

Hawkes punktprocessen: En model for jordskælv i Danmark

RUNE SLOTH CHRISTOPHERSEN
INSTITUT FOR MATEMATISKE FAG
AALBORG UNIVERSITET
31. MAJ 2011

Titel: Hawkes punktprocessen: En model for jordskælv i Danmark

Projektperiode: 1. feb.- 31. maj 2011

Deltagere:

Rune Sloth Christophersen

Vejleder: Jakob Gulddahl Rasmussen

Oplagstal: 4

Sidetæl: 77

Bilagsantal: 25 + CD

Afsluttet: 31. maj 2011

Rapportens indhold er frit tilgængeligt, men offentliggørelse (med kildeangivelse) må kun ske efter aftale med forfatterne.

Synopsis:

Rapportens formål er at give en statistisk analyse af Hawkes punktprocessen. For endvidere at højne forståelsen af Hawkes punktprocessen og dens anvendelser, eksemplificeres teorien igennem anvendelsen af Hawkes punktprocessen, som model for jordskælvshændelser i Danmark.

I rapporten præsenteres teori om punktprocesser på tidslinien, der blandt andet indeholder den betingede intensitetsfunktion, inferens og en præsentation af Hawkes punktprocessen. På denne baggrund konstrueres efterfølgende en model for jordskælvshændelser i Danmark.

For endvidere at kunne medtage styrken af hændelserne i modellen, udvides punktprocesserne på tidslinien til mærkede punktprocesser på tidslinien og de mærkede Hawkes punktprocesser præsenteres. Efterfølgende konstrueres en mærket model for jordskælvshændelserne i Danmark.

Rapporten afrundes med at præsentere perfekt simulering af Hawkes punktprocesser.

Summary

This report is mainly a statistical analysis of the Hawkes point process. To improve the understanding of the theory I will try to use both the marked and the unmarked Hawkes point processes as a model for the earthquakes in Denmark. The reason to why it is used on earthquake data is that the Hawkes point processes is self exciting and is thereby qualified in being able to handle the after quakes after an earthquake.

Initially, bits of general theory are presented. This contains the theory that is needed in the construction of the model, some general theory about spatial point processes and a presentation of a couple of processes.

The point process on the timeline is presented, and general properties of it are examined. The conditional intensity function is presented and it is proven how this can be used to construct the likelihood function of a point process. To be able to make inferens, the general n-dimensional Newton-Rhapson algorithm is presented and it is explained how it is possible to use the log-likelihoodfunction of a point process to estimate its parameters. Finally the Hawkes point process is defined, the likelihoodfunction found, and an algorithm for construction point patterns is presented.

The dataset is examined and a subset of the dataset is chosen to be used for the rest of the report. The reason behind throwing some of the data away, is because the machines that measure earthquakes have become increasingly more sophisticated and better. Therefore the intensity of earthquakes looks like an increasing function of the time. However since it can safely be presumed that the intensity of earthquakes is not influenced by time, I choose only to use the years from 2006-2009. Afterwards it is examined if the dataset is clustered, since the Hawkes point process isn't suited to simulate regular point patterns. Following the Hawkes point processes is used as a model for the earthquakes in Denmark. The parameters in the model are found by using a three dimension Newton-Rhapson algorithm and finally the model is examined too see if it fits reality.

The point processeses on the timeline is expanded to marked point processes on the timeline, different kinds of marks are presented and it is proven how it is possible to construct the likelihood function by using the conditional intensity function for marked point processes. The marked Hawkes point processes with unpredictable marks is then defined, the likelihood function is found and a simulation algorithm is presented.

Following the marked Hawkes point processes with unpredictable marks are used as a model for the earthquakes in Denmark. The parameters in the model are found by using a four dimension Newton-Rhapson algorithm and finally the model are examined too see if it fits the reality.

Finally the *normal* simulation algorithms do not necessarily simulate stationary point patterns on the timeline. The problem arises because of the edge problem, where events *before* the simulated point patterns should have an influence on the distribution of the events in it. It is possible to simulate stationary point patterns for the Hawkes point process directly through perfect simulation, and an algorithm for this is presented.

Forord

Rapporten er udarbejdet i samarbejde med vejleder Jakob Gulddahl Rasmussen i foråret 2011 af Rune Sloth Christophersen.

Denne rapport er mit speciale (MAT6-projekt) i statistik ved Institut for Matematiske Fag på Aalborg Universitet. Rapporten omhandler Hawkes punktprocessen på tidslinien og anvendelsen af disse. Rapporten indeholder blandt andet grundlæggende teori omkring punktprocesser på tidslinien, den betingede intensitets funktion, inferens med en n -dimensionel Newton-Rhapon algoritme, samt simulering. Endvidere er en del af rapporten konstruktionen af to modeller for jordskælvshændelser omkring Danmark. Disse modeller konstrueres på baggrund af et datasæt for jordskælvshændelser i området omkring Danmark.

Til modelkonstruktionen, samt undersøgelsen af datasættet, indgår en del databehandling. Al denne databehandling er foregået med statistikprogrammet R, samt med udvidelsespakken `spatstat`. Det er valgt at vedlægge R-scriptsene i appendikset til denne rapport, og der vil blive henvist til dem, når de anvendes. Årsagen til dette er at implementeringen af teorien i R, både er en tidskrævende og udfordrende proces. Samtidig fungerer de også som dokumentation for, at mine udregninger er korrekte.

Jeg har tidligere, på mit ottende semester, arbejdet med rumlige punktprocesser. For at denne rapport skal kunne læses selvstændigt, har det været nødvendigt enkelte steder at medtage teori, der er gennemgået i dette tidligere projekt. Hver gang sådan teori anvendes, vil der blive henvist til det tidligere projekt med [11]. Endvidere er dette tidligere projekt vedlagt på en CD.

I rapporten antages det, at læseren er bekendt med grundviden omkring sandsynlighedsteori samt rumlige punktprocesser.

Læsevejledning

Igennem rapporten vil der fremtræde kildehenvisninger, som henviser til en litteraturliste bagerst i rapporten.

Der vil i rapporten refereres til kilderne som $[x,y]$, hvor x vil være det nummer, som kilden står under i litteraturlisten, og y vil angive side eller linje i den pågældende kilde. Hvis kildehenvisningen er skrevet før punktum, henviser kilden til den pågældende linje. Hvis kildehenvisningen derimod står efter punktum, henvises der til hele det forudgående afsnit.

Bøger vil i litteraturlisten stå angivet med forfatter, titel, år og ISBN-nummer, mens internetsider vil være angivet med forfatter, titel og år. Hvis internetkilder er benyttet i rapporten, vil disse fremgå i teksten som $[x]$.

Igennem rapporten anvendes både $\exp(\cdot)$ og $\text{Exp}(\cdot)$. De har dog ikke den samme betydning, hvor $\exp(\cdot)$ er eksponentialfunktionen, så er $\text{Exp}(\cdot)$ den eksponentielle fordeling med parameter $\cdot$.

Når der i rapporten anvendes en logaritme vil det altid være den naturlige logaritme medmindre andet er nævnt.

1	Indledning	1
2	Grundlæggende teori	3
2.1	Jordskælv	3
2.2	Forgreningsprocesser	4
2.3	Den rumlige punktproces	7
3	Punktprocesser på tidslinien	12
3.1	Grundlæggende definitioner	12
3.2	Den betingede intensitetsfunktion	14
3.3	Inferens	16
3.4	Hawkes punktprocessen på tidslinien	23
3.5	Simulering	28
3.6	Modelkontrol	32
4	Datapræsentation	33
4.1	Datasæt	33
4.2	Udpluk af data	34
4.3	Fejl og mangler	35
4.4	Fordelingen af punkterne	36
5	Den umærkede Hawkes model	38
5.1	Likelihoodfunktionen for den umærkede model	39
5.2	Parameter estimation	40
5.3	Modelkontrol	44
6	Mærkede punktprocesser på tidslinien	46
6.1	Den mærkede Hawkes punktproces	48
6.2	Simulering af den mærkede Hawkes punktproces	50
7	Den mærkede model	53
7.1	Fordelingen af mærkerne	54

7.2	Likelihoodfunktionen for den mærkede model	55
7.3	Parameter estimation for den mærkede model	56
7.4	Modelkontrol af den mærkede model	62
8	Perfekt simulering	65
8.1	Fordelingen af længden af en klynge	66
8.2	Perfekt simulerings algoritme	68
8.3	Generering af I2	69
9	Model diskussion	72
10	Opsamling	74
	Litteratur	75
11	Bilag	a
11.1	Borelmængde	a
11.2	Additivitet	a
11.3	Tayler Polynomier	b
11.4	Pertielt afledte af den umærkede model	c
11.5	Pertielt afledte af den mærkede model	i
11.6	R-kode	p

KAPITEL 1

INDLEDNING

Denne rapport er hovedsageligt en statistik analyse af Hawkes punktprocessen. For yderligere at højne forståelsen af teorien, vælges det at anvende både en mærket og en umærket Hawkes punktproces, til at konstruere en model for jordskælvshændelser i Danmark.

