

Een Statistische Analyse van Klantgedrag
in Callcentra



Martijn van Zanten
BWI - Werkstuk

Statistische Analyse van Klantgedrag in Callcentra

Martijn van Zanten

Begeleider: Alex Roubos

**Vrije Universiteit Amsterdam
Faculteit der Exacte Wetenschappen
Studierichting Bedrijfswiskunde en Informatica
De Boelelaan 1081a
1081 HV Amsterdam**

September 2011

Voorwoord

Als afsluitend onderdeel voor de Master Business Mathematics and Informatics is deze scriptie geschreven. De Master Business Mathematics and Informatics is een academische opleiding gericht op het oplossen van bedrijfsproblemen. De opleiding is Engelstalig en multidisciplinair: In het theoretische deel van de opleiding doe je kennis op van de disciplines wiskunde, informatica en bedrijfseconomie. Deze drie vakgebieden vormen de basis bij het analyseren van bedrijfsproblemen en het zoeken naar optimale oplossing.

Het doel van deze scriptie is het beschrijven van een bedrijfskundig onderzoek waarin de drie verschillende disciplines zo duidelijk mogelijk aan bod komen. Hiervoor wordt uitgebreid literatuur onderzoek gedaan en waar nodig enig eigen onderzoek verricht.

Deze scriptie zal zich richten op het gebied van de callcentra. Callcentra kunnen als een, door het in literatuur vaak beschreven, wachtrijsysteem gemodelleerd worden. Om een kort maar duidelijk beeld te schetsen over het modelleren van callcentra zullen de verschillende modellen uit de literatuur in deze scriptie samengevat worden en daarbij toegepast worden op een dataset van een bestaand callcenter ter validatie.

Tot slot wil ik alvast van deze gelegenheid gebruik maken om mijn begeleider Alex Roubos te bedanken voor het opzetten van het onderzoeksplan en de steun en begeleiding tijdens het maken van dit werkstuk.

Martijn van Zanten
Beverwijk, 2011

Samenvatting

Callcentra zijn in deze tijd niet meer weg te denken als het gaat om de communicatie tussen een organisatie en haar klanten. Door goedlopende en klantvriendelijk 'contactmogelijkheden' aan te bieden proberen veel organisaties zich als klantvriendelijk te bestempelen. Om dit te kunnen verwezenlijken is dan ook hoge efficiëntie een vereiste. Het modelleren van een callcenter draagt bij aan een goed onderzoek voor dit optimalisatie proces.

Het modelleren van het proces binnen een callcenter gebeurt doormiddel van twee belangrijke processen met ieder hun parameter waarbij verschillende aannames worden gemaakt. Het eerste proces is het aankomst proces. Bij een wachtrijmodel wordt uitgegaan van aankomsten die plaatsvinden volgens een Poisson proces met parameter λ (lambda). Deze parameter geeft het gemiddeld aantal aankomsten per tijdsperiode aan. Waarbij aankomsten worden behandeld aan de hand van het "First Come First Served" (FCFS) principe.

Het tweede proces is het serviceproces met parameter μ (mu) die de servicetijd van de agents aangeeft. De servicetijd is Exponentieel verdeeld met een gemiddelde servicetijd van $\frac{1}{\mu}$.

In dit BWI-werkstuk is aan de hand van de modellen uit de literatuur onderzoek gedaan naar de toepasbaarheid van deze modellen op data van een bestaand callcenter. Zijn de modellen nog valide en dus te gebruiken of zijn er tegenstrijdige resultaten te ontdekken?

Uit dit onderzoek is gebleken dat bij het eerste proces de aankomstintensiteit het best per tijdsinterval berekend kan worden. Hierdoor is het aankomstproces nauwkeuriger te modelleren en is het verschil met de werkelijkheid het kleinst.

Het tweede proces, het service proces, is moeilijker te modelleren. Volgens modellen uit de literatuur is de servicetijd exponentieel verdeeld en zo te modelleren. Echter uit dit onderzoek is gebleken dat dit niet 100% te bewijzen is bij zowel de servicetijd over de hele dag als per tijdsinterval. Om een goed model te kunnen creëren is dus verder onderzoek hiernaar een vereiste. Tot slot is er onderzoek gedaan naar een apart proces, namelijk het wachtproces. Dit proces gaat over de wachttijd van een klant tot hij of zij geholpen wordt. In de literatuur gaat men uit van twee typen wachttijden. Er is een wachttijd die een klant minstens moet wachten voordat hij of zij geholpen wordt en er is een wachttijd die de klant over heeft om geholpen te worden. De eerste noemt men ook wel de virtuele wachttijd en de tweede is ook wel het geduld van klanten. In de literatuur gaat men er vanuit dat bij het modelleren van de wachttijd, die exponentieel verdeeld zou zijn, er geen sprake is van afhakers. Binnen dit onderzoek was er echter wel sprake van afhakers en is gebleken dat ook zonder deze afhakers er niet 100% sprake is van een wachttijd die exponentieel verdeeld is. Afhakers hebben daarbij een grote invloed op dit proces en er is gebleken dat afhaken afhankelijk is van factoren in het wachtproces, zoals een informatiebericht tijdens het wachten.

Inhoudsopgave

Voorwoord	iii
Samenvatting.....	v
1 Inleiding.....	1
1.1 Doel van de scriptie	2
2 Modellen	3
2.1 M/M/S	3
2.2 M(t)/M(t)/S(t).....	3
3 Dataset	5
3.1 Openingstijden	5
3.2 Typen aankomsten	5
3.3 Operationele veranderingen.....	5
3.4 Data bewerking	6
4 Het aankomstproces	7
4.1 Het homogene Poisson proces.....	7
4.2 Het tijd in-homogene Poisson proces.....	9
4.2.1 PS aankomsten tussen 07:00 en 24:00 per maand.....	10
4.2.2 PS aankomsten `s morgens, `s middags en `s avonds in januari-december.....	12
4.2.3 PS aankomsten tussen 07:00 en 24:00 in de maanden november en december	13
4.2.4 IN aankomsten tussen 07:00 en 24:00 in de maanden november en december	14
5 Het serviceproces.....	15
5.1 Korte servicetijden.....	15
5.2 De distributie van de servicetijden in November en December	16
5.3 De distributie van de servicetijden in de overige maanden	19
5.4 De servicetijd afhankelijk van de tijd.....	21
6 Het wachtproces	25
6.1 De invloed van afhakers op de gemiddelde wachttijd	25
6.2 Het geduld van klanten.....	28
6.3 De virtuele wachttijd	31
7 Conclusie	35
8 Referenties	37
9 Bijlagen.....	39
9.1 Methode voor het valideren van de nulhypothese voor het aankomstproces	39

1 Inleiding

Callcentra zijn in deze tijd niet meer weg te denken als het gaat om de communicatie tussen een organisatie en haar klanten. Klanten binden aan je organisatie is dan ook één van de belangrijkste middelen om je staande te houden in deze economische zwaardere tijden. Een callcenter is een afdeling of organisatie die is gespecialiseerd in het afhandelen van voornamelijk telefoonverkeer. Callcentra zijn in deze tijd kenmerkend door de vaak grote ruimtes met grote aantallen medewerkers, ook wel agents genoemd, en de hoge efficiëntie. Doormiddel van goedlopende en klantvriendelijke contactmogelijkheden aan te kunnen bieden proberen veel organisaties zich als klantvriendelijk te bestempelen. Om dit te kunnen verwezenlijken is dan ook hoge efficiëntie een vereiste. Deze hoge efficiëntie is het resultaat van veel onderzoek naar het proces binnen callcentra nadat in de jaren negentig de switch plaatsvond van callcentra naar contactcentra. Deze switch zorgde er voor dat men veel klantgericht wilde werken en er dus efficiënt gewerkt moest gaan worden. Er ontstonden normen waar bijvoorbeeld het niet langer dan een X aantal tijdsperiodes mag duren voordat iemand een agent aan de lijn krijgt.

Echter dit soort callcentra hebben hoge operationele kosten waarbij de personeelskosten vaak rond der 70 procent ligt. Door de normen, de groei van de callcentra industrie en de daarbij horende hoge kosten is er vraag naar een moderne en academische aanpak als het gaat om de optimalisatie van de inzet van personeel en technologische middelen.

Het modelleren van een callcenter draagt bij aan een goed onderzoek voor het optimalisatie proces. Goedlopende callcentra hebben een zo optimaal mogelijke verhouding tussen het inzetten van personeel en het bereiken van een bepaald service level voor de klant. Het zo optimaal modelleren van het proces dat zich afspeelt bij een callcenter is dus van groot belang om het proces te kunnen sturen en daardoor efficiënt en optimaal te kunnen werken.

Bij het modelleren van het proces van callcentra kijkt men naar de twee hoofdfactoren in dit proces. Als eerst heeft men te maken met de 'inbound calls'. Deze inkomende telefoontje/berichten zijn moeilijk te voorspellen maar in grote lijnen wel te modelleren zodat een schatting gedaan kan worden naar het aantal aankomsten per tijdsperiode.

Vervolgens is men geïnteresseerd in de 'service tijd', oftewel hoelang duurt de service zelf eigenlijk. De service tijd is het belangrijkste meetpunt voor een callcenter. Hoelang iemand moet wachten voordat hij of zijn geholpen wordt hangt onder andere af van de service tijd van zijn of haar voorganger.

In het volgende hoofdstuk zal kort beschreven worden welke twee verschillende modellen met deze factoren zullen worden bekeken. Het eerste model is een model op basis van vaste parameters dat bijna alle callcentra gebruiken. Bij het tweede model zal gekeken worden naar een meer realistischere aanpak, namelijk naar tijdsafhankelijke parameters. Dit er van uitgaande dat in de realiteit er pieken en dalen zijn wat betreft het aankomst en service proces. In het laatste hoofdstuk zal een apart onderzoek worden gedaan met betrekking tot virtuele wachttijden en zogenaamde afhakers.

1.1 Doel van de scriptie

Het doel van deze scriptie is het valideren van de modellen uit de literatuur.

Zijn de gegeven modellen uit de literatuur realistisch met betrekking tot echte data van een bestaand callcenter?

Als eerst zullen de modellen kort beschreven worden en vervolgens zullen ze getoetst worden doormiddel van de dataset. Tot slot worden de modellen uitgebreid tot een variant waarbij ongeduldige klanten een rol gaan spelen. *Oftewel is er iets te zeggen en/of modelleren als het gaat om het ongeduld van klanten?* Omdat hierover nog niet veel in de literatuur te vinden is zal daarom de dataset gebruikt worden om hier iets over te kunnen zeggen.

2 Modellen

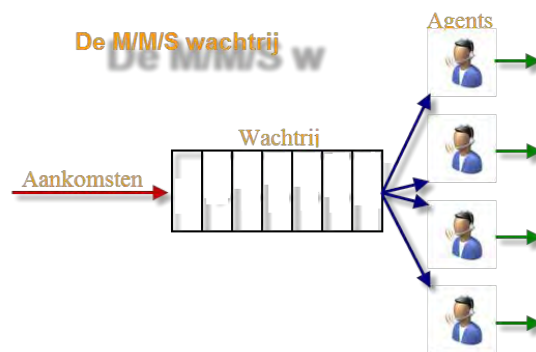
2.1 M/M/S

Het meest gebruikte model voor een callcenter is het M/M/S, ook wel Erlang C, wachtrijmodel. (figuur 1) Dit model bestaat uit aankomsten, een wachtrij en agents. In dit model gaan we uit van aankomsten, bijvoorbeeld telefoontjes of e-mail berichten, die achteraanluiten bij een (virtuele) 'wachtrij'. Wanneer de wachtrij leeg is en een agent vrij dan kan de aankomst direct worden behandeld. Wanneer alle agents bezet zijn of er nog andere aankomsten in de wachtrij zitten dan sluit de aankomst achteraan in de wachtrij. In dit model worden aankomsten dus behandeld via het "First Come First Served" (FCFS) principe.

Het modelleren van dit proces gebeurt doormiddel van drie parameters waarbij we verschillende aannames maken. De eerste parameter is die van de aankomsten. Bij een M/M/S wachtrijmodel wordt uitgegaan van aankomsten die plaatsvinden volgens een Poisson proces met parameter λ (lambda). Deze parameter geeft het gemiddeld aantal aankomsten per tijdsperiode aan.

Vervolgens is er de parameter μ (mu) die servicetijd van de agents aangeeft. De servicetijd is Exponentieel verdeeld met een gemiddelde servicetijd van $\frac{1}{\mu}$. Als laatste is de parameter S een waarde die aangeeft hoeveel agents er aan het werk zijn.

Om een M/M/S wachtrijmodel vervolgens toe te passen zullen deze parameters moeten worden verkregen uit historische data. In de volgende hoofdstukken zal uitgebreid ingegaan worden op de drie parameters en zal doormiddel van historische data geverifieerd worden of de aannames wel valide zijn.



Figuur 1: De M/M/S wachtrij

2.2 M(t)/M(t)/S(t)

Een meer realistischer model zou een aanpassing zijn van het vorige model waarbij er gekeken wordt naar de tijdsperiode. In deze situatie is er sprake van een M(t)/M(t)/S(t) wachtrijmodel waarbij het aantal aankomsten, de servicetijd en het aantal agents afhankelijk van de tijd zullen zijn. Het proces blijft hetzelfde, echter de parameters zijn nu tijdsafhankelijk.

Zoals hiervoor beschreven zullen ook deze aannames geverifieerd worden in de volgende hoofdstukken doormiddel van een dataset.

3 Dataset

Voor dit onderzoek wordt gebruik gemaakt van echte callcenter data. De data is verzameld door een callcenter van een bank uit Israël. De data is in alle maanden van 1999 verzameld en beschikbaar gesteld voor onderzoek. In dit hoofdstuk zal kort ingegaan worden op de data en op welke punten gelet moet worden.

3.1 *Openingstijden*

In Israël zijn werkdagen anders als bijvoorbeeld bij ons in Nederland. Daar worden namelijk de dagen zondag tot en met donderdag beschouwd als werkdagen. Tijdens deze werkdagen is het callcenter open van zeven uur `s morgens tot middernacht. Het is dus van belang tijdens dit onderzoek alleen die data te gebruiken die tijdens werkdagen is verzameld omdat er dan uit gegaan kan worden van volledige operationele openingstijden.

