing in the number of water tenders available.

The model we proposed, is a basic model, which has to be extended further, before drawing useful concluding for real fire departments, as the BWAA. A good extension to this model would be to include the geographical distribution according to which incidents take place, so the distance between an incident and a fire station is not the same for all fire stations. The model can also be extended by considering a time distribution according to which incidents take place, so we can use a continuous time scale, instead of considering a time scale with time epochs based on the occurrence of an incident. Another extension would be to take dependence into account, leading to not only deciding how many water tenders should be sent, but also which water tenders should be sent, when every fire station has only one water tender.

Bibliography

- Archibald, R., Walker, W., Chaiken, J., Ignall, E., and Archibald, R. (1979). Managing change in the fire department. *Fire Department Deployment Analysis: A Public Policy Analysis Case Study, North-Holland, New York*.
- Budge, S., Ingolfsson, A., and Zerom, D. (2010). Empirical Analysis of Ambulance Travel Times: The Case of Calgary Emergency Medical Services. *Management Science*, 56(4):716–723.
- Church, R. L. and Davis, R. R. (1974). The fixed charge maximal covering location problem. *Papers in Regional Science*, 32(1):101–118.
- Ignall, E., Carter, G., and Rider, K. (1982). An Algorithm for the Initial Dispatch of Fire Companies. *Management Science*, 28(4):366–378.
- Ingolfsson, A., Budge, S., and Erkut, E. (2008). Optimal ambulance location with random delays and travel times. *Health Care Management Science*, 11(3):262–274.
- Ingolfsson, A., Erkut, E., and Budge, S. (2003). Simulation of Single Start Station for Edmonton EMS Author(s): A Simulation of single start station for Edmonton EMS. *The Journal of the Operational Research Society journal of the Operational Research Society Journal of the Operational Research Society*, 54(54):736–746.
- Jagtenberg, C., Bhulai, S., and Van der Mei, R. (2015). An efficient heuristic for real-time ambulance redeployment. *Operations Research for Health Care*, 4:27–35.
- Jenelius, E. and Koutsopoulos, H. N. (2013). Travel time estimation for urban road networks using low frequency probe vehicle data. *Transportation Research Part B: Methodological*, 53:64–81.
- Johnson, N. L., Kotz, N., and Balakrishnan, N. (1994). Continuous Univariate Distributions. *New York, Wiley*, 2:375.
- Kolesar, P. J. (2012). Some Lessons on Operations Management Model Implementation Drawn from the RAND Fire Project. *Manufacturing & Service Operations Management*, 14(1):1–6.
- Kolesar, P. J., Walker, W., and Hausner, J. (1975). Determining the Relation between Fire Engine Travel Times and Travel Distances in New York City. *Operations Research*, 23(4).

- Krishnamurthy, V. (2016). *Partially Observed Markov Decision Processes: From Filtering to Controlled Sensing*. Cambridge University Press.
- Larson, R. C. (1972). *Urban police patrol analysis*, volume 28. MIT Press Cambridge, MA.
- MATLAB (2015). *MATLAB version 8.5.0.197613 (R2015a)*. The Mathworks, Inc., Natick, Massachusetts.
- Maxwell, M. S., Henderson, S. G., and Topaloglu, H. (2009). Ambulance Redeployment: An Approximate Dynamic Programming Approach. In *Winter Simulation Conference, WSC '09*, pages 1850–1860.
- Papadaki, K. and Powell, W. B. (2007). Monotonicity in multidimensional markov decision processes for the batch dispatch problem. *Operations research letters*, 35(2):267–272.
- QGIS Development Team (2009). *QGIS Geographic Information System*. Open Source Geospatial Foundation.
- Rigby, R. A. and Stasinopoulos, D. M. (2006). Using the Box-Cox t distribution in GAMLSS to model skewness and kurtosis. *Statistical Modelling*, 6(3):209–229.
- RStudio Team (2015). *RStudio: Integrated Development Environment for R*. RStudio, Inc., Boston, MA.
- Swersey, A. J. (1982). A Markovian Decision Model for Deciding How Many Fire Companies to Dispatch. *Management Science*, 28(4):352–365.
- Van Barneveld, T., Bhulai, S., and Van der Mei, R. (2015). A dynamic ambulance management model for rural areas. *Health care management science*, pages 1–22.
- Van Barneveld, T., Bhulai, S., and Van der Mei, R. (2016). The effect of ambulance relocations on the performance of ambulance service providers. *European Journal of Operational Research*, 252(1):257–269.
- Van den Berg, P. L. (2016). *Logistics of Emergency Response Vehicles*. PhD thesis, TU Delft.
- Westgate, B. S., Woodard, D. B., Matteson, D. S., and Henderson, S. G. (2013). Travel time estimation for ambulances using Bayesian data augmentation. *Annals of Applied Statistics*, 7(2):1139–1161.

A. Matlab code

Listing A.1: Best policy

```
1 function [Best_policy]=Best_policy(N,alpha,p,c)
2 %Finds the optimal policy for the optimal discounted value function
3 %V(x)=min_{a<=x}{(c(a)+alpha*sum_{i=0}^{N-x}(p^i (1-p)^{N-x-i} BCD(N-x)(i)*
4 %V(x-a+i))
5 %for all x in (0,...,N).
6 P=[1:N];
7 Q=zeros(1,N);
8 maxit=1000;
9 j=1;
10 while abs(sum(P-Q))>0 && j<maxit
11     Q=P;
12     V=sym(zeros(1,N));
13     syms V_0
14     for i=1:N
15         V(i) = strcat('V_',num2str(i) );
16     end
17     V_g=[0,V];
18     Eqn=sym(zeros(1,N+1));
19     Eqn(1)=V_0 == binsum(V_g,0,0,N,alpha,p,c);
20     for i=2:N+1
21         Eqn(i) = V_g(i) == binsum(V_g,i-1,P(i-1),N,alpha,p,c);
22     end
23     [A,B] = equationsToMatrix(Eqn,[V_0,V]);
24     X=linsolve(A,B);
25     V=double(X);
26     P = policy(V,p,alpha,c,N);
27     j=j+1
28 end
29 Best_policy=P;
30 end
```

Listing A.2: policy

```

1 function [P]=policy(V,p,alpha,c,N)
2 %Gives a solution to
3 % $P(x)=\operatorname{argmin}_{a \leq x} \{ (c(a) + \sum_{i=0}^{N-x} (p^i (1-p)^{N-x-i}) \operatorname{BCD}(N-x)(i) *$ 
4 % $V(x-a+i)$ 
5 %for all  $x$  in  $(0, \dots, N)$ .
6 P=zeros(1,N);
7 for i = 1:N
8     R=ones(1,i+1);
9     for j=1:length(R)
10        R(j)=binsum(V,i,j-1,N,alpha,p,c);
11    end
12    [argvalue,argmin]=min(R);
13    P(i)=argmin-1;
14 end

```

Listing A.3: binsum

```

1 function [value]=binsum(V,x,a,N,alpha,p,c)
2 %Gives a solution to
3 % $(c(a) + \alpha * \sum_{i=0}^{N-x} (p^i (1-p)^{N-x-i}) \operatorname{BCD}(N-x)(i) * V(x-a+i))$ .
4 Y=sym(zeros(1,N-x+1));
5     for i=0:N-x
6         Y(i+1)=binopdf(i,N-x,p)*V(x-a+i+1);
7     end
8 value_start=sum(Y);
9 value=c(a)+alpha*value_start;
10 end

```