Først betragtes lidt generel teori. Denne teori er blandt andet grundlæggende teori om rumlige punktprocesser, samt generel teori om jordskælvshændelser, der anvendes i den senere modelkonstruktion.

Efterfølgende præsenteres punktprocessen på tidslinien, og generelle egenskaber for denne undersøges. Den betingede intensitetsfunktion præsenteres og det bevises, at denne kan anvendes til at konstruere likelihoodfunktionen for en punktproces. Der gennemgås hvorledes det er muligt, at bestemme ukendte parametre ved inferens, via den n-dimensionelle Newton-Rhapson algoritme. Hawkes punktprocessen defineres, og dennes likelihoodfunktion bestemmes, og en simuleringsalgoritme præsenteres.

Før det er muligt at konstruere den umærkede model, er det nødvendigt med en indledende undersøgelse af det anvendte datasæt. Efter det er blevet klarlagt, at der er en sammenklumpning af punkterne i datasættet, er det muligt at anvende Hawkes punktprocessen til at konstruere en model for jordskælvshændelser i Danmark. Parametrene for modellen findes med en tre-dimensionel Newton-Rhapson algoritme, og det undersøges om modellen passer til virkeligheden.

Punktprocesserne på tidslinien udvides til mærkede punktprocesser på tidslinien. Der bliver præsenteret forskellige typer af mærker, samt hvorledes det er muligt at konstruere likelihoodfunktionen, med den betingede intensitetsfunktion, for en mærket punktproces på tidslinien. Den mærkede Hawkes punktproces med uforudsigelige mærker præsenteres, likelihoodfunktionen findes og en simuleringsalgoritme præsenteres.

Følgende udvides modellen for jordskælvshændelserne i Danmark, til at være en mærket Hawkes punktproces med uforudsigelige mærker. Parametrene i denne bliver bestemt ved brug af en fire-dimensionel Newton-Rhapson algoritme, og det undersøges om modellen passer til virkeligheden.

Endelig gælder der, at de *normale* simuleringsalgoritmer ikke nødvendigvis simulerer stationære punktmønstre på tidslinien. Problematikken opstår idet der er et kantproblem, hvor hændelser *før* det simulerede punktmønster, burde påvirke fordelingen af punkter i dette. Derfor præsenteres en algoritme, der gør det muligt direkte at simulere stationære punktmønstre for Hawkes punktprocessen.

Afslutningsvis rundes rapporten af med en kort diskussion af modellen og en opsamling.

Det er blevet klarlagt at der i denne rapport blandt andet bliver konstrueret en model, der kan simulere jordskælvshændelser i Danmark. For at være i stand til at opbygge en model der inkorporerer de forskellige aspekter ved jordskælvshændelser, bliver der derfor kort præsenteret generel teori omkring jordskælvshændelser, med det formål at opbygge et forståelsesgrundlag for naturfænomenet.

Der vil også blive defineret hvad forgreningsprocesser er, og der vil blive præsenteret et par grundlæggende egenskaber for disse. En forgreningsproces kan bruges, hvis det ønskes at generere data, hvor det er muligt at betragte forskellige *generationer* af hændelser. I en forgreningsproces er der dermed hændelser, der er afkom af andre hændelser. En sådan proces kan f.eks. bruges til at beskrive reproduktion i økologiske systemer.

Endelig vil den grundlæggende teori afsluttes med en kort og generel beskrivelse af rumlige punktprocesser, således at læseren er klædt på til at forstå punktprocesserne på tidslinien. Jeg har tidligere, på mit ottende semester, medvirket i en projektgruppe der har udfærdiget en grundlæggende rapport om rumlige punktprocesser, [11]. Denne kan findes som pdf på den vedlagte cd.

Målet med dette afsnit er ikke at give en grundig beskrivelse af de teoretiske emner, men blot at præsentere dem, idet de anvendes i beviser og argumenter senere i projektet.

2.1 Jordskælv

I dette afsnit vil der kort blive præsenteret lidt generel teori omkring jordskælv. Målet er ikke at gå i dybden med emnet, men blot klarlægge enkelte egenskaber, der spiller en rolle for den senere modelkonstruktion. Der vil derfor f.eks. ikke blive gennemgået hvorfor eller hvorledes jordskælv opstår.

Intensiteten og styrken af forskellige naturfænomener er ikke altid statisk, nogle af dem ændres naturligt over tid, imens andre påvirkes af f.eks. mennesket. Det vælges dog, i denne rapport, at gå ud fra at der ikke er nogen forskel på intensiteten eller styrkerne af jordskælv-

ene, som funktion af tiden. Det bemærkes dog, at det måske er muligt for mennesket at fremprovokere jordskælv ved meget store byggerier, som f.eks. dæmninger, men dette ses der også bort fra i denne rapport.

Når der forekommer jordskælvshændelser, kan disse være i stand til at fremprovokere flere mindre jordskælv i tidsperioden efter. Disse mindre skælv kaldes for efterskælv. Der er dog også mulighed for, at der kan være mindre jordskælv i en tidsperiode op til en store jordskælvshændelse, disse kaldes for forskælv.

Hvis det ønskes at sammenligne styrken af forskellige jordskælvshændelser, er det muligt at sammenligne amplituden af de seismiske bølger. På baggrund af denne amplitude, er det muligt at kvantificere de forskellige jordskælvshændelser til et tal fra 0 til 10+. Dette kaldes for Richter-skalaen, og er en logaritmisk skala med en base på 30, dvs. at et jordskælv med en styrke på 2 på Richter-skalaen er 30 gange så kraftig, som en hændelse med styrke 1. Følgende er det muligt at se skadesvirkningerne ved jordskælv af forskellig styrke i tabel 2.1.

Tabel 2.1: Richter-skalaen, baseret på [1, s. 259]

Richter-tallet	Effekt	Gennemsnitligt antal per år
< 3,4	Umærkelige	800000
3,5 – 4,2	Kan mærkes inden døre	30000
4,3 – 4,8	Vinduer klirrer	4800
4,8 – 5,4	Døre bevæger sig, ting falder på gulvet	1400
5,5 – 6,1	Lette bygningsskader	500
6,2 – 6,9	Betragtelige bygningsskader	100
7 – 7,3	Alvorlige bygningsskader	15
7,4 – 7,9	Svære bygningsskader, de fleste bygninger falder sammen	4
> 8	Totale bygningsskader	1

Til sammenligning målte detoneringen af et thermo nukleart våben i 1992 6,6 på Richter-skalaen.

[1, s. 254-255, 261-266]

2.2 Forgreningsprocesser

I dette afsnit vil der kort blive redegjort for forgreningsprocesser¹, samt hvornår de er endelige.

Først defineres forgreningsprocesserne.

¹Direkte oversat fra det engelske *Branching processes*

Definition 2.1. – Forgreningsproces

Hvis der haves en proces X hvor følgende gælder:

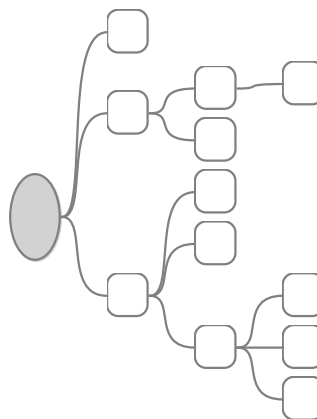
1. Det er muligt, at opdele de genererede hændelser i to forskellige kategorier den oprindelige immigrant og afkom. Afkommet er de hændelser der på en eller anden måde er affødt af en af de tidligere hændelser, hvorimod immigranten er den hændelse, der ikke har nogen form for forfar.
2. Alle hændelser har den samme mulighed for at få nyt afkom, uafhængigt af andre hændelser, ligegyldigt om det er en immigrant eller afkom fra andre hændelser.

Så siges X at være en forgrenende proces.

[9, s. 433-434]

På baggrund af punkt 2. i definition 2.1, er det muligt at betragte antallet af afkom, fra hver hændelse i en forgreningsproces, som en stokastisk variabel trukket fra afkomsfordelingen. Afkomsfordelingen er tæthedsfunktionen for den diskrete stokastiske variabel der angiver antallet af afkom. Yderligere er det muligt, at opstille hændelser genereret via en forgreningsproces som enten et træ eller en skov².