3.2 *Typen aankomsten*

In de twaalf maanden van 1999 zijn ongeveer 1,5 miljoen telefoontjes binnen gekomen bij het callcenter. Van deze telefoontjes zijn verschillende typen te classificeren. Elk type heeft een ander doel voor ogen wanneer hij of zij naar het callcenter belt. De vier belangrijkste en meest voorkomende typen voor dit onderzoek zijn:

- PS: Deze reguliere telefoontjes hebben een algemene vraag met betrekking tot diensten van de bank.
- NE: Dit zijn telefoontjes die betrekking hebben op aandelen transacties.
- NW: Deze type telefoontjes zijn nieuwe potentiële klanten voor de bank.
- IN: Dit gaat om klanten die internet assistentie nodig hebben.

In dit onderzoek zal wanneer dit het geval is alleen gekeken worden naar deze vier typen telefoontjes.

3.3 *Operationele veranderingen*

Door het jaar heen zijn twee belangrijke operationele veranderingen doorgevoerd bij het callcenter. De eerste heeft plaatsgevonden in de maand augustus. Vanaf de maand augustus werden alle IN aankomsten behandeld door speciale agents en niet meer door agents die alle typen aankomsten behandelden.

De tweede verandering vond plaats eind oktober. Vanaf eind oktober heeft men data omtrent service tijd anders opgeslagen en vandaar dat we hierin ook verschil zullen zien.

3.4 Data bewerking

In het volgende hoofdstuk zal gebruik worden gemaakt van de Kolmogorov-Smirnov toets. Voordat deze toets gebruikt kan worden is er een algoritme geschreven en enig databewerking verricht. Het algoritme zal, zoals ook naar voren komt in hoofdstuk 4, als bijlage worden toegevoegd.

Voordat dit algoritme gebruikt kan worden zijn alle tijdstippen opgezet naar aantallen seconden. Echter omdat deze tijdstippen zijn afgerond en er anders nullen ontstaan in het algoritme bij het totaal aantal seconden nog eens een random waarde $[0,1]$ opgeteld.

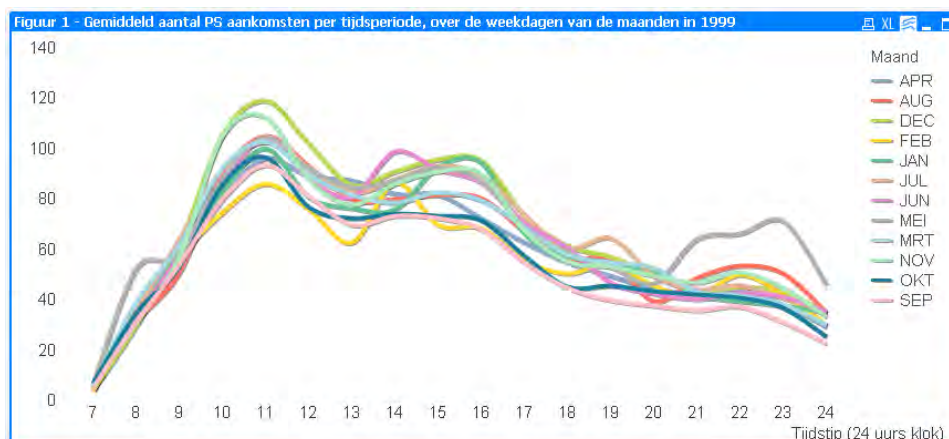
4 Het aankomstproces

Het aankomstproces beschrijft het proces van telefoontjes of e-mails bij een callcenter. Deze aankomsten zijn willekeurig en kunnen als volgt beschreven worden. Er zijn een groot aantal potentiële, statistisch identieke bellers die onafhankelijk van elkaar met een bepaalde kans naar het callcenter zullen bellen op een willekeurig gekozen moment. Volgens deze beschrijving is het mogelijk het aankomst proces als een Poisson proces met een gemiddelde aankomst van λ per tijdsperiode te modelleren

4.1 Het homogene Poisson proces

Zoals vaak in de literatuur beschreven wordt er bij callcentra uitgegaan van een homogeen Poisson proces met intensiteit λ als het gaat om de aankomsten. Bij een homogeen Poisson proces gaat men uit van één vaste waarde met betrekking tot de intensiteit λ .

In figuur 1 is het gemiddeld aantal reguliere (PS in de dataset) aankomsten per tijdsperiode getoond over de maanden van 1999.

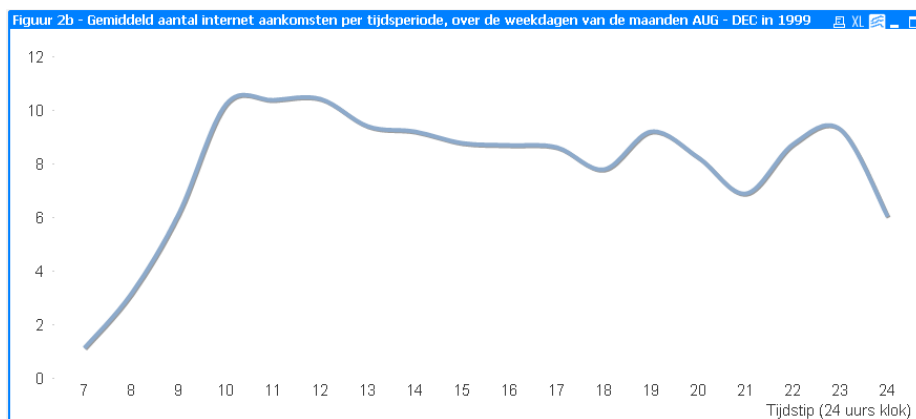
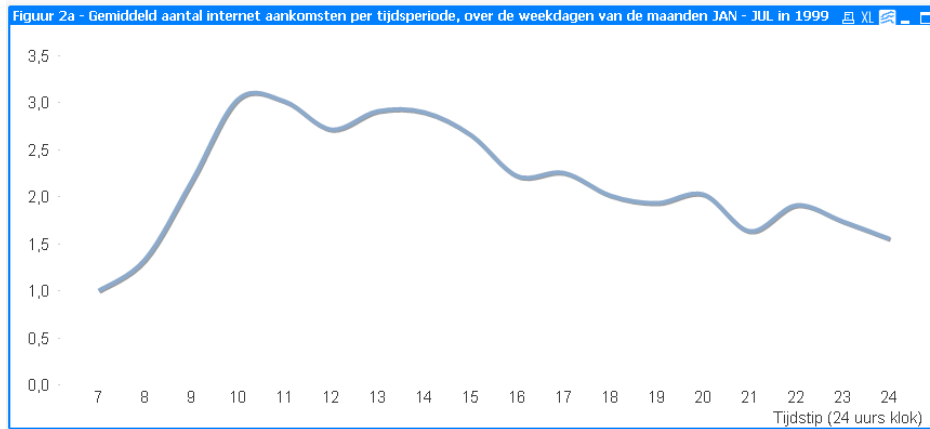


Duidelijk zijn de twee meest voorkomende pieken rond elf en drie uur. In de literatuur is dit vaak beschreven als een Bimodal patroon. Daarbij is het duidelijk dat dit geen simpel homogeen Poisson proces is. In dat geval zou over de tijdsperiodes heen de gemiddelde aankomstintensiteit ongeveer gelijk moeten blijven en dat is hier duidelijk niet het geval.

Verder is bij alle maanden, en met name in de maand mei, een duidelijke piek te zien rond tien uur 's avonds. Dit is misschien te verklaren uit het feit dat dit ongeveer het laatste tijdstip is dat mensen nog even bellen en men daarna van de avond gaat genieten.

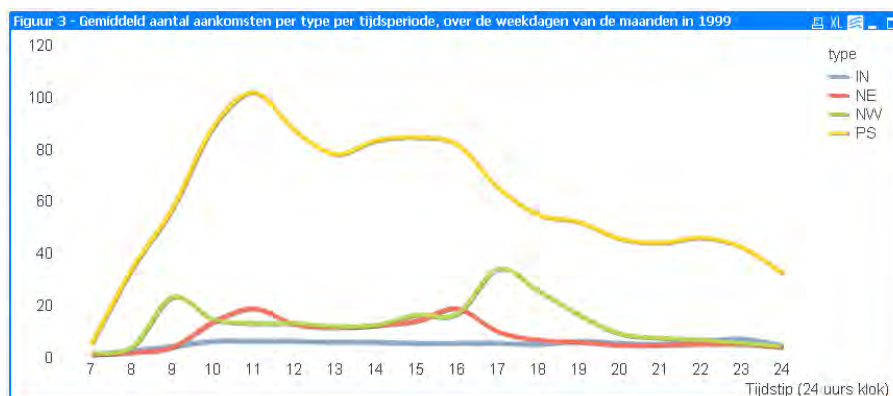
Om verder te onderzoeken of er echt geen sprake is van een homogeen Poisson proces zijn bepaalde situaties verder onderzocht. In de maand augustus heeft er namelijk een verandering plaatsgevonden zoals beschreven in het vorige hoofdstuk. Door deze verandering in het callcenter werden internet aankomsten (IN) door speciale IN agents behandeld en niet meer door agents die alle aankomsten behandelden.

In figuur 2a is te zien hoe de aankomstintensiteit van internet aankomsten er uit zag in de maanden JAN – JUL. Figuur 2b toont precies dezelfde intensiteit maar dan over de maanden AUG – DEC, en dus over der periode na de verandering.



Duidelijk verschil is dat de aankomst intensiteit hoger ligt in de nieuwe situatie maar dat er nog geen duidelijk homogeen Poisson proces is te tonen over de dag heen. Dat de internet intensiteit hoger ligt komt waarschijnlijk doordat men eerder geholpen wordt omdat er in de nieuwe situatie aparte agents zijn voor internet aankomsten.

Tot slot is er nog in figuur 3 een overzicht gegeven van de verschillende typen aankomsten om te bestuderen of er dan wel sprake is van een homogeen Poisson proces.



Duidelijk is dat de PS (normale telefoontjes) aankomsten het meest schommelt en het minst lijkt op een homogeen Poisson proces. In de volgende sectie zal er ingegaan worden op deze PS aankomsten.

4.2 Het tijd in-homogene Poisson proces

In de vorige paragraaf is naar voren gekomen dat de in de literatuur beschreven homogeen Poisson proces met betrekking tot de aankomst intensiteit niet van toepassing is op de gebruikte dataset. Een verbetering op het hiervoor beschreven homogene Poisson aankomstproces is het gebruiken van een tijdsafhankelijke λ . Dit zogenaamde tijd in-homogeen Poisson proces houdt in dat in kleine stukjes interval de λ bepaald zal worden. Hierdoor past de aankomstintensiteit in het model zich als het ware aan aan de specifieke tijdsperiode.

Om dit aan te tonen zal er een test uitgevoerd moeten worden die dit kan valideren. Deze test zal uitwijzen of in dit geval een bepaalde nulhypothese van toepassing is of dat deze wordt verworpen. Voor deze test is de volgende nulhypothese gedefinieerd:

De verschillende typen aankomsten vormen een in-homogeen Poisson proces met een stuksgewijs constante intensiteit.

Om de test, die straks zal worden beschreven, te kunnen gebruiken is gebruik gemaakt van een specifiek algoritme. Dit stukje algoritme is te vinden in de bijlagen.

De eerste stap in dit algoritme is het opdelen van een dag in kleine gelijke, met tijds lengte L , stukjes. Dit moet zo klein dat binnen een tijdsblokje er geen grote verandering in de aankomstintensiteit kan plaatsvinden. Om dit te valideren zal het nodig zijn te testen op het homogeen zijn van aankomsten binnen deze tijdsblokjes.

Hiervoor is het volgende gedefinieerd:

Definieer T_{ij} als de j -de aankomsttijd in tijdsblok i , met $i = 1, \dots, I$.

Dus $T_{i1} \leq \dots \leq T_{iJ(i)}$, met $J(i)$ het totaal aantal aankomsten in tijdsblok i .

Nu definiëren we $T_{i0} = 0$ en $R_{ij} = (J(i) + 1 - j) \left(-\log \left(\frac{L - T_{ij}}{L - T_{i,j-1}} \right) \right)$, met $j = 1, \dots, J(i)$

Onder de nulhypothese dat de aankomstintensiteit constant is in elk tijdsblokje, de $\{R_{ij}\}$ zullen onafhankelijke standaard exponentiële variabelen zijn. Wat ons dus nog rest is te testen of dit valide is en of de nulhypothese dus geldt.

Valideren of de $\{R_{ij}\}$ standaard exponentiële variabelen zijn gebeurt aan de hand van de zogenaamde Kolmogorov-Smirnov test. De Kolmogorov-Smirnovtoets is een statistische toets gebaseerd op een maat voor het verschil in twee verdelingen. In de vorm voor één steekproef, is het een aanpassingstoets, waarmee onderzocht wordt of de verdeling waaruit de steekproef getrokken is, afwijkt van een bekende verdeling zoals de normale verdeling, de uniforme verdeling, de Poisson-verdeling, de exponentiële verdeling, e.d. In de vorm voor twee steekproeven wordt nagegaan of de verdelingen waaruit de steekproeven afkomstig zijn, van elkaar verschillen. Wanneer we in ons geval onze experimenteel verkregen, call center, data vergelijken met random exponentieel getrokken data (met gemiddelde gelijk aan gemiddelde van experimenteel verkregen data), en de test verworpen de nulhypothese (beide vectoren aannemelijk uit eenzelfde verdeling) niet, dan kan gezegd worden dat de experimenteel verkregen data aannemelijk is gedistribueerd als een exponentiële verdeling.

Formeler, laat x en y twee vectoren zijn met respectievelijke stapsgewijze empirische distributiefuncties $F_x(c)$ en $F_y(c)$:

$H_0 := x$ en y hebben aannemelijk dezelfde verdeling

$H_1 := x$ en y hebben NIET dezelfde verdeling

Toetsingsgrootheid : $\max_c | F_x(c) - F_y(c) |$ Na het uitvoeren van de Kolmogorov-Smirnov toets is het mogelijk ook grafisch nog eens te checken of we te maken hebben met aankomsten van een exponentiële verdeling. Dit zal gebeuren aan de hand van een QQ-plot van de data van het callcenter tegenover een lid van de exponentiële familie. Een QQ-plot of quantile-quantile plot is een grafische techniek voor het checken of twee datasets uit dezelfde populatie komen met een gemeenschappelijke distributie. Het plot de kwantielen van de eerste dataset tegen de kwantielen van de tweede dataset. Als de twee datasets inderdaad uit dezelfde populatie komen met een gemeenschappelijke distributie dan zal een "rechte" lijn te zien zijn van 45 graden. Des te groter de afwijking van een rechte lijn van 45 graden de te kleiner de kans dat de dataset uit dezelfde populatie komen met een gemeenschappelijke distributie.