Eksempler på hændelser, der følger denne type af adfærd er f.eks. bakterier. På figur 2.1 er det muligt at se et grafisk eksempel på en forgreningsproces.



Figur 2.1: Et eksempel på et træ, der illustrerer en forgrenende proces. Figuren er lavet i mindmap programmet FreeMind.

For at afgøre om antallet af hændelser i en forgreningsproces er endelig, undersøges det om populationen på et eller andet tidspunkt *uddør*. Altså om der på et eller andet tidspunkt er en generation, hvor der ikke er noget afkom. Først betragtes strukturen, og det bemærkes at i et træ for en forgræningsproces, vil roden altid være immigranten. Det er dermed muligt, at betragte de forskellige hændelser som forskellige generationer efter *ankomsten* af immigranten. Immigranterne betragtes således som 0. generation (Z_0), deres afkom som 1. generation (Z_1) afkom af afkommet af immigranten som 2. generation (Z_2) og så videre.

²Et træ er en sammenhængende graf hvor der er $n - 1$ kanter hvis der er n knuder. En skov er en mængde af træer.

Det er muligt at udregne antallet af hændelser i den n 'te generation rekursivt, som en sum af antallet af afkom for den $(n-1)$ 'te generation,

$$Z_n = \sum_{i=1}^{Z_{n-1}} X_i^{(n-1)},$$

hvor $X_i^{(n-1)}$ er antallet af afkom for den i 'te hændelse i den $(n-1)$ 'te generation. At spørge om en forgreningsproces er endelig, er dermed det samme som at spørge om der er en generation $Z_n = 0$, for $n \in \mathbb{Z}_+$. Lad hændelsen E være at den forgrenende proces uddør.

Det viser sig, at det eneste der er nødvendigt for at sikre endelighed af en forgreningsproces er, at det gennemsnitlige antal afkom er mindre end en altså, at $v < 1$. For at vise dette anvendes den sandsynlighedsgenererende funktionen.

Definition 2.2. – Sandsynlighedsgenererende funktion

Lad den stokastiske heltallig variabel X være et heltal og positivt. Så kaldes funktionen

$$G_X(s) = E[s^X], \quad 0 \leq s \leq 1$$

for den sandsynlighedsgenererendefunktionen for X .

[9, s. 236]

Der gælder endvidere for den sandsynlighedsgenererende funktioner at $G_X(1) = 1$, og at $G_X(0) = p_X(0)$, hvor $p_X(0)$ er sandsynligheden for at den stokastiske variabel er lig 0, jvf. [9, s. 237-238]. Bemærk at p_X er tæthedsfunktionen for den stokastiske variabel X . Følgende gælder for forgreningsprocesser.

Sætning 2.3.

Hvis der haves en forgreningsproces med sandsynlighedsgenererende funktion G , så er $q = \Pr\{E\}$ den mindste løsning til $s = G(s)$, hvor $s \in [0, 1]$.

[9, s. 436]

Denne sætning vil ikke blive vist i denne rapport. Ligeledes vil den næste sætning heller ikke blive bevist.

Sætning 2.4.

Hvis X har sandsynlighedsgenererendefunktionen G_X , så gælder der, at

$$E[X] = G'_X(1).$$

[9, s. 238]

I den følgende sætning er det muligt, at vise at hvis $v < 1$, så vil antallet af hændelser være endeligt.

Sætning 2.5.

Hvis der haves en forgreningsproces $\mathbf{X}$, med gennemsnitligt antal afkom ν , så gælder der:

$$\nu < 1 \Rightarrow P(E) = 1$$

$$\nu > 1 \Rightarrow P(E) < 1$$

[9, s. 437-438]

Bevis. Hvis sandsynligheden for ikke at have nogle afkom er lig 0 så er $P(E) = 0$.

Antag at sandsynligheden for ikke at have nogle afkom er større end 0 ($p_x(0) > 0$). Så gælder der, at $\nu = G'(1)$, jvf. sætning 2.4. Da der gælder at $G(1) = 1$ og $G(0) = p_x(0)$, og da $p_x(0) > 0$ må $G(0) > 0$. Jvf. [9, s. 264 opgave 133] gælder endvidere at $G(s)$ er konveks og voksende, derfor må $G(0)$ være voksende til $G(1)$. Hvis $\nu > 1$ gælder der at hældningen i $G(1)$ er større end 1, og den konvekse og voksende $G(s)$, er derfor nødt til at krydse identitetslinien $y = s$. Hvis $\nu < 1$, så gælder der at hældningen i $G(1)$ er mindre end 1, og den konvekse og voksende $G(s)$ kan derfor ikke skære identitetslinien $y = s$.

Der gælder jvf. sætning 2.3, at sandsynligheden for at uddø er givet ved den mindste løsning til systemet $s = G(s)$, hvor $s \in [0, 1]$. Dermed vil den mindste løsning s være mindre end 1, når $\nu > 1$, og den vil være 1 når $\nu < 1$. $\square$

Bemærk at i tilfældet hvor $\nu = 1$ er det ikke sikkert at $P(E) = 1$, men det er ikke umuligt.

[9, s. 236-238, 433-440]

2.3 Den rumlige punktproces

Dette afsnit indeholder en kort præsentation af den rumlige punktproces. Dele af dette afsnit er lånt fra et tidligere projekt [11].

Hvis der haves et observationsvindue W , og der er observeret en mængde af hændelser indenfor dette vindue, er dette et punktmønster.

Definition 2.6. – Punktmønster

Et punktmønster er en mængde af punkter $\{u_1, u_2, \dots\}$ indeholdt i et observationsvindue W .

Hvis det ønskes at generere punktmønstre, er dette muligt ved at anvende en rumlig punktproces, hvilket er en stokastisk proces, der genererer punktmønstre i et rum S . Fremover i rapporten vil der kun blive betragtet punktmønstre med et udfaldsrum, der er en delmængde af $\mathbb{R}^d$. Følgende i rapporten betegnes en punktproces $\mathbf{X}$, og dens punktmønstre for $\mathbf{x}$.

Nu defineres to forskellige tællemaal.

Definition 2.7. – Tællemålet n

Hvis der haves et punktmønster $\mathbf{x}$, så gælder der at tællemålet $n(\mathbf{x})$ angiver antallet af punkter i punktmønstret. Hvis

$$n(x) < \infty, \text{ når } S \subseteq B$$

hvor B er en borelmængde^a, så siges $\mathbf{x}$ at være lokalt endeligt.

^aDefinitionen på en borelmængde kan ses i appendiks 11.1.

Definition 2.8. – Tællemålet N

Hvis der haves en borelmængde $B \subseteq S$, så angiver $N(B)$ antallet af hændelser i borelmængden B .

Hvis der gælder, at der ikke er nogle punkter i borelmængden B haves der en void-hændelse.

Definition 2.9. – Void-sandsynlighed

Void-sandsynligheden for en begrænset Borelmængde er givet ved

$$v(B) = P(N(B) = 0),$$

hvor $B \in \mathcal{B}_0$.
[8, s. 9]

Der gælder endvidere, at

Sætning 2.10.

En punktproces $\mathbf{X}$ er entydigt bestemt ved dennes void-sandsynlighed.

$$v(B) = P(N(B) = 0), \quad \text{for } B \in \mathcal{B}_0$$

[8, s. 9]

Beviset for denne sætning vil ikke blive gennemgået i rapporten.

Det er muligt at snakke om stationaritet af punktprocesser.

Definition 2.11. – Stationaritet og isotropi

En punktproces $\mathbf{X}$ på $\mathbb{R}^d$ er stationær, hvis dens fordeling er invariant under translation, dvs. at

$$\mathbf{X} \sim \{u + s | u \in \mathbf{X}\}, \quad s \in \mathbb{R}^d.$$

En punktproces $\mathbf{X}$ er isotropisk, hvis dens fordeling er invariant under rotation, dvs. at

$$\mathbf{X} \sim \{Ru | u \in \mathbf{X}\},$$

hvor R betegner en rotation omkring origo.

[8, s. 14-15]

Følgende præsenteres to egenskaber ved punktprocesser.

Definition 2.12. – Superposition

Hvis $\mathbf{X}_i$ er disjunkte punktprocesser for $i = 1, 2, \dots$, kaldes $\bigcup_{i=1}^{\infty} \mathbf{X}_i$ for en superposition.

[8, s. 22]

Definition 2.13. – Udtynding

Lad $\mathbf{X}$ være en punktproces på S , og lad p være en funktion, således at $p : S \rightarrow [0, 1]$. Den udtyndede punktproces $\mathbf{X}_{thin} \subseteq \mathbf{X}$ fremkommer ved, at ethvert punkt $u \in \mathbf{X}$ uafhængigt af hinanden beholdes med sandsynlighed $p(u)$, eller forkastes med sandsynlighed $1 - p(u)$.

[8, s. 22]

Ofte haves der et givent punktmønster for en virkelig hændelse, og det ønskes at finde en punktproces der er i stand til at generere punktmønstre, der minder om det givne punktmønster. Der gælder dog, at et givent punktmønster kan have forskellige tendenser, f.eks. kan punkterne være samlet i små klynger, eller det kan være meget regulært, således at der ikke er nogen punkter, der ligger tæt på hinanden.

Konstruktionen af en rumlig punktproces er altafgørende for hvilken type af punktmønstre den genererer. Det er derfor nødvendigt på forhånd at vide, hvilken type af punktmønstre der ønskes genereret, inden det er muligt at udvælge en rumlig punktproces til dette. For at undersøge et rumligt punktmønster, er det muligt at anvende forskellige deskriptive størrelser, disse vil dog ikke videre blive behandlet i denne rapport. Årsagen til dette er, at en senere indskrænkning til kun at betragte punktprocesser på tidslinien, giver andre muligheder for at vurdere om der er tendenser til f.eks. klyngedannelse i et punktmønster. Det er dog muligt, at læse mere om deskriptive størrelser i den vedlagte rapport [11].

De rumlige punktprocesser kan groft opdeles i tre forskellige klasser, alt efter hvilken type af punktmønstre de genererer, nemlig de klyngede, tilfældige og regulære. Følgende er et par eksempler på forskellige punktprocesser, der genererer forskellige typer af punktmønstre

- Klyngede punktmønstre: Klyngeprocesser.
- Tilfældige punktmønstre: Poisson punktprocessen.
- Regulære punktmønstre: Markov punktprocessen.

Her er det nødvendigt specielt at bemærke Poisson punktprocessen, der er i stand til at generere punktmønstre, hvor punkternes placering er fuldstændigt tilfældige. I det vedlagte projekt, fra mit ottende semester, blev den rumlige udgave af Poisson punktprocessen præsenteret, og beviset for eksistensen samt entydigheden blev udført. Der gælder dog, at Poisson punktprocessen på tidslinien har en noget simplere form. Derfor vil der ikke blive gået mere i dybden med den rumlige udgave her.

Der er dog en type af rumlige punktprocesser det er nødvendigt at gøre lidt mere ud af, da den bliver anvendt senere i projektet. Dette er klyngeprocessen.

[8, s. 7-9, 22-23]

2.3.1 Klyngeprocesser

Hvis det ønskes at danne punktmønstre, der er klyngedannende, er det ikke muligt blot at anvende Poisson punktprocessen. Punktmønstre på denne form kan konstrueres ved først at finde klyngernes centre med en punktproces, hvorefter punkterne i hver klynge konstrueres med en anden. Den endelige punktproces bliver da en superposition af punkterne i alle klyngerne. En punktproces af denne type kaldes for en klyngeproces.

Definition 2.14. – Klyngeproces

Hvis der gælder følgende:

1. *Der findes en centerproces bestående af en punktproces $\mathbf{C}$ på $\mathcal{Y}$.*
2. *Der findes en familie af punktprocesser $\{N(\cdot|y)|y \in \mathcal{Y}\}$, der genererer klyngerne på X .*
3. *Der gælder, at $\mathbf{X}$ er en superposition af klyngerne.*

Så siges $\mathbf{X}$, at være en klyngeproces, hvis for ethvert begrænset $A \in \mathcal{B}_X$:

$$N(A) = \sum_{y_i \in \mathbf{C}} N(A|y_i) < \infty.$$

[3, s. 176]

I dette projekt arbejdes der kun med de klyngeprocesser, hvor der gælder at punktprocesserne $N(\cdot|y)$ er uafhængige af hinanden. Det bemærkes, at de punktprocesser der anvendes til konstruktionen af klyngeprocesserne ikke behøver at være Poisson punktprocesser.

Følgende ønskes det at bevise eksistensen af en sådan punktproces. For at en klyngeproces eksisterer, er det nødvendigt at den kun indeholder endeligt mange punkter.

Lemma 2.15.

En klyngeproces eksisterer hvis og kun hvis, for enhver begrænset borelmængde $A \in S$,

$$\sum_{y_i \in \mathbf{C}} p_A(y_i) < \infty,$$

hvor $p_A(y) = \Pr\{N(A|y) > 0\}$ for $y \in \mathcal{Y}$.

[3, s. 177]

Bevis. For at bevise lemmet tages der udgangspunkt i biimplikationen, og det vises begge veje.

Antag først at en klyngeproces eksisterer. Der gælder da at $\sum_{y_i \in \mathbf{C}} p_A(y_i) < \infty$ for enhver begrænset borelmængde $A \subseteq S$ pr. definition 2.14.

Antag dernæst, at der gælder at der findes en given center punktproces $\mathbf{C}$, hvor der gælder, at $\sum_{y_i \in \mathbf{C}} p_A(y_i) < \infty$ for enhver begrænset borelmængde $A \subseteq S$. Det ønskes at anvende Borel-

cantelli lemmaet [17]. Dette siger, at for E_n uafhængige hændelser gælder der, at hvis

$$\sum_{n=1}^{\infty} Pr\{E_n\} < \infty \quad \text{så} \quad Pr\{\limsup_{n \rightarrow \infty} E_n\} = 0$$

Betragt de uafhængige hændelserne $E_i = \{\text{Klynge } i \text{ tilfører mindst et punkt til } A\}$, da

$$\sum_{n=1}^{\infty} Pr\{E_n\} = \sum_{y_i \in \mathbf{C}} p_A(y_i) < \infty,$$

følger det, at der vil være endeligt mange punkter i mængden A . □

Klyngeprocessen er blot en blandt mange punktprocesser, der er i stand til at danne punktmønstre med klynger. F.eks. er Cox-punktprocesserne også i stand til at danne klyngede punktmønstre. Følgende i rapporten vil der blive arbejdet med punktprocesser på tidslinien, og udelukkende punktprocesser af en type der danner klynger.

[3, s. 175-181]

KAPITEL 3

PUNKTPROCESSER PÅ TIDSLINIEN

Mange hændelser kan afbildes som en funktion af tiden. Dette kunne f.eks. være ankomsten af en ny kunde til en butik, eller tidspunktet for jordskælvshændelser. Det der er væsentligt for alle disse forskellige hændelser er, at de har evolutionær karakter. Det vil sige, at en hændelse kan være afhængig af tidligere hændelser, men den kan *ikke* være afhængig af fremtidige hændelser. Det ønskes, i dette afsnit, at præsentere punktprocesser på tidslinien, der er punktprocesser der opfylder at $S \subseteq \mathbb{R}$, og som har evolutionær karakter¹. Først først præsenteres kort lidt generel teori om blandt andet intensiteten og stationaritet. Efterfølgende betragtes det, hvorledes der laves inferens for punktprocesser på tidslinien.

Der findes mange forskellige klasser af punktprocesser på tidslinien, i dette afsnit præsenteres to af disse, Poisson punktprocessen og Hawkes punktprocessen, hvilket er punktprocesser der danner punktmønstre, hvor punkterne er hhv. fuldstændigt tilfældigt fordelte eller fordelt i klynger. Kapitlet rundes af med at betragte, hvorledes det er muligt at simulere punktmønstre fra en Hawkes punktproces.

3.1 Grundlæggende definitioner

En punktproces på tidslinien er en punktproces der genererer punktmønstre på tidslinien. Der er dog flere forskellige ækvivalente måder, at definere en punktproces på tidslinien på.

Definition 3.1. – Punktproces på tidslinien

En punktproces på tidslinien defineres ved ét af følgende tre ækvivalente punkter:

1. En stokastisk følge T på $\mathbb{R}$.
2. En stokastisk følge af intervaller $\mathcal{T}$ på $[0, \infty)$.
3. Et stokastisk tællemaal N på $\mathbb{R}$.

[3, s. 41]

¹Jordskælv kan have forskælv, men dette ses der bort fra i modellen.

For yderligere, at undgå at punktprocesser der har uendeligt mange punkter, betragtes kun punktprocesser hvor der gælder at $N(B) < \infty$ for enhver begrænset borelmængde $B \subseteq \mathbb{R}$.

Det vælges også kun at betragte simple punktprocesser på tidslinien. Dette er punktprocesser der danner punktmønstre, hvor det ikke er muligt, at der forekommer to hændelser på nøjagtig samme tidspunkt.