In de volgende secties zal ingegaan worden op verschillende selecties van aankomsttijden die interessant zijn om te onderzoeken. Bij elke selectie is gebruik gemaakt van de in bijlagen beschreven methode, de Kolmogorov-Smirnov toets en een QQ-plot om te controleren of de nulhypothese voor de desbetreffende selectie geldt.

4.2.1 PS aankomsten tussen 07:00 en 24:00 per maand.

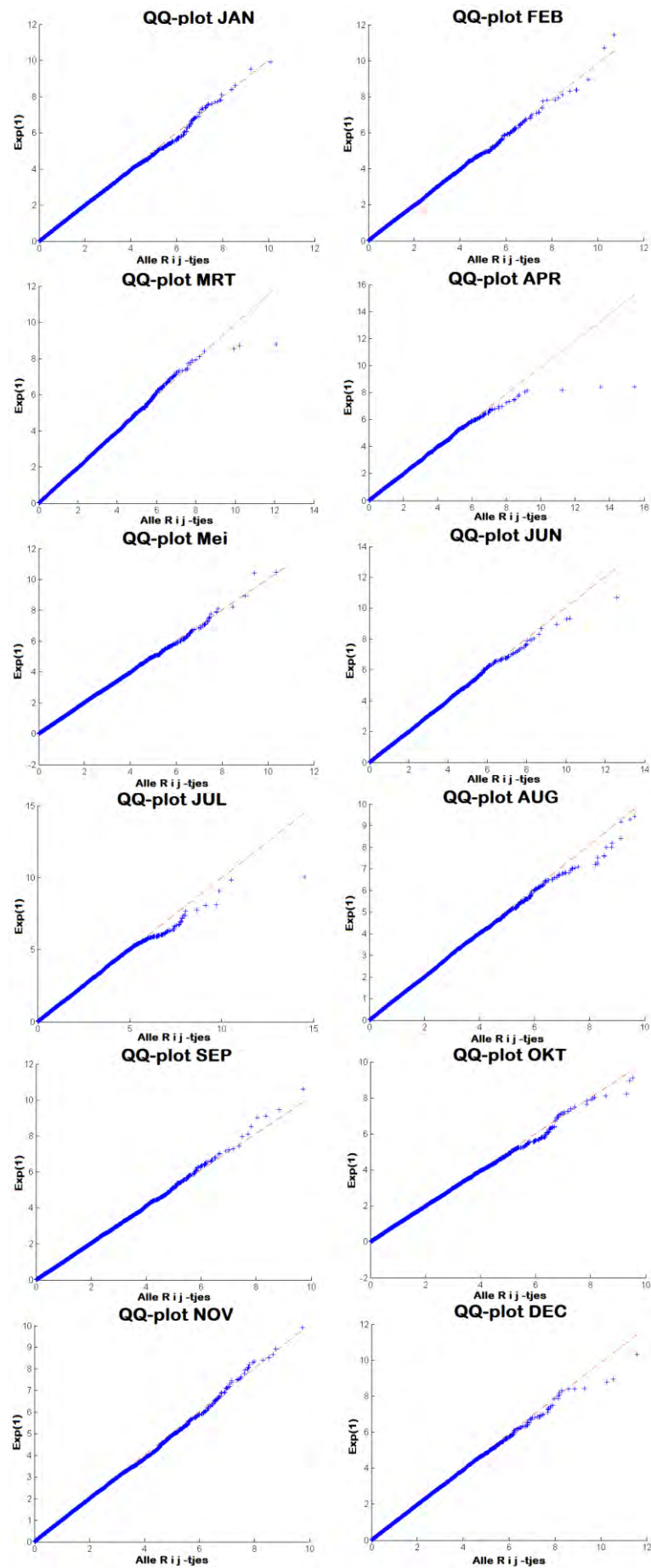
Als eerst is het interessant om te kijken hoe de verschillende maanden onderling verschillen. Met name door de veranderingen binnen het callcenter is het interessant te kijken of dit verschil oplevert tussen de maanden. In tabel 1 is de uitkomst van de Kolmogorov-Smirnov toets gegeven van de verschillende maanden.

Maand	H	P	K	Nulhypothese
Januari	0	0.9801	0.0033	Niet verworpen
Februari	1	0.0406	0.0100	Verworpen
Maart	0	0.9210	0.0036	Niet verworpen
April	0	0.3095	0.0069	Niet verworpen
Mei	0	0.76520	0.0041	Niet verworpen
Juni	0	0.7863	0.0042	Niet verworpen
Juli	0	0.5471	0.0051	Niet verworpen
Augustus	0	0.1132	0.0075	Niet verworpen
September	0	0.4817	0.0061	Niet verworpen
Oktober	0	0.1353	0.0081	Niet verworpen
November	0	0.1748	0.0070	Niet verworpen
December	0	0.3510	0.0058	Niet verworpen

Tabel 1 : Kolmogorov-Smirnov toets per maand

Opvallend is te zien dat in de maand februari de nulhypothese net aan is verworpen. De rede hiervan is niet duidelijk aangezien er geen rare verandering heeft plaatsgevonden en er niets bekend is over bijvoorbeeld een vakantieperiode.

In het volgende figuur is een overzicht met alle QQ-plots te vinden van elke maand. Opvallend is dat hier geen significante afwijking in de maand februari is te ontdekken.



Al met al blijkt dus dat wanneer er gekeken wordt naar alle PS aankomsten tussen 07:00 en 24:00 en dit verdelen in tijdsperiodes van 6 minuten dat de nulhypothese bij bijna alle maanden geldt.

De PS aankomsten vormen een in-homogeen Poisson proces met een stuksgewijs constante intensiteit.

4.2.2 PS aankomsten `s morgens, `s middags en `s avonds in januari-december.

Een tweede interessante kant om te onderzoeken is een situatie waar gekeken wordt naar de verschillende tijdstippen van de dag. Hiervoor zullen de PS aankomsten per ochtend, middag en avond interessant zijn om te onderzoeken. In deze situatie zal er uitgegaan worden van de volgende dagdelen:

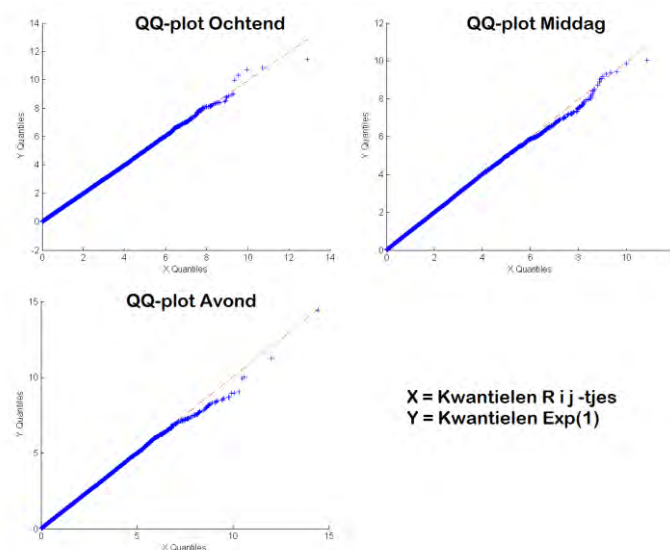
- Ochtend: 07:00 – 13:00 = 6 uur
- Middag: 13:00 – 18:00 = 5 uur
- Avond: 18:00 – 24:00 = 6 uur

In tabel 2 is per dagdeel de Kolmogorov-Smirnov toets toegepast. Ook hieruit blijkt dat wederom de nulhypothese niet verworpen kan worden.

Uur	H	P	K	Nulhypothese
07:00-12:00	0	0,9225	0.0016	Niet verworpen
12:00-19:00	0	0.6270	0.0025	Niet verworpen
19:00-24:00	0	0.6980	0,0028	Niet verworpen

Tabel 2: Kolmogorov-Smirnov toets per dagdeel

In het figuur hieronder zijn alle QQ-plots weergegeven en ook daar is duidelijk te zien dat wanneer er gekeken wordt naar het dagdeel dat wederom de nulhypothese geldt.



4.2.3 PS aankomsten tussen 07:00 en 24:00 in de maanden november en december

De aankomsten in de maanden november en december zijn interessant om te onderzoeken omdat dan de veranderingen zijn geweest zoals beschreven in hoofdstuk 3.

Uit de Kolmogorov-Smirnov toets kwam de volgende uitkomst.

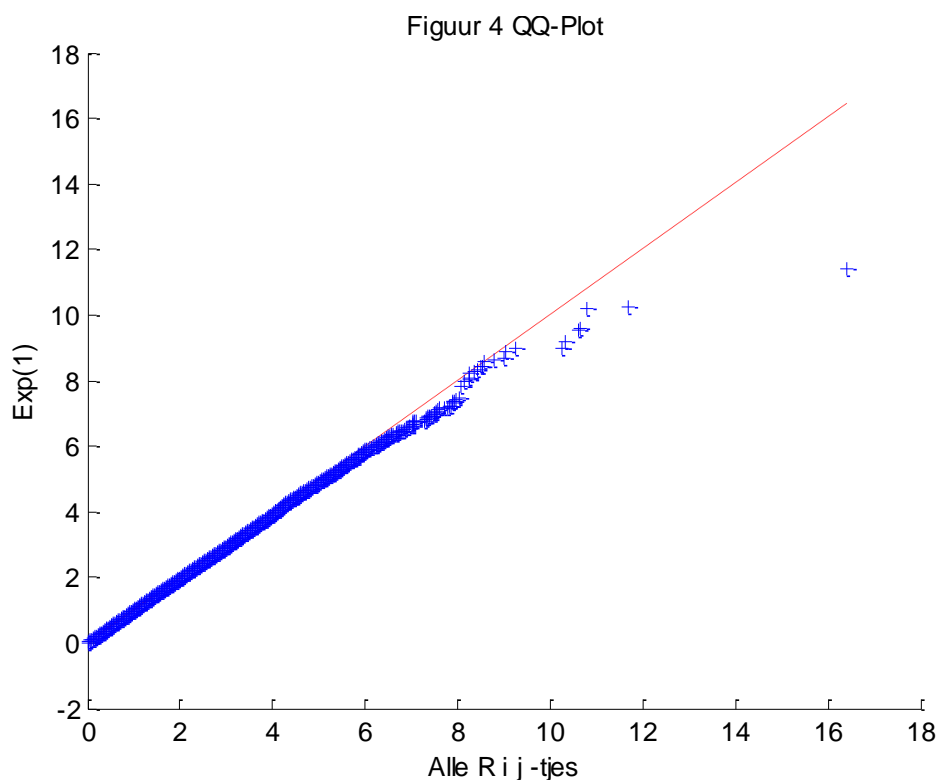
$$h = 0$$

$$p = 0.4184$$

$$k = 0.0039$$

Hierdoor is de nulhypothese niet verworpen en kan nog steeds uitgegaan worden van aankomsten die exponentieel verdeeld zijn.

In figuur 4 is de QQ-plot getoond van de selectie tegen een dataset uit een populatie met exponentiële verdeling.



Zowel uit de Kolmogorov-Smirnov toets en de QQ-plot is er geen rede om de nulhypothese, dat de PS aankomsten in november en december tussen 07:00 en 24:00 een inhomogeen Poisson proces vormen, te verwerpen. Ondanks de verandering kan men dus wederom uitgaan van een inhomogeen Poisson proces.

4.2.4 IN aankomsten tussen 07:00 en 24:00 in de maanden november en december

Omdat de IN aankomsten veel minder vaak voorkomen is in dit geval gekozen voor tijdsblokjes van 60 minuten. Deze aankomsten zijn ook interessant om te onderzoeken in november en december omdat ook hier een verandering heeft plaatsgevonden. En we de nieuwe situatie zoals die nu is willen bekijken.

In deze situatie blijkt uit de Kolmogorov-Smirnov toets het volgende:

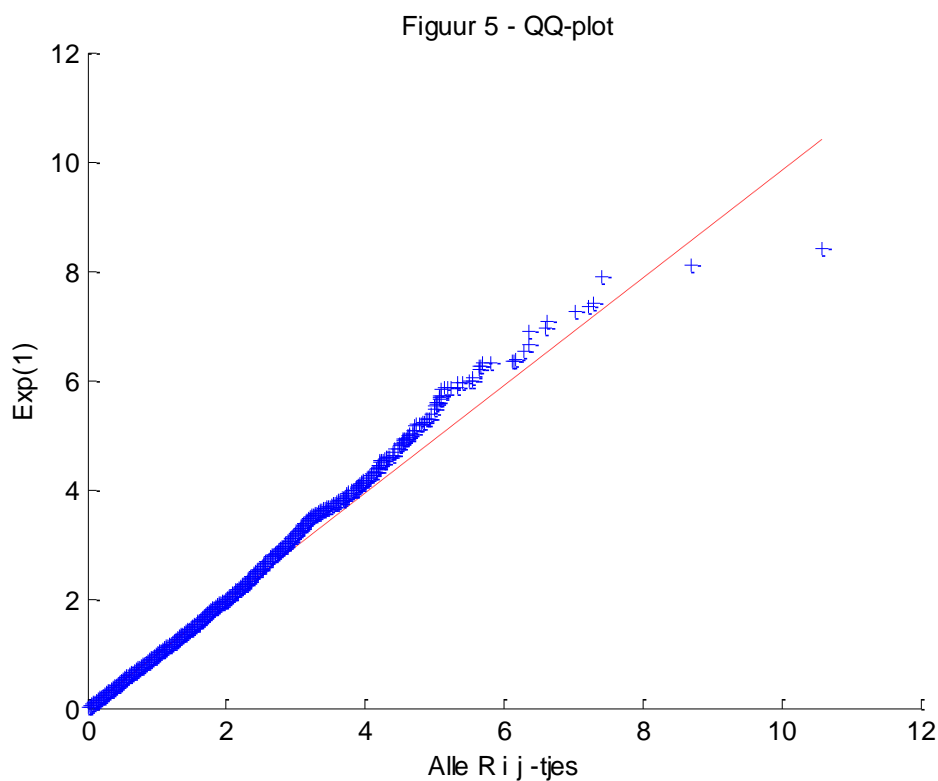
$$h = 0$$

$$p = 0.8089$$

$$k = 0.0079$$

Wederom is de nulhypothese niet verworpen met deze toets.

In figuur 5 is de QQ-plot getoond van de selectie tegen een dataset uit een populatie met exponentiële verdeling.