Definition 3.2. – Simple punktprocesser på tidslinien

En punktproces $\mathbf{X}$ siges at være en simpel punktproces, hvis der gælder følgende for tællemalet $N(t)$:

$$\Pr\{N(\{t\}) = 0 \text{ eller } 1 \text{ for alle } t \in \mathbb{R}\} = 1.$$

[3, s. 47]

Det ønskes at bestemme et stationaritets begreb for punktprocesser på tidslinien. Der er dog to forskellige måder, at betragte stationaritet for en punktproces på tidslinien på. Dette er alt efter om det er fordelingen af tællemalet $N(B)$, eller fordelingen af ventetiderne. der er stationær.

Definition 3.3.

En punktproces $\mathbf{X}$ på tidslinien, siges at være stationær, hvis der for ethvert n gælder, at mængden

$$\{N(B_1 + t), \dots, N(B_n + t)\}$$

er uafhængigt af $t \in \mathbb{R}$ for de begrænsede borelmængder $B_1 \subseteq S, \dots, B_n \subseteq S$.

[3, s. 45]

Definition 3.4.

En punktproces $\mathbf{X}$ på tidslinien, siges at være interval stationær, hvis der for ethvert n , og heltal $i_1, \dots, i_n$ gælder, at mængden af intervalllængderne

$$\{\tau_{i_1+k}, \dots, \tau_{i_n+k}\},$$

er uafhængigt af k .

[3, s. 45]

For en stationær punktproces på tidslinien er det muligt, at finde *intensiteten* af punkter.

Sætning 3.5. – Kinchin's eksistenssætning

For en stationær punktproces, findes grænseværdien

$$\lambda = \lim_{h \rightarrow 0} \frac{\Pr\{N(0, h] > 0\}}{h},$$

omend den kan være uendelig.

[3, s. 46]

Bevis. Betragt funktionen ϕ , hvor

$$\phi(x) = \Pr\{N(0, x] > 0\}.$$

Hvis $x \rightarrow 0$ vil det medføre at $\phi(x) \rightarrow 0$ idet $\Pr\{N(0, x] > 0\} \rightarrow 0$. Yderligere gælder der for $x, y > 0$,

$$\begin{aligned}\phi(x+y) &= \Pr\{N(0, x+y] > 0\} \\ &= \Pr\{N(0, x] > 0\} + \Pr\{N(0, x] = 0, N(x, x+y] > 0\} \\ &\leq \Pr\{N(0, x] > 0\} + \Pr\{N(x, x+y] > 0\} \\ &= \phi(x) + \phi(y).\end{aligned}$$

Dermed gælder der at $\phi(x)$ er subadditiv, jvf. appendiks 11.2, og der gælder at

$$\frac{\phi(h)}{h} = \frac{\Pr\{N(0, h] > 0\}}{h},$$

findes, omend den kan være uendelig. □

Hvis der ikke haves en stationær punktproces, vil $\Pr\{N(x, x+h] > 0\}$ ikke være en konstant funktion, men derimod afhængig af hvor på tidslinien x er placeret. Dermed kan $\Pr\{N(x, x+h] > 0\}$ erstattes af en tæthedsfunktion $f(t)$, og intensiteten er ikke længere konstant, men derimod blevet til en funktion af tiden t . I det tilfælde hvor intensiteten ikke er konstant, men er en funktion af tiden t , kaldes den for intensitetsfunktionen.

[3, s. 41-50, 64-65]

3.2 Den betingede intensitetsfunktion

Da punktprocesser på tidslinien har evolutionær karakter gælder der, at tæthedsfunktionen og intensitetsfunktionen kun kan være betinget af tidligere begivenheder. Den betingede tæthedsfunktion er givet ved $f(t|\mathcal{H}_t)$, hvor historikken $\mathcal{H}_t$ er alle hændelser, der er sket før tidspunktet t . Ved hjælp af den betingede tæthedsfunktion er det muligt, at definere den betingede intensitetsfunktion.

Definition 3.6. – Den betingede intensitetsfunktion

Ud fra den betingede tæthed $f(t|\mathcal{H}_t)$ og den tilhørende fordelingsfunktion $F(t|\mathcal{H}_t)$, er den betingede intensitetsfunktion givet ved

$$\lambda^*(t)dt = \frac{f(t|\mathcal{H}_t)dt}{1 - F(t|\mathcal{H}_t)}.$$

[3, s. 231]

Den intuitive forståelse af den betingede intensitetsfunktion er, at hvis dt er et lille interval fra tidspunktet t og hændelserne før tiden t er givet ved mængden $\{\dots, t_{n-1}, t_n\}$, så giver

$\lambda^*(t) \cdot dt$ middeværdien af antallet af punkter i intervallet, idet:

$$\begin{aligned}\lambda^*(t)dt &= \frac{f(t|\mathcal{H}_t)dt}{1-F(t|\mathcal{H}_t)} \\ &= \frac{Pr\{t_{n+1} \in [t, t+dt]|\mathcal{H}_t\}}{Pr\{t_{n+1} \notin (t_n, t)|\mathcal{H}_t\}} \\ &= \frac{Pr\{t_{n+1} \in [t, t+dt], t_{n+1} \notin (t_n, t)|\mathcal{H}_t\}}{Pr\{t_{n+1} \notin (t_n, t)|\mathcal{H}_t\}} \\ &= Pr\{t_{n+1} \in [t, t+dt] | t_{n+1} \notin (t_n, t), \mathcal{H}_t\} \quad (3.1)\end{aligned}$$

$$\begin{aligned}&= Pr\{t_{n+1} \in [t, t+dt]|\mathcal{H}_{t-}\} \\ &= E[N([t, t+dt])|H_{t-}]. \quad (3.2)\end{aligned}$$

For at få ligning (3.1) anvendes definitionen på den betingede sandsynlighed [9, s. 29]. Ligning (3.2) følger af antagelsen om, at punktmønstrene er simple, så der er enten ingen punkter eller ét punkt i et interval, når dette er tilstrækkeligt lille. Bemærk at historikken $\mathcal{H}_{t-}$ også medtager tiden op til t fra den sidste hændelse t_n og ikke kun de tidligere hændelser.

Følgende gælder for $f(t|\mathcal{H}_t)$ og $F(t|\mathcal{H}_t)$.

Lemma 3.7.

Størrelserne $f(t|\mathcal{H}_t)$ og $F(t|\mathcal{H}_t)$ er givet ved

$$f(t|\mathcal{H}_t) = \lambda^*(t) \exp\left(-\int_{t_n}^t \lambda^*(s)ds\right) \quad (3.3)$$

og

$$F(t|\mathcal{H}_t) = 1 - \exp\left(-\int_{t_n}^t \lambda^*(s)ds\right), \quad (3.4)$$

hvor t_n er det sidste punkt før t .

Dette er blevet vist i et tidligere projekt [11], men vil dog for sammenhængens skyld blive præsenteret her aligevel.

Bevis. Det haves ud fra definition 3.6, at

$$\begin{aligned}\lambda^*(t) &= \frac{f(t|\mathcal{H}_t)}{1-F(t|\mathcal{H}_t)} \\ &= \frac{\frac{d}{dt}F(t|\mathcal{H}_t)}{1-F(t|\mathcal{H}_t)} \\ &= -\frac{d}{dt} \log(1-F(t|\mathcal{H}_t)).\end{aligned}$$

Hvis der integreres på begge sider, fås det ud fra Analysens Fundamentalsætning [16, s. 127], at

$$\begin{aligned}\int_{t_n}^t \lambda^*(s)ds &= -(\log(1-F(t|\mathcal{H}_t)) + \log(1-F(t_n|\mathcal{H}_t))) \\ &= -\log(1-F(t|\mathcal{H}_t)).\end{aligned}$$

Da det er antaget, at $F(t|\mathcal{H}_t)$ er en simpel punktproces, er sandsynligheden for, at den rammer i samme punkt lig 0, så $F(t_n|\mathcal{H}_t) = 0$. Herefter isoleres $F(t|\mathcal{H}_t)$:

$$\begin{aligned} -\log(1 - F(t|\mathcal{H}_t)) &= \int_{t_n}^t \lambda^*(s) ds \\ \Leftrightarrow 1 - F(t|\mathcal{H}_t) &= \exp\left(-\int_{t_n}^t \lambda^*(s) ds\right) \\ \Leftrightarrow F(t|\mathcal{H}_t) &= 1 - \exp\left(-\int_{t_n}^t \lambda^*(s) ds\right), \end{aligned}$$

hvormed ligning (3.4) er opnået. Ud fra Analysens Fundamentalsætning fås ligning (3.3) ved at differentiere $F(t|\mathcal{H}_t)$ [11, s. 46]. $\square$

Der gælder, at den betingede intensitetsfunktion entydigt bestemmer en punktproces.

Lemma 3.8.