In het geval van IN aankomsten is er wederom spraken van de nulhypothese. Dit betekend dat met grote zekerheid de IN aankomsten exponentieel verdeeld zijn met deze manier van werken volgens het inhomogene Poisson proces.

5 Het serviceproces

Het service proces is het belangrijkste proces voor een callcenter. Het doel van het contact vinden met een callcenter is namelijk het gebruik maken van de service van het callcenter.

In tabel 3 is een overzicht gegeven van het gemiddelde, de standaard deviatie en de mediaan van de servicetijd. De aankomsten waarbij de service langer dan één uur inhiel is hier buiten beschouwing gelaten. Het is onrealistisch deze mee te nemen aangezien servicetijden langer dan één uur wel erg lang zouden zijn.

IN aankomsten hebben gemiddeld de langste servicetijd en daarna de NE aankomsten. Potentiële klanten, oftewel de NW aankomsten, hebben gemiddeld de kortste service tijd.

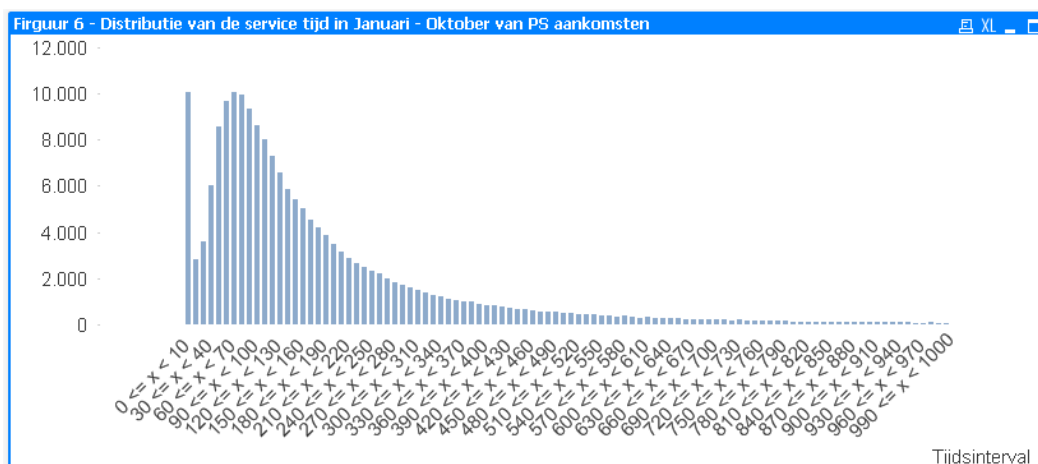
	Reguliere Aankomsten (PS)	Potentiële klant (NW)	Internet consult (IN)	Aandelen transactie (NE)
Gemiddelde	177,64	114,28	399,87	268,72
SD	189,46	145,69	472,82	303,04
Mediaan	119	72	220	175

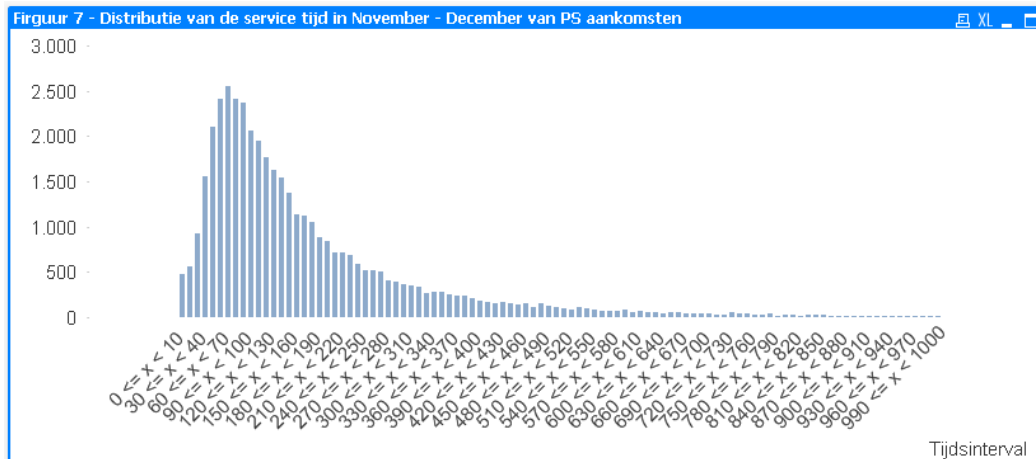
Tabel 3: Servicetijd gegevens in november en december ($t > 0$ en $t \leq 3600$)

Dat potentiële nieuwe klanten gemiddeld de kortste servicetijd hebben zou kunnen uit het feit dat dit type telefoontje een standaard behandeling vraagt. En er dus al bekend is bij het callcenter wat er stap voor stap moet worden gedaan. Internet en aandelen transacties kunnen per keer erg verschillen qua behandeling en daarom zou een gemiddeld langere servicetijd niet uitzonderlijk kunnen zijn.

5.1 Korte servicetijden

Uit de data van het callcenter blijkt dat PS aankomsten het meest voorkomen. In deze paragraaf zal dan ook deze groep uitgebreid bekeken worden. In figuur 6 is een histogram weergegeven van alle PS servicetijden samen in januari tot en met oktober. Eenzelfde histogram voor de maanden november en december is te zien in figuur 7.





Doormiddel van deze twee histogrammen is duidelijk te zien dat in de maanden november en december het percentage korte servicetijden, die korter zijn dan tien seconden, veel lager is dan die van de maanden er voor. Servicetijden korter dan tien seconden zijn erg onrealistisch aangezien dan weinig tot niets besproken zou kunnen zijn.

En inderdaad, uit gegevens van het callcenter blijkt dat sommige agents in de maanden januari tot en met oktober expres snel op hingen zodat zij van wat extra pauze konden genieten. Omdat de manager van het callcenter hierachter kwam is er in oktober een waarschuwing naar de agents gestuurd. Deze waarschuwing heeft kennelijk geleid tot het gewenste resultaat en in de rest van dit hoofdstuk zal dan ook alleen gekeken worden naar deze twee maanden.

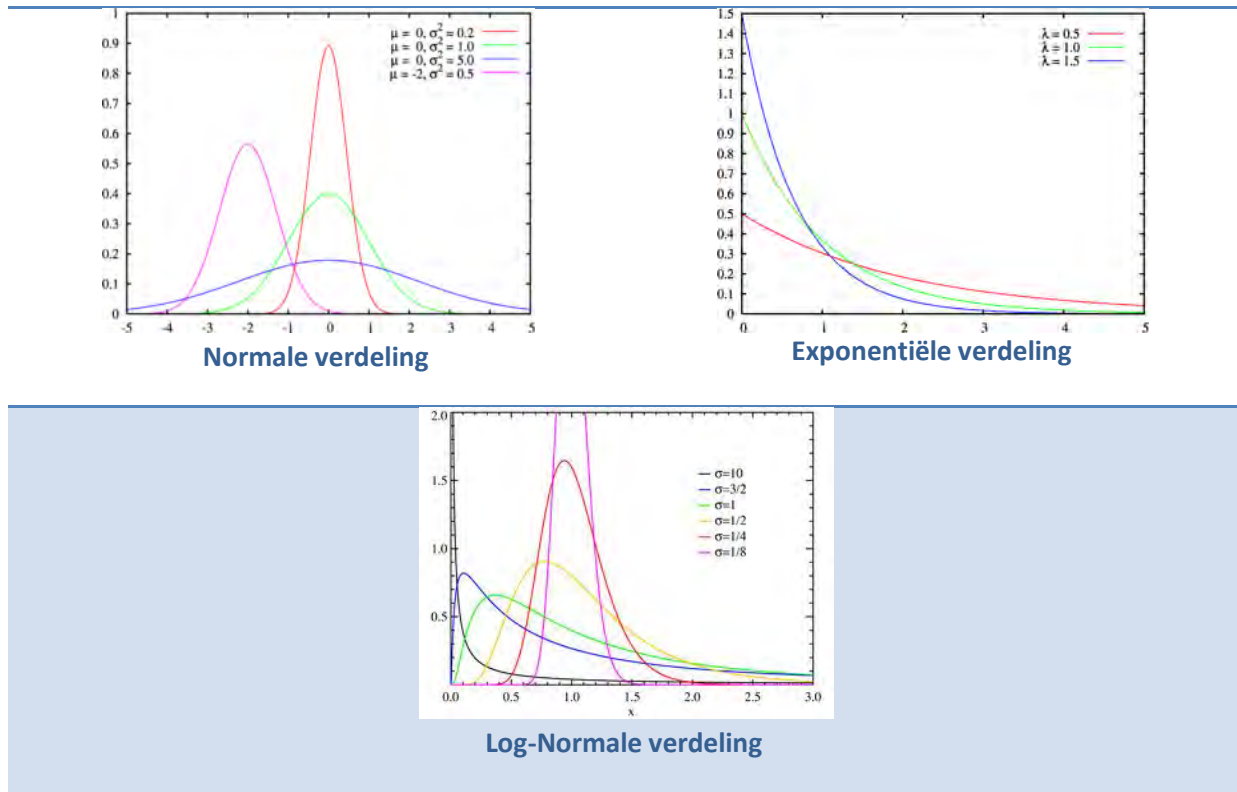
5.2 De distributie van de servicetijden in November en December

In de standaard literatuur over de servicetijden bij de wachtrij theorie gaat men vaak uit van exponentieel verdeelde servicetijden. Echter kijkend naar de vorige twee figuren dan is duidelijk te zien dat het hier niet gaat om zogenaamde exponentieel verdeelde servicetijden. De vraag is nu:

Van welke distributie is met betrekking tot de servicetijd dan wel sprake?

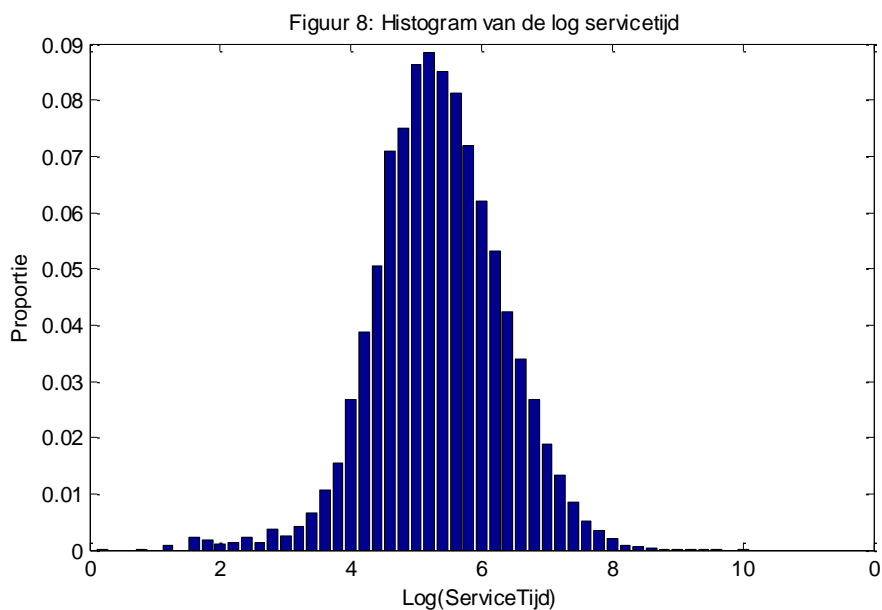
Voor het vinden van de best passende distributie is als eerst grafisch onderzoek van belang. Hierdoor zullen bepaalde distributies kunnen worden uitgesloten en de juiste distributie misschien al aan het licht komen. Tot slot kunnen we dit valideren aan de hand van de in hoofdstuk 4 beschreven Q-Q plot en Kolmogorov-Smirnov toets.

In de tabel 4 op de volgende pagina zijn drie distributies getoond die zullen worden vergeleken met de hiervoor getoonde figuur 7.



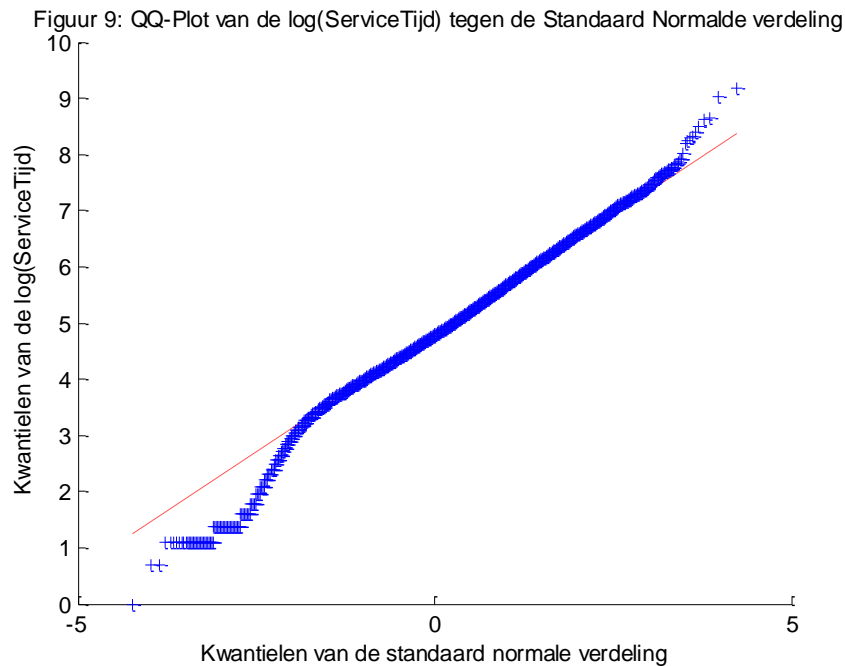
Tabel 4: Tabel met de PDF's van drie verdelingen.

Duidelijk is de dat de eerste verdeling, de normale verdeling, niet goed past bij de hiervoor getoonde figuur 7. Nog duidelijker is dat de exponentiele verdeling totaal niet past bij figuur 7. Echter de laatste in de tabel, de log-normale verdeling, heeft een redelijk gelijke vorm in vergelijking met figuur 7. Verder onderzoek met betrekking tot de log-normale verdeling zal nu worden beschreven.



In figuur 8 is een histogram te zien van de log(service tijd) voor de maanden november en december. Duidelijk is natuurlijk de vorm van een normale verdeling. Om dit kracht bij te zetten is op de volgende pagina ter validatie een QQ-plot en de Kolmogorov-Smirnov toets beschreven.