Den betingede intensitetsfunktion $\lambda^(t)$ definerer entydigt en punktproces, hvis den opfylder følgende betingelser for alle t , og alle mulige punktmønstre før t :*

1. $\lambda^*(t)$ er veldefineret og ikke-negativ.
2. Integralet $\int_{t_n}^t \lambda^*(s) ds$ er veldefineret.
3. $\int_{t_n}^t \lambda^*(s) ds \rightarrow \infty$ for $t \rightarrow \infty$.

[13, Kursusgang 4 slide 8]

Ligesom lemma 3.7 er beviset for lemma 3.8 blevet behandlet i [11]. I dette tilfælde vil beviset dog ikke blive gentaget her.

Det er muligt at integrere den betingede intensitetsfunktion.

Definition 3.9.

Den integrerede betingede intensitetsfunktion er givet ved

$$\Lambda^*(t) = \int_0^t \lambda^*(s) ds,$$

hvor λ^ er den betingede intensitetsfunktion.*

[3, s. 258]

Det der gør den betingede intensitetsfunktion specielt interessant er, at den kan bruges til at bestemme likelihoodfunktionen, og dermed muliggør inferens.

[3, s. 229-233]

3.3 Inferens

Hvis der haves et datasæt, i form af et punktmønster og en model for en punktproces på tidslinien, er det ofte ønskværdigt at kunne bestemme de parametre i modellen der med

størst sandsynlighed, vil være i stand til, at producere et punktmønster, der ligner det givne datasæt. I dette afsnit præsenteres én metode til inferens, først er det dog nødvendigt at introducere likelihoodfunktionen.

3.3.1 Likelihood og log-likelihood

Hvis der haves en punktproces og et punktmønster er det interessant at kende sandsynligheden for, at punktprocessen er i stand til at frembringe et sådant punktmønster, som funktion af parametrene i modellen. En sådan funktion kaldes for en likelihoodfunktion

Definition 3.10. – Likelihoodfunktionen

Likelihoodfunktionen er givet ved:

$$L(\theta|x) = p(x|\theta) = f_{\theta}(x).$$

Hvor $f_{\theta}(x)$ er tæthedsfunktionen for det fremkomne punktmønster.

Log-likelihoodfunktionen er givet ved logaritmen til likelihoodfunktionen:

$$\mathcal{L}(\theta|x) = \log(L(\theta|x)).$$

I rapporten vil det blot blive benævnt $L(\theta)$ og $\mathcal{L}(\theta)$.

[9, s. 329]

Likelihoodfunktionen siger dermed noget om, hvor stor sandsynligheden er for, at få et bestemt udfald x , når parametrene θ er kendte.

Det der gør den betingede intensitetsfunktion specielt interessant er, at det er muligt at konstruere likelihoodfunktionen på baggrund af denne.

Sætning 3.11.

Hvis der haves en punktproces $\mathbf{X}$ på tidslinien er dennes likelihoodfunktion givet ved

$$L(\theta) = \left(\prod_{i=1}^n \lambda^*(t_i) \right) \exp \left(- \int_0^T \lambda^*(s) ds \right),$$

hvor $\lambda^(t)$ er den betingede intensitetsfunktion, og $\{t_1, t_2, \dots, t_n\}$ er punkterne på tidsintervallet $[0, T)$.*

[3, s. 232]

Dette er bevist i et tidligere projekt [11], men præsenteres her i en kortere udgave.

Bevis. Likelihoodfunktionen er givet ved tæthedsfunktion $f_{\theta}(x)$. Men da $\mathbf{X}$ er en punktproces på tidslinien, og derfor har evolutionær karakter, kan tætheden for et punkt t_i udtrykkes som den betingede tæthedsfunktion $f(t_i|\mathcal{H}_{t_i})$. Dermed er det muligt, at faktorisere $f_{\theta}(x)$ op i den betingede tæthedsfunktion $f(t_i|\mathcal{H}_{t_i})$ for $i = 1, 2, \dots, n$. Bemærk, at på den sidste del af intervallet, altså efter punkt t_n , vil der ikke være noget punkt, og der skal derfor

multipliseres med $(1 - F(T|\mathcal{H}_T))$. Dette er sandsynligheden for ikke at have nogle punkter i den sidste del af intervallet.

$$L(\theta) = f_\theta(x) = f(t_1|\mathcal{H}_1) \cdots f(t_n|\mathcal{H}_n)(1 - F(T|\mathcal{H}_T)), \quad (3.5)$$

Definition 3.6 og lemma 3.7 anvendes på ligning (3.5).

$$\begin{aligned} L(\theta) &= \left(\prod_{i=1}^n f(t_i|H_{t_i}) \right) \frac{f(T|\mathcal{H}_T)}{\lambda^*(T)} \\ &= \left(\prod_{i=1}^n \lambda^*(t_i) \exp\left(-\int_{t_{i-1}}^{t_i} \lambda^*(s)ds\right) \right) \frac{\lambda^*(T) \exp\left(-\int_{t_n}^T \lambda^*(s)ds\right)}{\lambda^*(T)} \\ &= \left(\prod_{i=1}^n \lambda^*(t_i) \right) \exp\left(-\int_0^T \lambda^*(s)ds\right), \end{aligned}$$

hvor $t_0 = 0$. □

[3, s. 232-233][9, s. 329]

3.3.2 Maksimum likelihood estimation

Det er relevant at have metoder, hvormed det er muligt, at finde de værdier for θ , der maksimerer likelihoodfunktionen, idet det vil være de parameterværdier der med størst sandsynlighed giver hændelser, der ligner de oprindelige data.

Hvis der haves en punktproces $\mathbf{X}$ med ukendte parametre θ , men med en tæthedsfunktion givet ved $f_\theta(x)$, vil likelihoodfunktionen, jvf. definition 3.10, være givet ved:

$$L(\theta) = f_\theta(x).$$

Der gælder, at de parametre der maksimerer $L(\theta)$ også maksimerer $\mathcal{L}(\theta)$, da logaritmen ikke ændrer ved placeringen af maksimum. Ydermere gælder der, da logaritmen er en konkav funktion, at et eventuelt entydigt maksimum vil have en hældning på 0 i $L(\theta)$ og i $\mathcal{L}(\theta)$. Derfor er det nok at løse ligningen $u(\hat{\theta}) = 0$, for

$$u(\theta) = \frac{d}{d\theta} \mathcal{L}(\theta).$$

Ydermere gælder der, at $j(\theta)$ er givet ved

$$j(\theta) = -\left(\frac{d}{d\theta}\right)^2 \mathcal{L}(\theta) = -\frac{d}{d\theta} u(\theta).$$

Bemærk at $u(\theta)$ er scorefunktionen, og $j(\theta)$ er den observerede fisher information.

Eksempel 3.12. Lad $\mathbf{X}$ være en homogen Poisson punktproces på tidslinien i intervallet $I = [0, T]$ med betinget intensitetsfunktion intensitet $\lambda^*(t) = \theta$. Ydermere hvis der haves et punktmønster $\{x_1, \dots, x_n\}$, da gælder der, at den tilhørende likelihood- og log-likelihoodfunktion

er givet ved:

$$\begin{aligned}
 L_{\theta}(T) &= \left(\prod_{i=1}^n \lambda^*(t_i) \right) \exp \left(- \int_0^T \lambda^*(s) ds \right) \\
 &= \left(\prod_{i=1}^n \theta \right) \cdot \exp \left(- \int_0^T \theta ds \right) \\
 &= \theta^n \cdot \exp(-\theta T). \\
 \mathcal{L}_{\theta}(T) &= \log(L_{\theta}(T)) \\
 &= \log(\theta^n \cdot \exp(-\theta T)) \\
 &= n \cdot \log(\theta) + \log(\exp(-\theta T)) = n \cdot \log(\theta) - \theta T.
 \end{aligned}$$

Da likelihoodfunktionen og log-likelihoodfunktionen har samme maksimum er det ligegyldigt, hvilken af de to der maksimeres. Da der gælder, at de er konkave funktioner, gælder der at maksimum vil findes der hvor differentialen giver 0. Derfor differentieres $\mathcal{L}(\theta)$ i forhold til T og sættes lig 0.

$$\begin{aligned}
 \frac{d}{d\theta} \mathcal{L}_{\theta}(T) &= 0 \\
 \Leftrightarrow \frac{n}{\theta} - T &= 0 \\
 \Leftrightarrow \theta &= \frac{n}{T}
 \end{aligned}$$

Dermed gælder der, at MLE er givet ved $\hat{\theta} = \frac{n}{T}$.