In figuur 9 is een QQ-plot te zien van de $\log(\text{service tijd})$ tegen de normale verdeling. Duidelijk is hier de bulk te zien in een rechte lijn. Dit geeft al aan dat met grote waarschijnlijkheid hier sprake is van service tijden die log-normaal verdeeld zijn.



Tot slot blijkt uit de Kolmogorov-Smirnov toets dat de nulhypothese, dat de $\log(\text{service tijden})$ normaal verdeeld zijn, wel wordt verworpen. De nulhypothese wordt verworpen omdat de uitkomst als volgt is:

$$H = 1$$

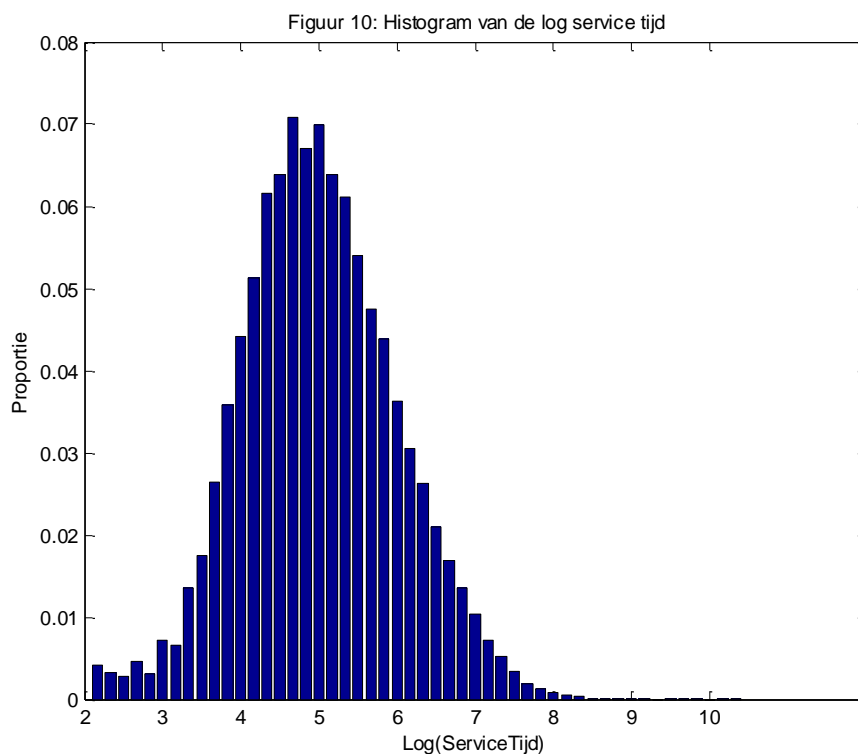
$$P < 0.05$$

Dit betekent dat het niet 100% procent zeker is of de servicetijden log-normaal verdeeld zijn ondanks de duidelijke grafische bewijzen. Hieruit is dus te concluderen dat de distributie van de service tijden in november en december erg dicht tegen de distributie ligt van de log-normale verdeling maar dit niet exact te bewijzen is.

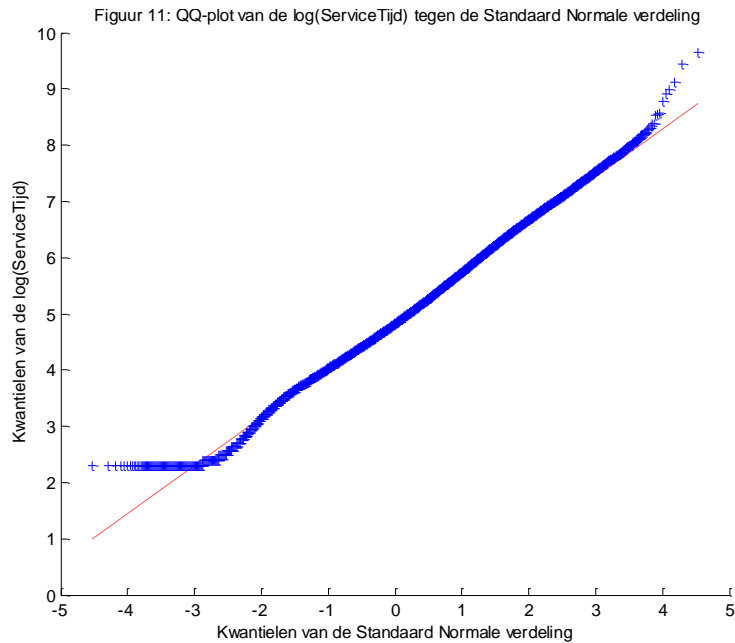
5.3 De distributie van de servicetijden in de overige maanden

Uit de vorige paragraaf blijkt dat de distributie van de servicetijden in november en december dicht tegen de distributie ligt van de log-normale verdeling. Als eerst zijn deze twee maanden onderzocht omdat hierin er geen ruis is van het zogenaamde “ophang gedrag” van de agents tijdens de eerste tien seconden, zoals beschreven in paragraaf 5.1.

Hiervan uitgaande zou de distributie van de overige maanden ook dicht tegen de distributie van de log-normale verdeling moeten liggen wanneer hierin het “ophang gedrag” wordt weggelaten. In deze paragraaf zal dan ook kort onderzoek worden gedaan naar de maanden januari tot en met oktober waarbij servicetijden korter dan tien seconden zullen worden weggelaten.



In figuur 10 is het histogram getoond van de maanden januari tot en met oktober waarbij de servicetijden korter dan 10 seconden zijn weggelagen. Ook dit histogram toont al een vorm van een normale verdeling waardoor het mogelijk is dat ook in dit geval de service tijden log-normaal verdeeld zijn.



Ook de QQ-plot, in figuur 11, voor de maanden januari tot en met oktober tonen een redelijk mooie rechte lijn. Dit geeft al aan dat met grote waarschijnlijkheid hier sprake is van service tijden die log-normaal verdeeld zijn.

Voor een goed eindoordeel zal gekeken moeten worden naar de uitkomst van de Kolmogorov-Smirnov toets. Wederom is in dit geval de nulhypothese, dat de log(service tijden) normaal verdeeld zijn, verworpen. Ook in dit geval is de nulhypothese verworpen omdat de uitkomst van de Kolmogorov-Smirnov toets als volgt is:

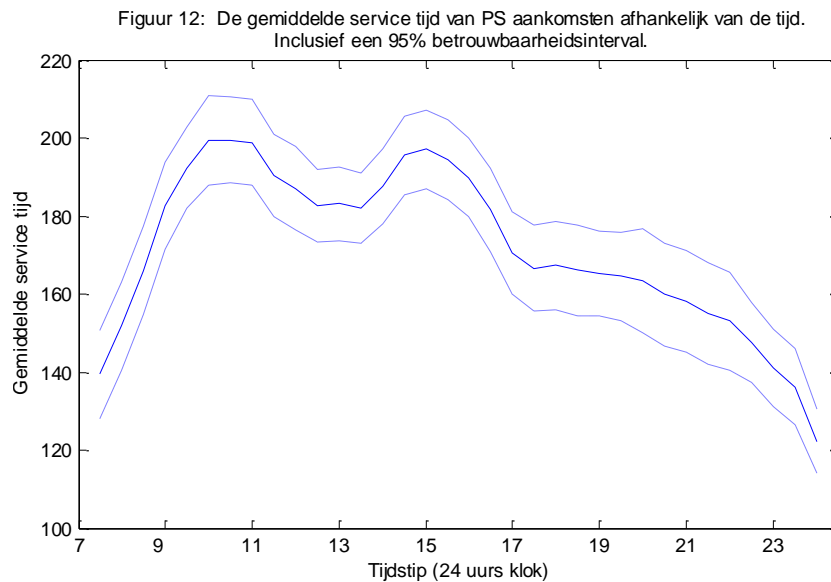
$$H = 1$$

$$P < 0.05$$

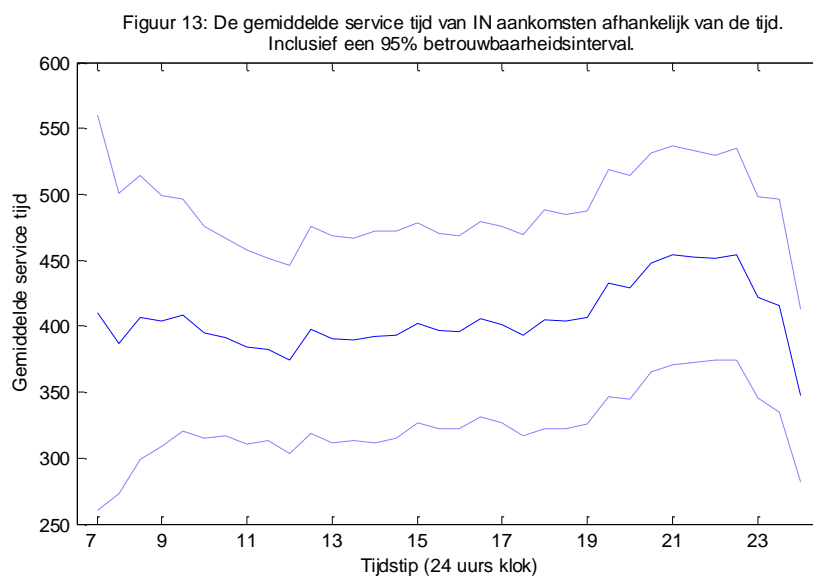
Dit betekent dat het ook voor de maanden januari tot en met oktober niet 100% procent zeker is of de servicetijden log-normaal verdeeld zijn ondanks de duidelijke grafische bewijzen. Hieruit is te concluderen dat de distributie van de service tijden ook in deze maanden dicht tegen de distributie ligt van de log-normale verdeling maar niet exact te bewijzen is.

5.4 De servicetijd afhankelijk van de tijd

In de vorige paragraaf was te concluderen dat de service tijd over het algemeen log-normaal verdeeld is. Een exactere methode om de service tijd te bekijken zou zijn de gemiddelde service tijd per tijdsperiode te berekenen. Dus wat is de service tijd afhankelijk van de tijd. In deze paragraaf zal hier kort op worden ingegaan en zal aangetoond worden dat dit een preciezere manier is om de gemiddeld servicetijd te bekijken.



In figuur 12 is de gemiddelde service tijd voor PS aankomsten per tijdstip van de dag te zien inclusief een 95 procent betrouwbaarheidsinterval. Opvallend is dat het patroon vergelijkbaar is met het Bimodal patroon bij de aankomsten. Daarbij zijn de twee toegevoegde stippellijnen met betrekking tot het betrouwbaarheidsinterval bijna identiek qua vorm in vergelijking met de gemiddelde service tijd.



Voor de IN aankomsten is in figuur 13 de gemiddelde service tijd aangegeven per tijdstip met wederom twee lijnen met betrekking tot het betrouwbaarheidsinterval.

Opvallend is het grote verschil tussen de twee figuren. Bij de IN aankomsten is geen duidelijk verband met een Bimodal patroon en ligt de gemiddelde service tijd veel hoger als bij de PS aankomsten. Daarbij liggen de twee stippelijnen bij de IN aankomsten veel verder uit elkaar. Dit betekent dat de IN aankomsten over het algemeen een veel langere maar vooral fluctuerende service tijd hebben, zoals geconcludeerd aan het begin van dit hoofdstuk.

In beide gevallen is uitgegaan van dezelfde periodes, van een half uur, waarvoor de gemiddelde service tijd is berekend. Interessant is nu het onderzoeken van deze kleine stukjes periode. Zijn deze net zoals aan het begin van dit hoofdstuk redelijke log-normaal verdeeld of wordt wederom de nulhypothese verworpen?

In tabel 5 is per tijdsperiode de uitkomst van de Kolmogorov-Smirnov toets weergegeven.

Periode	H	P	K	Nulhypothese
7:00 – 7:30	1	< 0.05	0.9504	Verworpen
7:30 – 8:00	1	< 0.05	0.9471	Verworpen
8:00 – 8:30	1	< 0.05	0.9577	Verworpen
8:30 – 9:00	1	< 0.05	0.9627	Verworpen
9:00 – 9:30	1	< 0.05	0.9789	Verworpen
9:30 – 10:00	1	< 0.05	0.9677	Verworpen
10:00 – 10:30	1	< 0.05	0.9766	Verworpen
10:30 – 11:00	1	< 0.05	0.9763	Verworpen
11:00 – 11:30	1	< 0.05	0.9726	Verworpen
11:30 – 12:00	1	< 0.05	0.9774	Verworpen
12:00 – 12:30	1	< 0.05	0.9817	Verworpen
12:30 – 13:00	1	< 0.05	0.9794	Verworpen
13:00 – 13:30	1	< 0.05	0.9774	Verworpen
13:30 – 14:00	1	< 0.05	0.9776	Verworpen
14:00 – 14:30	1	< 0.05	0.9801	Verworpen
14:30 – 15:00	1	< 0.05	0.9784	Verworpen
15:00 – 15:30	1	< 0.05	0.9855	Verworpen
15:30 – 16:00	1	< 0.05	0.9767	Verworpen
16:00 – 16:30	1	< 0.05	0.9723	Verworpen
16:30 – 17:00	1	< 0.05	0.9678	Verworpen
17:00 – 17:30	1	< 0.05	0.9729	Verworpen
17:30 – 18:00	1	< 0.05	0.9701	Verworpen
18:00 – 18:30	1	< 0.05	0.9789	Verworpen
18:30 – 19:00	1	< 0.05	0.9786	Verworpen
19:00 – 19:30	1	< 0.05	0.9711	Verworpen
19:30 – 20:00	1	< 0.05	0.9688	Verworpen
20:00 – 20:30	1	< 0.05	0.9747	Verworpen
20:30 – 21:00	1	< 0.05	0.9618	Verworpen
21:00 – 21:30	1	< 0.05	0.9721	Verworpen
21:30 – 22:00	1	< 0.05	0.9605	Verworpen
22:00 – 22:30	1	< 0.05	0.9648	Verworpen
22:30 – 23:00	1	< 0.05	0.9665	Verworpen
23:00 – 23:30	1	< 0.05	0.9607	Verworpen
23:30 – 24:00	1	< 0.05	0.9739	Verworpen

Tabel 5: Kolmogorov-Smirnov toets per tijdsperiode met PS-aankomsten

Wederom blijkt uit de Kolmogorov-Smirnov toets de servicetijd per tijdsperiode niet exact log-normaal verdeeld is. Het is dus alleen via grafische methodes mogelijk te concluderen dat de servicetijd log-normaal verdeeld is.