Det er dog langt fra alle punktprocesser, hvor det er muligt at estimere parametrene analytisk ved MLE som i eksempel 3.12. Det er derfor relevant at betragte, hvorledes det i stedet er muligt at approksimere parametrene. I denne rapport vil der blive betragtet én metode til at approksimere parametrene baseret på en Newton-Rhapson algoritme. Der findes andre måder at approksimere parametre på, f.eks. kan nævnes baysiansk statistik eller ved minimum kontrast estimering.

[8, s. 146-147]

3.3.3 Newton-Rhapson approksimation

Hvis der haves en funktion $f(\theta)$, og det ønskes at approksimere $\hat{\theta}$ således at $f(\hat{\theta}) = 0$ kan dette gøres med en Newton-Rhapson algoritme. Først præsenteres Newton-Rhapson approksimation, hvor antallet af ukendte parametre er lig én, hvorefter der udvides til en mere generel form.

De fleste læsere er sikkert bekendte med den én-dimensionelle udgave af Newton-Rhapson

udledt af et første ordens Taylor polynomium, jvf. appendiks 11.3, hvor $f(\theta) = 0$.

$$\begin{aligned} f(\theta) &\approx f(\theta^{(0)}) + (\theta - \theta^{(0)})f'(\theta^{(0)}) \\ 0 &\approx f(\theta^{(0)}) + (\theta - \theta^{(0)})f'(\theta^{(0)}) \\ \theta &\approx \frac{-f(\theta^{(0)})}{f'(\theta^{(0)})} + \theta^{(0)}. \end{aligned}$$

Dette kan udvides rekursivt til følgende

$$\theta^{(n)} = \theta^{(n-1)} - \frac{f(\theta^{(n-1)})}{f'(\theta^{(n-1)})}, \quad (3.6)$$

hvor der gælder at $\theta_n \rightarrow \hat{\theta}$ når $n \rightarrow \infty$. Det er heller ikke noget problem at anvende ligning (3.6) på punktprocesser med én ukendt variabel. Betragt f.eks. en likelihoodfunktion med én ukendt parameter, hvor toppunktet for likelihoodfunktionen er givet ved $\hat{\theta}$, hvor $\frac{d}{d\theta} \mathcal{L}(\hat{\theta}) = 0$. Da er det muligt at anvende den én-dimensionelle udgave af Newton-Rhaphson fra ligning (3.6).

$$\theta^{(m+1)} = \theta^{(m)} + \frac{u(\theta^{(m)})}{j(\theta^{(m)})}, \quad \text{for } m = 0, 1, \dots,$$

hvor $\theta^m \rightarrow \hat{\theta}$ når $m \rightarrow \infty$.

Der gælder dog, at mange punktprocesser har mere end en ukendt variabel. Det ønskes derfor at finde en Newton-Rhaphson algoritme, der fungerer når antallet af parametre overstiger én. Før den n -dimensionelle Newton-Rhaphson betrædes er det dog først nødvendigt at definere Jacobi- og Hessian-matricerne, idet disse væsentligt letter opskrivningen af algoritmen.

Definition 3.13. – Jacobimatrix

Hvis der haves m funktioner, med n ubekendte, er Jacobimatricen givet ved:

$$J(f_1, \dots, f_m) = \begin{bmatrix} \frac{\partial f_1}{\partial \theta_1}(x) & \frac{\partial f_1}{\partial \theta_2}(x) & \dots & \frac{\partial f_1}{\partial \theta_n}(x) \\ \frac{\partial f_2}{\partial \theta_1}(x) & \frac{\partial f_2}{\partial \theta_2}(x) & \dots & \frac{\partial f_2}{\partial \theta_n}(x) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial \theta_1}(x) & \frac{\partial f_m}{\partial \theta_2}(x) & \dots & \frac{\partial f_m}{\partial \theta_n}(x) \end{bmatrix}$$

[4, s. 1003]

Definition 3.14. – Hessian-matrix

Hvis der haves en funktion $f(x_1, x_2, \dots, x_n)$, så er Hessian-matricen givet ved:

$$H(f) = \begin{bmatrix} \frac{\partial^2 f}{\partial \theta_1^2}(x) & \frac{\partial^2 f}{\partial \theta_1 \partial \theta_2}(x) & \cdots & \frac{\partial^2 f}{\partial \theta_1 \partial \theta_n}(x) \\ \frac{\partial^2 f}{\partial \theta_2 \partial \theta_1}(x) & \frac{\partial^2 f}{\partial \theta_2^2}(x) & \cdots & \frac{\partial^2 f}{\partial \theta_2 \partial \theta_n}(x) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial \theta_n \partial \theta_1}(x) & \frac{\partial^2 f}{\partial \theta_n \partial \theta_2}(x) & \cdots & \frac{\partial^2 f}{\partial \theta_n^2}(x) \end{bmatrix}$$

[18]

For at finde de n ukendte parametre betragtes n ligninger

$$\begin{aligned} f_1(\theta_1, \theta_2, \dots, \theta_n), \\ f_2(\theta_1, \theta_2, \dots, \theta_n), \\ \vdots \\ f_n(\theta_1, \theta_2, \dots, \theta_n), \end{aligned}$$

med n ubekendte. Det ønskes at finde $(\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_n)$, således at

$$\begin{aligned} f_1(\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_n) &= 0, \\ f_2(\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_n) &= 0, \\ \vdots \\ f_n(\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_n) &= 0. \end{aligned}$$

Altså et ligningssystem med n ligninger og n ubekendte. Lad nu

$$\theta = \begin{bmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_n \end{bmatrix} \quad \text{og} \quad \mathbf{f}(\theta) = \begin{bmatrix} f_1(\theta_1, \theta_2, \dots, \theta_n) \\ f_2(\theta_1, \theta_2, \dots, \theta_n) \\ \vdots \\ f_n(\theta_1, \theta_2, \dots, \theta_n) \end{bmatrix}$$

Følgende er det muligt, at opskrive Taylorudvidelsen, jvf. appendiks 11.3.

$$\begin{aligned} \mathbf{f}(\theta) &= \mathbf{f}(\theta^{(0)}) + J(\mathbf{f}(\theta^{(0)})) \cdot (\theta - \theta^{(0)}) + R. \\ \mathbf{f}(\theta) &\approx \mathbf{f}(\theta^{(0)}) + J(\mathbf{f}(\theta^{(0)})) \cdot (\theta - \theta^{(0)}), \end{aligned} \tag{3.7}$$

hvor R er restleddet, $J(\mathbf{f}(\theta^{(0)}))$ er Jacobimatricen, $\theta^{(0)T} = [\theta_1^{(0)}, \theta_2^{(0)}, \dots, \theta_n^{(0)}]$ er et gæt indenfor parameterrummet. Bemærk, at hvor den én-dimensionelle Newton-Rhapon anvender tangenten til at udføre approksimationen, anvender den n -dimensionelle Newton-Rhapon det n -dimensionelle tangentplan.

Da $\mathbf{f}(\boldsymbol{\theta}) = \mathbf{0}$, kan ligning (3.7) omskrives.

$$\begin{aligned}\mathbf{f}(\boldsymbol{\theta}) &\approx \mathbf{f}(\boldsymbol{\theta}^{(0)}) + J(\mathbf{f}(\boldsymbol{\theta}^{(0)})) \cdot (\boldsymbol{\theta} - \boldsymbol{\theta}^{(0)}) \\ \Leftrightarrow \mathbf{0} &\approx \mathbf{f}(\boldsymbol{\theta}^{(0)}) + J(\mathbf{f}(\boldsymbol{\theta}^{(0)})) \cdot (\boldsymbol{\theta} - \boldsymbol{\theta}^{(0)}) \\ \Leftrightarrow \boldsymbol{\theta} &\approx \boldsymbol{\theta}_0 - J(\mathbf{f}(\boldsymbol{\theta}^{(0)}))^{-1} \mathbf{f}(\boldsymbol{\theta}^{(0)}).\end{aligned}$$

Dermed fremkommer en iterativ formel af samme type som (3.6).

$$\boldsymbol{\theta}^{(k)} = \boldsymbol{\theta}^{(k-1)} - J(\mathbf{f}(\boldsymbol{\theta}^{(k-1)}))^{-1} \mathbf{f}(\boldsymbol{\theta}^{(k-1)}). \quad (3.8)$$

Hvor der gælder, at $\boldsymbol{\theta}_k \rightarrow \hat{\boldsymbol{\theta}}$ når $k \rightarrow \infty$.