Al met al is er dus nog niet een 100% waterdicht bewijs over de verdeling van de servicetijd. Echter de in de literatuur beschreven exponentieel verdeelde servicetijd kan word beschreven als niet van toepassing bij dit callcenter.

6 Het wachtproces

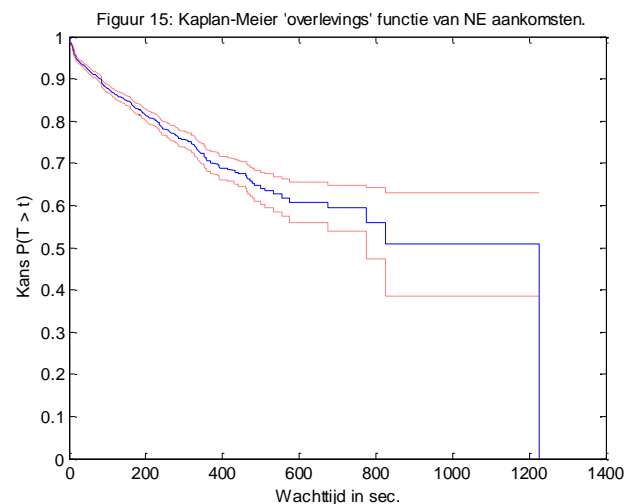
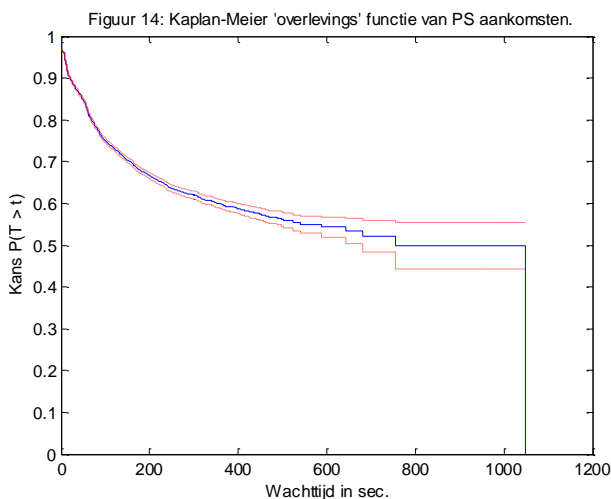
In de vorige hoofdstukken is ingegaan op het aankomst en het service proces. Beide processen zijn primaire processen binnen de wachtrij theorie. In dit laatste hoofdstuk zal ingegaan worden op een proces wat ook een grote rol speelt binnen de wachtrij theorie, namelijk het afhaakproces. Klanten die bellen naar een callcenter zullen in de meeste gevallen moeten wachten tot zij geholpen worden. Dit wachten is eigenlijk een afhaakproces. Want hoelang wil men wachten tot hij of zij geholpen wordt? Hoeveel tijd of eigenlijk geduld heeft een klant voor een bepaalde service of dienst van het callcenter?

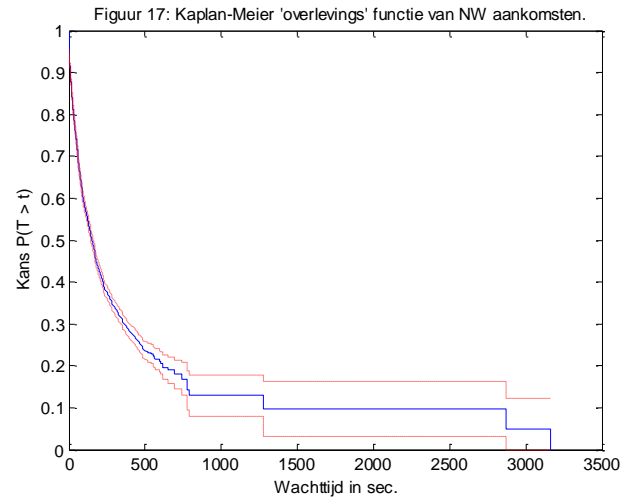
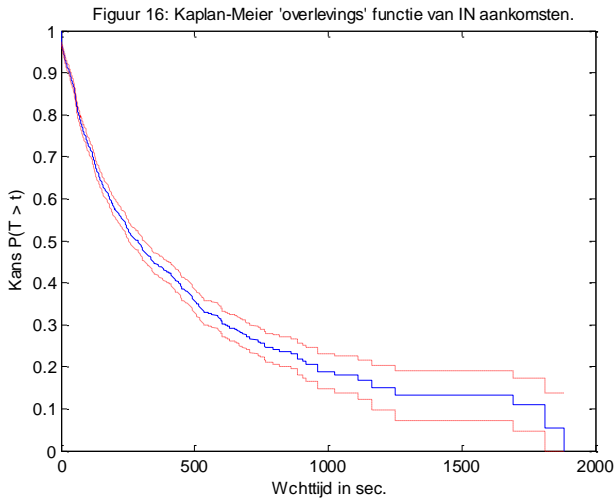
In de literatuur gaat men uit van twee typen wachttijden. Er is een wachttijd die een klant minstens moet wachten voordat hij of zij geholpen wordt en er is een wachttijd die de klant over heeft om geholpen te worden. De eerste noemt men ook wel de virtuele wachttijd en de tweede is ook wel het geduld van klanten.

Beide zijn duidelijk van grote invloed op het hele wachtrij proces. Echter zijn beide niet vaak beschreven in de literatuur en niet eenvoudig te berekenen. In dit hoofdstuk zal gekeken worden welke conclusies we kunnen trekken over de wachttijd en het geduld van klanten.

6.1 De invloed van afhakers op de gemiddelde wachttijd

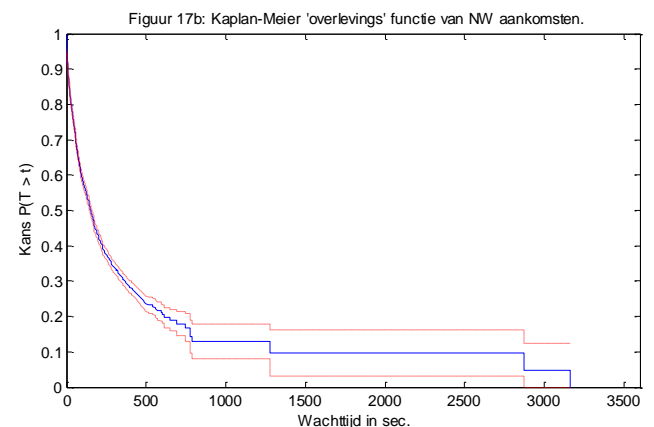
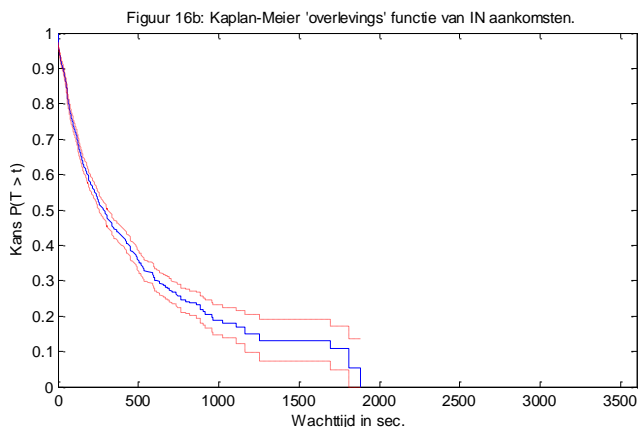
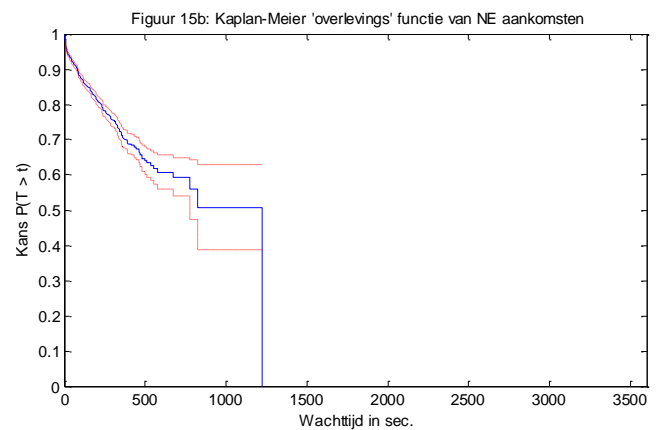
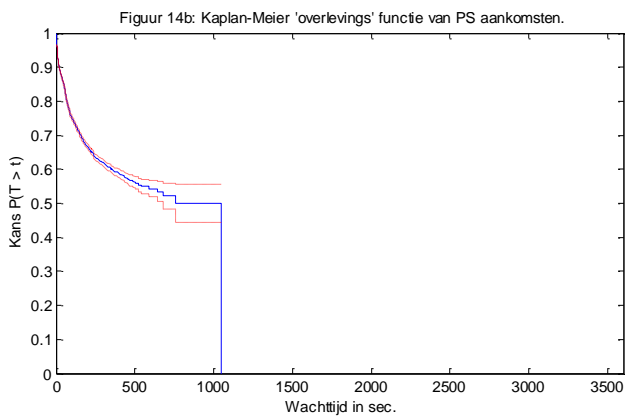
Onderzoek zal als eerst worden gedaan middels de standaard Kaplan-Meier product-limit toets. Deze toets is een vaak gebruikte toets in onderzoeken en daarom ook belangrijk om in dit onderzoek te gebruiken ter validatie. De Kaplan-Meier methode is een statistische methode om 'overlevingsdata' te analyseren. Overlevingsdata zijn niet noodzakelijk gerelateerd aan mortaliteit, maar kunnen ook gerelateerd zijn aan bijvoorbeeld een afhaakproces. Met deze methode is het mogelijk grafisch inzichtelijk te krijgen hoe groot de kans is dat een klant die belt een bepaalde tijd t wacht tot hij of zij wordt geholpen.





In de figuren 14 tot en met 17 zijn de belangrijkste typen aankomsten te zien met hun Kaplan-Meier 'overlevings' functie. Per figuur is getoond, middels de blauwe lijn, hoe groot de kans is dat een klant wacht tot hij of zij geholpen wordt dan wel afhaakt. De twee rode lijnen in elk figuur tonen daarbij de boven en ondergrens van het 95%-betrouwbaarheidsinterval.

Als eerst is opvallend dat het bereik op de x-as totaal verschillend is tussen de vier figuren. Om toch de vier verschillende typen aankomsten goed te kunne vergelijken is in de figuren 14b tot en met 17b de x-as gelijk genomen.



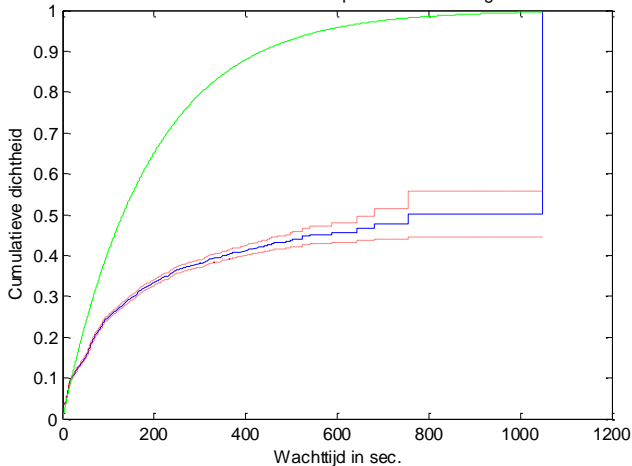
Uit de figuren 14b tot en met 17b valt als eerst op dat de reguliere klanten en de klanten voor aandelen transacties ongeveer dezelfde Kaplan-Meier 'overlevings' functie vertonen. Echter zijn de klanten voor aandelen transacties in het algemeen iets langer bereid te wachten. Dit zou te verklaren zijn uit het feit dat zij op dat moment echt hulp nodig hebben en niet terug willen bellen op een later moment.

Opvallend is ook de snelle daling en de lange staart van de Kaplan-Meier 'overlevings' functie van nieuwe klanten. Oftewel het grootse percentage van nieuwe klanten wacht niet lang en als hij of zij wel wacht is men bereid lang te wachten.

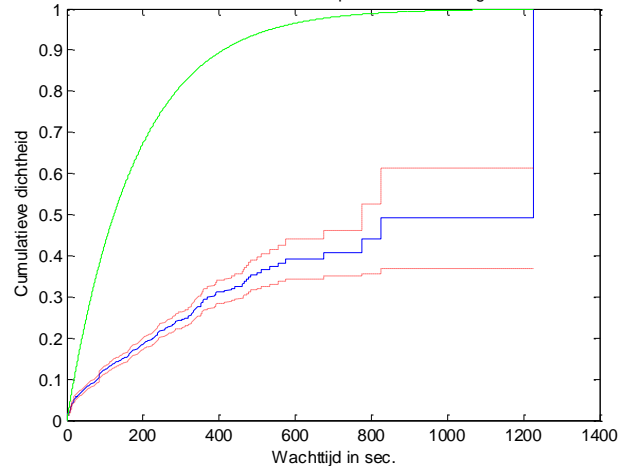
Tot slot kunnen we met de standaard Kaplan-Meier product-limit toets nog bekijken wat nou de invloed is van de ongeduldige klanten op de wachttijd. In de wachtrij theorie gaat men er namelijk van uit dat de wachttijd exponentieel verdeeld is. Echter zijn in die gevallen de randvoorwaarden dat er geen 'afhakkers' zijn binnen het systeem. Afhakkers zijn klanten waarbij de tijd van het geduld korter is dan de virtuele wachttijd. Deze klanten hangen dus op omdat zijn het te lang vinden duren voordat zij geholpen worden en dus ongeduldig worden.

Om de invloed van de afhakkers grafisch weer te geven zijn in de figuren 14c tot en met 17c middels de Kaplan-Meier toets de cumulatieve distributie functie van de wachttijd weergegeven samen met de bijbehorende cumulatieve distributie functie van de exponentiële verdeling uit de theorie.

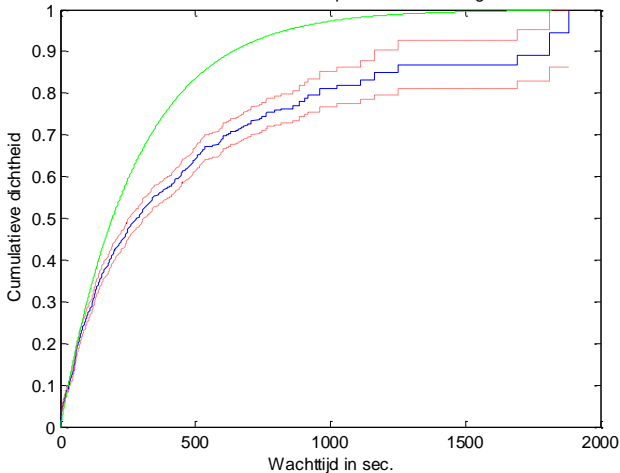
Figuur 14c: Kaplan-Meier 'overlevings' functie van PS aankomsten.
Inclusief CDF van Exponentiële verdeling.



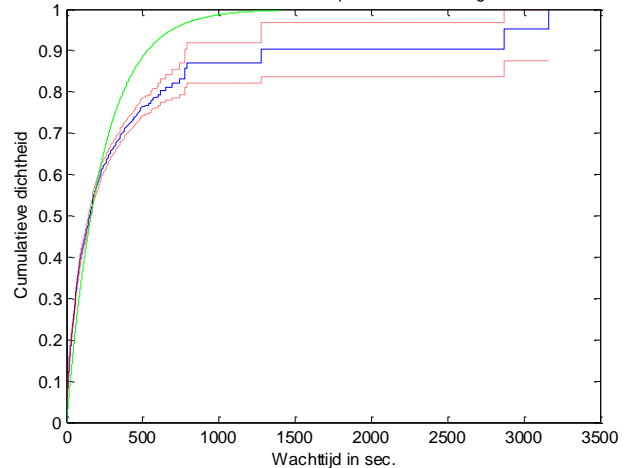
Figuur 15c: Kaplan-Meier 'overlevings' functie van NE aankomsten.
Inclusief CDF van Exponentiële verdeling.



Figuur 16c: Kaplan-Meier 'overlevings' functie van IN aankomsten.
Inclusief CDF van Exponentiële verdeling.



Figuur 16d: Kaplan-Meier 'overlevings' functie van NW aankomsten.
Inclusief CDF van Exponentiële verdeling.



Wederom is in de figuren 14c en 15c te zien dat de PS aankomsten redelijk gelijk liggen qua wachttijd proces in vergelijking met de NE aankomsten. Daarbij liggen ze beide behoorlijk ver van de exponentiële cumulatieve verdelingsfunctie, de groene lijn. Dit houdt in dat het vroegtijdig ophangen van klanten, ook wel afhakers, een grote rol speelt en dit dus de wachttijd gegevens ‘verstoord’. De figuren van de IN en NW aankomsten daarentegen vertonen een bijna zelfde vorm als de exponentiële cumulatieve verdelingsfunctie. Dit houdt in dat er meer sprake is van een , zoals in de literatuur beschreven, wachttijdsituatie zonder afhakers. De invloed van het kleine percentage afhakers is hier dus minder van invloed op het wachttijd proces.

6.2 Het geduld van klanten.

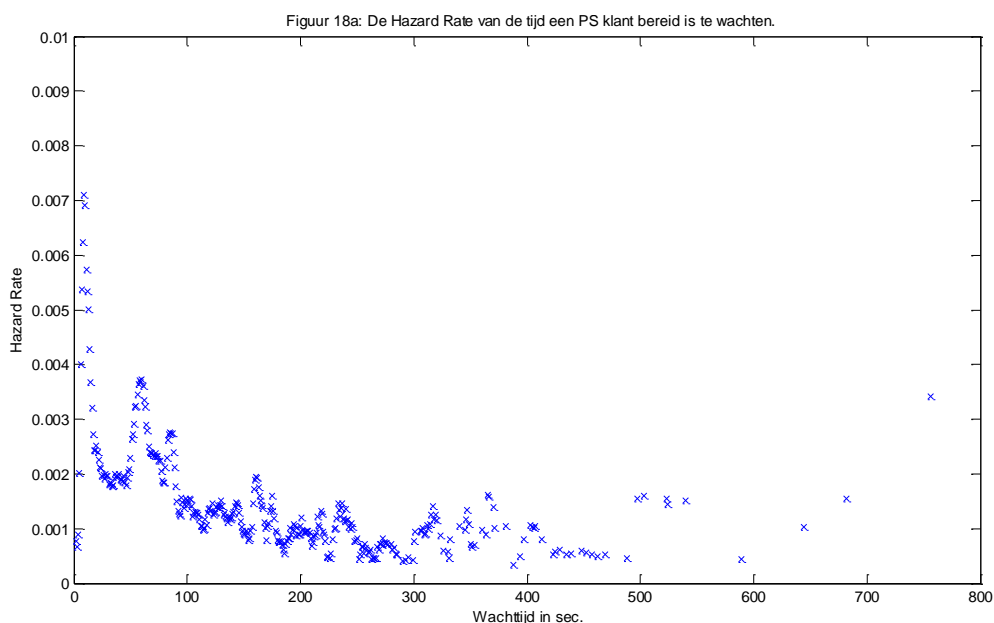
Zoals in de vorige paragraaf beschreven is het afhaakproces binnen de wachttijd theorie van grote invloed. Het afhaakproces is namelijk van invloed op de wachttijdsituatie zoals beschreven in de literatuur. Het verder bestuderen van dit speciale proces gaat in de literatuur aan de hand van de zogenaamde Hazard Rate.

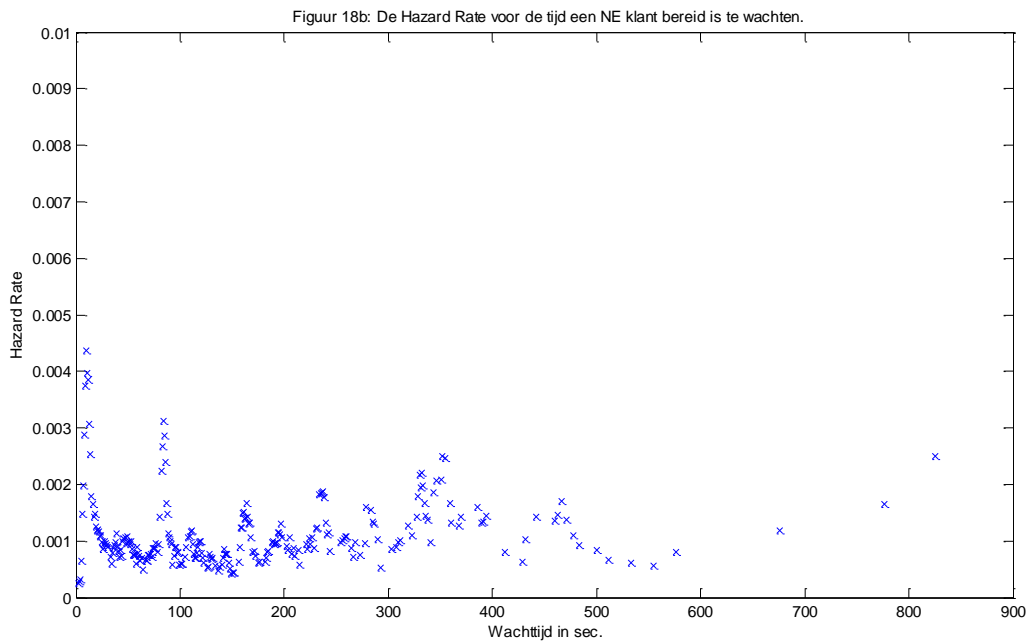
Het idee van de Hazard Rate komt uit de wereld van systeemonderhoud, waar het belangrijk is om te weten hoelang een onderdeel in het system het nog doet gegeven dat het op dat moment nog werkt. In het geval van een callcenter betekend dit dat men wil weten hoelang iemand nog aan de lijn blijft en dus niet ophang gegeven dat hij of zij nog niet opgehangen heeft. In deze paragraaf zal ingegaan worden op de klanten die ophangen. In de volgende paragraaf zal, middels de Hazard Rate, ingegaan worden op de klanten die niet hebben opgehangen en dus te maken hebben gehad met een virtuele wachttijd.

De Hazard Rate wordt berekend aan de hand van de volgende formule

$$\text{Hazard Rate} = \frac{f(x)}{(1-F(x))}$$

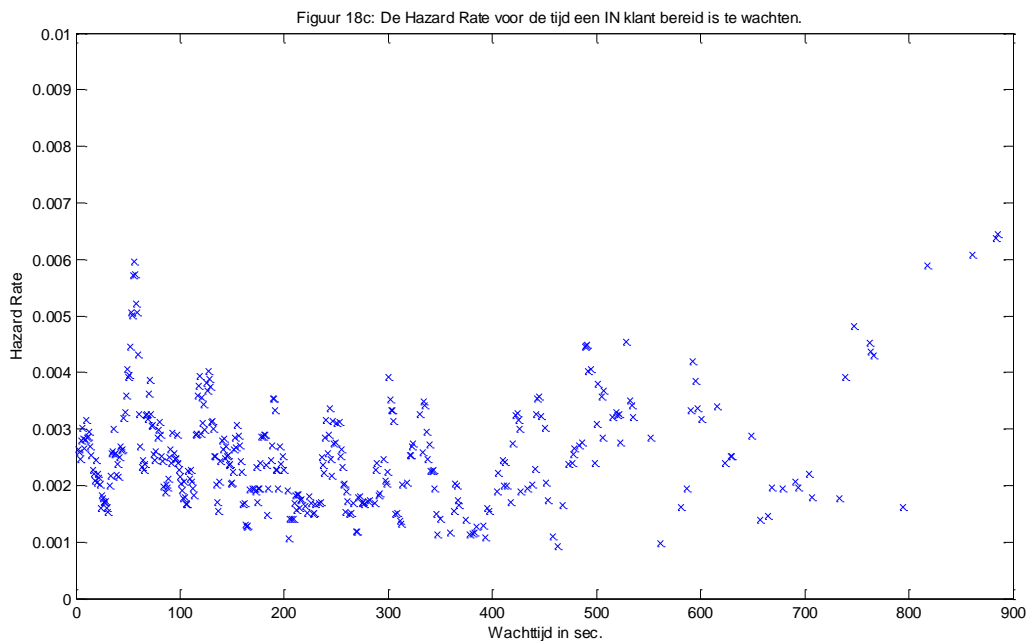
In figuur 18a en 18b is de Hazard Rate van respectievelijk de PS en IN klanten weergegeven. In de vorige paragrafen kwam naar voren dat deze twee typen aankomsten hetzelfde gedrag vertonen en ook met de Hazard Rate is te zien dat hier inderdaad ook sprake is van een vergelijkend beeld.

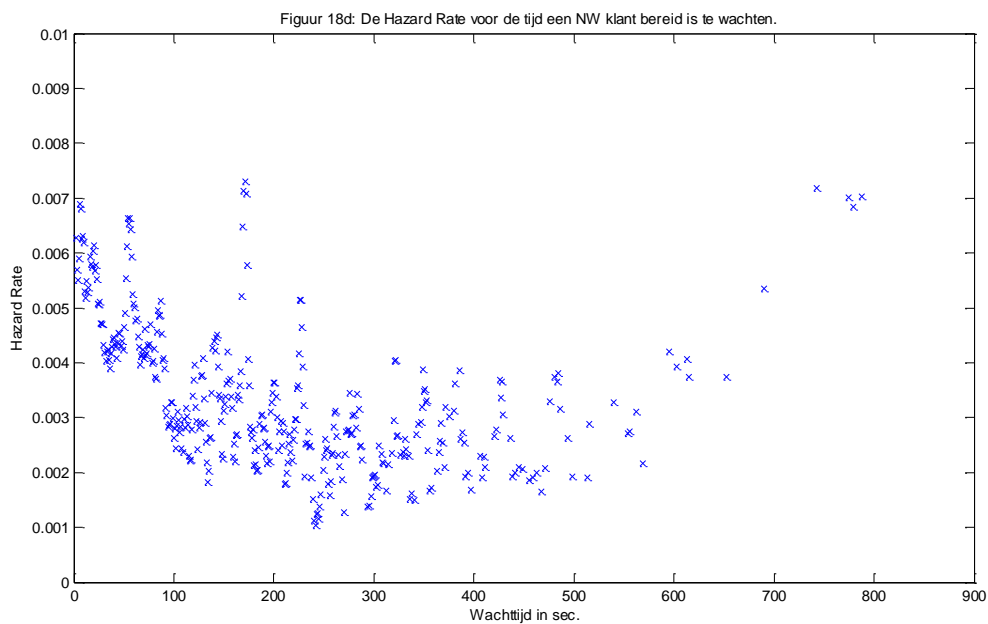




Beide figuren hebben twee pieken, namelijk eentje in het begin en een na ongeveer 60 seconden. De eerste piek is te verklaren uit het feit dat men ophangt aan het begin omdat er dan verteld wordt dat hij of zij in de wachtrij zit. De tweede piek is het vaak de herhalende mededeling dat hij of zij nog steeds in de wachtrij zit. Bij beide figuren is dan ook te zien dat steeds na ongeveer 60 seconden er een klein piekje is. Waarschijnlijk is de mededeling dat hij of zij in de wachtrij zit zo demotiverend dat de klant ophangt.

In de figuren 18c en 18d, respectievelijk die voor de internet hulp (IN) klanten en nieuwe (NW) klanten, is nog duidelijker het effect van het algemene wachtrij bericht te zien. Bij de vorige twee figuren was vooral aan het begin een hoge piek en werden de pieken daarop steeds minder. Bij de figuren van de IN en NW klanten is echter te zien dat het piek patroon behoorlijk lang doorgaat en dat de pieken twee keer zo hoog zijn als bij de vorige twee figuren.



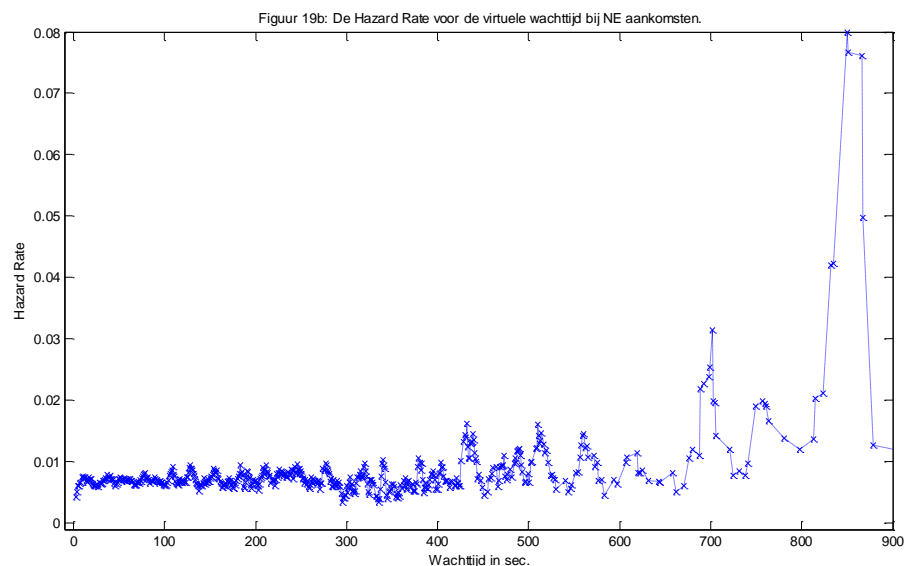
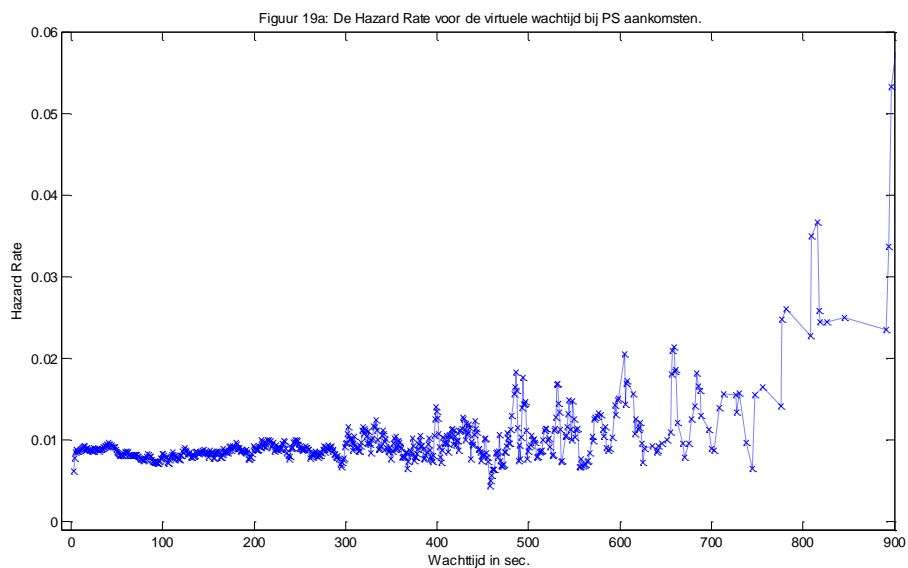


Wat uit de vorige figuren te concluderen valt is dat de invloed van een informatie bericht over de wachtrij de 'ophangkans' vergroot en dus van grote invloed is op de wachttijd gegevens. Dit zou kunnen omdat men, wanneer er om de een minuut een bericht te horen is, het idee heeft als lang aan de lijn te hangen. De klant hoort tenslotte vaak hetzelfde bericht dus zal hij of zij al lang aan de lijn hangen denk men. Een goede test voor het callcenter zou zijn het tijdsinterval tussen de berichten te vergroten en dan te bestuderen wat de gevolgen hiervan zijn. Als dit van goede invloed is op de situatie zal de theoretische benaderingen veel beter tot zijn recht komen.

6.3 De virtuele wachttijd

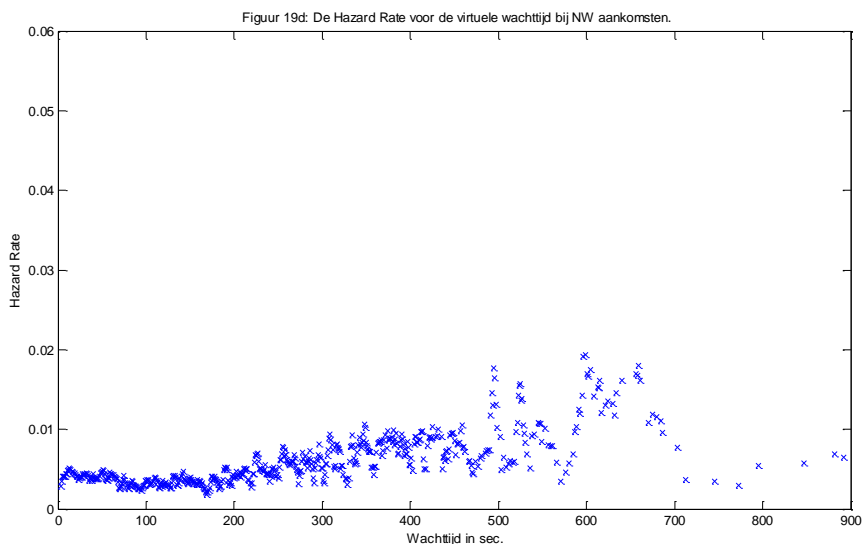
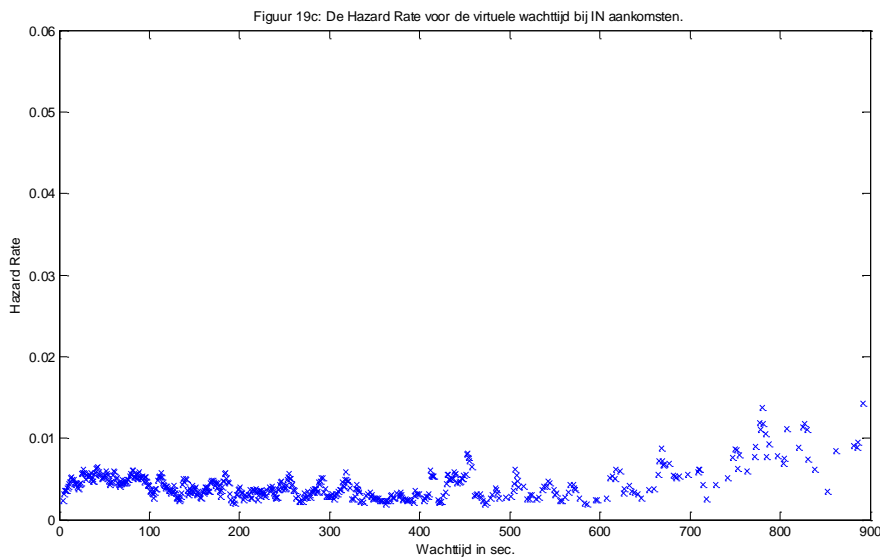
Als afsluiting van dit hoofdstuk zal in deze paragraaf de virtuele wachttijd bekeken worden. De virtuele wachttijd is, zoals aan het begin van dit hoofdstuk beschreven, de tijd een klant heeft gewacht in de wachtrij voordat hij of zij geholpen werd. De klant heeft dus niet opgehangen en is dus uiteindelijk geholpen.

Wederom is dit te bekijken via de Hazard Rate. In de figuren 19a en 19b is de Hazard Rate met betrekking tot de virtuele wachttijd weergegeven voor PS en NE aankomsten.



Duidelijk is dat de eerst vijf minuten de Hazard Rate behoorlijk constant is rond de 0.007. Pas naar vijf minuten beginnen de schommelingen steeds groter te worden. Met betrekking tot de virtuele wachttijd is zes minuten redelijk normaal in het geval van PS en NE aankomsten.

In figuren 19c en 19d is tot slot de Hazard Rate weergegeven van de IN en NW aankomsten.



In de figuren 19c en 19d is te zien dat de Hazard Rate lager ligt dan de vorige twee, namelijk rond de 0.004, en zijn de schommeling ook veel minder groot. Bij de IN aankomsten ligt de grens van redelijk constant net zoals bij de vorige twee figuren bij de zes minuten. Die van de NW aankomsten daarentegen begint al redelijk te schommelen bij de vier minuten.

Bij alle vier de figuren is wel duidelijk dat de massa van de aankomsten in de eerste minuten ligt en dat de schommelingen zo klein zijn dat de Hazard Rate zelfs constant te noemen is. Na de eerste vier tot zes minuten worden de schommelingen groter omdat er minder data punten zijn bij de overige wachttijden. Al met al is wel duidelijk te zien dat wanneer men kijkt naar de massa dat bij alle vier de grafieken de lijn eerst omhoog schiet in de eerste paar seconden en daar langzaam afneemt. Dit afnemen gebeurt bij de één wat langer dan de anders maar er is een mogelijkheid dat dit redelijk fit op een exponentiële distributie. Echter is verder onderzoek nodig om dit te kunnen concluderen.

Om tot slot wat meer over de wachttijden te kunnen zeggen is in tabel 6 een overzicht te zien van de wachttijden bij PS, NW, IN en NE aankomsten.

Gemiddelde wachttijd	Reguliere Aankomsten (PS)	Potentiele klant (NW)	Internet consult (IN)	Aandelen transactie (NE)
Alle tijden	66,77 sec.	87,38 sec.	82,28 sec.	89,42 sec.
Van afhakers	50,35 sec.	82,21 sec.	131,36 sec.	78,90 sec.
Van virtuele wachttijd	70,12 sec.	90,63 sec.	68,74 sec.	89,59 sec.

Tabel 6: Wachttijd gegevens in november en december inclusief afhakers

Opvallend uit de tabel hierboven is dat de gemiddelde wachttijd voor de afhakers bij de PS, NW en NE aankomsten lager ligt dan voor de virtuele wachttijd. Alleen bij de IN aankomsten is de wachttijd van de afhakers hoger. Bij de reguliere aankomsten is in het geval van ophangen het geduld het kortst. Bij internet aankomsten daarentegen ligt de gemiddelde wachttijd van afhakers veel hoger. Men is dus bereid lang te wachten maar er blijft natuurlijk een grens van geduld.

Al met al is uit dit hoofdstuk niet te concluderen of de wachttijd nou exponentieel verdeeld is. De invloed van afhakers is zo van belang dat een exponentiële fit niet mogelijk is. Ondanks dat de Hazard Rate een redelijk constante lijn toont bij de figuren 19a tot en met 19d is het bewijs nog niet 100%. Dit vraagt om een dieper onderzoek wanneer men dit 100% zou willen vaststellen.

7 Conclusie

Aan de hand van de literatuur is geprobeerd de onderzoeken of alle feiten uit de literatuur met betrekking tot de wachtrij theorie van toepassing is op de data van een bestaand callcenter. Hierbij zijn de volgende conclusies getrokken.

De aankomstintensiteit is een tijd in-homogeen Poisson proces.

Met betrekking tot het aankomst proces is onderzoek gedaan naar de aankomst intensiteit. Hieruit bleek dat de aankomstintensiteit niet netjes verdeeld is volgens een homogeen Poisson proces. Echter wanneer er gekeken wordt naar een constante aankomstintensiteit per tijdsperiode was wel sprake van een homogeen Poisson proces in alle onderzochte gevallen.

De servicetijd is niet exponentieel verdeeld.

In de literatuur is de servicetijd vaak beschreven als tijden die exponentieel verdeeld zijn. Echter is uit dit onderzoek duidelijk gebleken dat dit niet het geval is bij dit callcenter. Er is onderzoek gedaan naar van welke verdeling er wel sprake zou zijn. De best mogelijke verdeling is in dit geval de log-normale verdeling. De log-normale verdeling fitte het best op de servicetijd echter was dit niet 100% te bewijzen volgens de Kolmogorov-Smirnov toets. Een dieper onderzoek in de toekomst met betrekking tot de servicetijd zou een optie kunnen zijn.

Afhakers hebben een grote invloed op het wachttijd proces.

Uit het onderzoek kwam naar voren dat afhakers een grote invloed hebben op het goed kunnen onderzoeken van het wachttijd proces. In de literatuur wordt het wachttijdproces behandeld als een proces waarbij er geen sprake is van afhakers waardoor men kan vaststellen dat de wachttijd exponentieel verdeeld is. Echter uit dit onderzoek is niet 100% te concluderen dat de wachttijd exponentieel verdeeld is. Het afhaakproces is van zo'n grote invloed dat hiermee wel degelijk rekening mee moet worden gehouden bij een dieper onderzoek van de wachttijd. Een groot aantal afhakers is te verwijten aan het feit dat een informatie bericht over de wachttijd een negatieve invloed heeft op de klant waardoor de kans groter wordt dat hij of zij ophangt. In de toekomst zou men dit kunnen verlagen door het informatiebericht minder vaak af te spelen of te vervangen door een ander 'vriendelijker' bericht.

8 Referenties

Lawrence Brown, Noah gans, Avishai Mandelbaum, Anat Sakov, Haipeng Shen, Sergey Zeltyn and Linda Zhao (Revised 2004), "Statistical Analysis of a Telephone Call Center: A Queueing-Science Perspective"

Ger Koole, Avishai Mandelbaum (2001), "Queueing Models of Call Centers An Introduction"

Aalen, O.O., and Gjessing, H. (2001), "Understanding the Shape of the Hazard Rate: A Process Point of View"

Figuren in tabel 4: http://nl.wikipedia.org/wiki/Exponentiële_verdeling,
http://nl.wikipedia.org/wiki/Normale_verdeling, http://nl.wikipedia.org/wiki/Lognormale_verdeling

9 Bijlagen

9.1 Methode voor het valideren van de nulhypothese voor het aankomstproces

Voor het valideren van de nulhypothese van het aankomstproces is gebruik gemaakt van de volgende methode.

1. Selecteer de aankomsten die gevalideerd moet worden onder de nulhypothese en plaats deze in een vector X.
2. Maak een vector met allemaal nullen met dezelfde lengte als X en maak een teller aan voor de procedure.
 - a. $alleR = \text{zeros}(\text{length}(X), 1)$
 - b. teller = 1;
3. Draai de volgende procedure en bereken daarbij eerst de volgende twee waardes en zet die in de procedure.
 - “tot” = “aantal intervallen van 6 minuten (in geval van PS aankomsten) of 60 minuten (in geval van andere aankomsten) van de periode die je gaat gebruiken.
 - “start periode” = eerste tijdstip van de periode in seconden.

```
for i=0: "tot"
    result = X(X >= ("start periode" + (i * 360)) & X < ("start periode" + ((i+1)*360)))
    result2 = result - ("start periode" + (i*360))
    lengte = length(result2)
    for j=1:lengte
        if j == 1
            alleR(teller) = (lengte + (1-j)) * (-log((360-result2(j))/(360-0)))
        else
            alleR(teller) = (lengte + (1-j)) * (-log((360-result2(j))/(360-result2(j-1))))
        end
        teller = teller + 1
    end
end
```

4. Voer de Kolmogorov-Smirnov toets uit met de gevulde vector alleR.
 - a. $\text{Kstest}(\text{alleR})$