Hvis det ønskes at anvende en Newton-Rhapson algoritme til at approksimere MLE for en punktproces med n variable, ønskes det igen at finde $\hat{\boldsymbol{\theta}}$, således at de partielt afledte af $\mathcal{L}(\hat{\boldsymbol{\theta}})$ bliver 0. Dermed bliver $\mathbf{f}(\boldsymbol{\theta}^{(k)})$ fra ligning (3.8) en vektor med de partielt afledte af $\mathcal{L}(\boldsymbol{\theta}^{(k)})$, denne vektor kaldes $\Delta(\mathcal{L}(\boldsymbol{\theta}^{(k)}))$.

$$\Delta(\mathcal{L}(\boldsymbol{\theta}^{(k)})) = \begin{bmatrix} \frac{\partial}{\partial \theta_1} \mathcal{L}(\boldsymbol{\theta}^{(k)}) \\ \frac{\partial}{\partial \theta_2} \mathcal{L}(\boldsymbol{\theta}^{(k)}) \\ \vdots \\ \frac{\partial}{\partial \theta_n} \mathcal{L}(\boldsymbol{\theta}^{(k)}) \end{bmatrix}$$

Det er nødvendigt, at finde Jacobimatricen for $\Delta(\mathcal{L}(\boldsymbol{\theta}^{(k)}))$.

$$\begin{aligned}J(\Delta(\mathcal{L}(\boldsymbol{\theta}^{(k)}))) &= \begin{bmatrix} \frac{\partial}{\partial \theta_1} \frac{\partial}{\partial \theta_1} \mathcal{L}(\boldsymbol{\theta}^{(k)}) (x) & \frac{\partial}{\partial \theta_1} \frac{\partial}{\partial \theta_2} \mathcal{L}(\boldsymbol{\theta}^{(k)}) (x) & \dots & \frac{\partial}{\partial \theta_1} \frac{\partial}{\partial \theta_n} \mathcal{L}(\boldsymbol{\theta}^{(k)}) (x) \\ \frac{\partial}{\partial \theta_2} \frac{\partial}{\partial \theta_1} \mathcal{L}(\boldsymbol{\theta}^{(k)}) (x) & \frac{\partial}{\partial \theta_2} \frac{\partial}{\partial \theta_2} \mathcal{L}(\boldsymbol{\theta}^{(k)}) (x) & \dots & \frac{\partial}{\partial \theta_2} \frac{\partial}{\partial \theta_n} \mathcal{L}(\boldsymbol{\theta}^{(k)}) (x) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial}{\partial \theta_n} \frac{\partial}{\partial \theta_1} \mathcal{L}(\boldsymbol{\theta}^{(k)}) (x) & \frac{\partial}{\partial \theta_n} \frac{\partial}{\partial \theta_2} \mathcal{L}(\boldsymbol{\theta}^{(k)}) (x) & \dots & \frac{\partial}{\partial \theta_n} \frac{\partial}{\partial \theta_n} \mathcal{L}(\boldsymbol{\theta}^{(k)}) (x) \end{bmatrix} \\ &= \begin{bmatrix} \frac{\partial^2 \mathcal{L}(\boldsymbol{\theta}^{(k)})}{\partial \theta_1^2} & \frac{\partial^2 \mathcal{L}(\boldsymbol{\theta}^{(k)})}{\partial \theta_1 \partial \theta_2} & \dots & \frac{\partial^2 \mathcal{L}(\boldsymbol{\theta}^{(k)})}{\partial \theta_1 \partial \theta_n} \\ \frac{\partial^2 \mathcal{L}(\boldsymbol{\theta}^{(k)})}{\partial \theta_2 \partial \theta_1} & \frac{\partial^2 \mathcal{L}(\boldsymbol{\theta}^{(k)})}{\partial \theta_2^2} & \dots & \frac{\partial^2 \mathcal{L}(\boldsymbol{\theta}^{(k)})}{\partial \theta_2 \partial \theta_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 \mathcal{L}(\boldsymbol{\theta}^{(k)})}{\partial \theta_n \partial \theta_1} & \frac{\partial^2 \mathcal{L}(\boldsymbol{\theta}^{(k)})}{\partial \theta_n \partial \theta_2} & \dots & \frac{\partial^2 \mathcal{L}(\boldsymbol{\theta}^{(k)})}{\partial \theta_n^2} \end{bmatrix} = H(\mathcal{L}(\boldsymbol{\theta}^{(k)}))\end{aligned}$$

Dermed er Jacobimatricen for vektoren med de partielt afledte af log-likelihoodfunktionen lig Hessian-matricen for log-likelihoodfunktionen.

Dermed er det muligt, at omskrive ligning (3.8).

$$\boldsymbol{\theta}^{(n)} = \boldsymbol{\theta}^{(n-1)} - H^{-1}(\mathcal{L}(\boldsymbol{\theta}^{(n-1)})) \Delta(\boldsymbol{\theta}^{(n-1)}).$$

Hvor $\boldsymbol{\theta}^n \rightarrow \hat{\boldsymbol{\theta}}$ når $n \rightarrow \infty$.

Det er nu blevet betragtet hvorledes der kan udføres inferens, for punktprocesser på tidslinien. Får at kunne udføre inferens er det dog nødvendigt, at have en model. Derfor præsenteres følgende konkrete punktprocesser på tidslinien.

[15, s. 45-49]

3.4 Hawkes punktprocessen på tidslinien

Der er nu blevet gennemgået en del generel teori omkring punktprocesser på tidslinien og inferens for disse. Derfor vil der nu blive præsenteret en klasse af punktprocesser på tidslinien kaldet Hawkes punktprocesser. I konstruktionen af Hawkes punktprocessen anvendes Poisson punktprocessen på tidslinien og derfor præsenteres disse først.

3.4.1 Poisson punktprocessen på tidslinien

Det ønskes at konstruere punktprocesser på tidslinien, der kan generere punktmønstre hvor punkterne er fuldstændigt tilfældigt fordelt, og hvor det gennemsnitlige antal punkter i intervallet $I = [0, T]$ er givet ved $\lambda \cdot T$. Da en punktproces på tidslinien kan defineres som en stokastisk følge af intervaller, er det nødvendigt at finde en måde at generere disse intervaller, således at punkterne bliver fuldstændigt tilfældigt fordelt, men med det samme gennemsnitlige antal punkter. Her udnyttes den hukommelsesløse egenskab ved den eksponentielle fordeling til at konstruere hvert tidsinterval som et udtræk fra denne, således at punkterne bliver fuldstændigt tilfældigt fordelt. Hvis endvidere parametren af den eksponentielle fordeling er λ , så er middelværdien, af intervalstørrelsen, givet ved λ^{-1} . Middelantallet af punkter i intervallet I , bliver da:

$$E[N(I)] = \frac{T}{\lambda^{-1}} = T \cdot \lambda.$$

Det er dermed muligt at definere en Poisson punktproces på tidslinien.

Definition 3.15. – Poisson punktproces på tidslinien

Hvis der gives en punktproces $\mathbf{X}$, hvor størrelsen af hvert interval $\tau_i \sim \text{Exp}(\lambda)$ for $i = 1, 2, \dots$, så siges $\mathbf{X}$ er være en Poisson punktproces på tidslinien med intensitet λ .

Hvis intensitetsfunktionen for en Poisson punktproces er konstant siges Poisson punktprocessen at være homogen, hvis den derimod ikke er konstant siges Poisson punktprocessen at være inhomogen.

Poisson punktprocessen på tidslinien er en byggesten i konstruktionen af andre punktprocesser på tidslinien, f.eks. Hawkes punktprocessen.

[3, s. 19-21]

3.4.2 Hawkes punktprocessen

Hvis det ønskes at konstruere punktmønstre på tidslinien, der er mere klyngedannende end de fuldstændige tilfældige punktmønstre, er det nødvendigt, at anvende en anden type af

punktprocesser end Poisson punktprocessen. Et eksempel på en klyngedannende punktproces er Hawkes punktprocessen. En Hawkes punktproces består af to forskellige typer af punkter, *immigranter* og *afkom*. Immigranterne, $\{y\}$, er kendetegnet ved, at de ikke er afkom af nogle af de punkter der findes i punktprocessen, men derimod er *ankommet* ved en Poisson punktproces $\mathbf{C}$ med intensitet $\mu(t)$. Afkommet skabes ved, at ethvert punkt i punktprocessen har en mulighed for at skabe nye punkter, fordelt efter en endelig inhomogen Poisson punktproces. Da afkom af immigranternes generation er i stand til at skabe nyt afkom, kan der dermed skelnes mellem afkom af forskellige generationer, hvor de oprindelige immigranter er 0'te generation. Et punkt t_i genererer afkom efter en Poisson punktproces med intensitetsfunktion $\alpha\gamma(t - t_i)$, både hvis punktet er en immigrant eller tilhører en generation større end 0. Yderligere ønskes det at klyngerne skal være endelige, og derfor er det påkrævet at det gennemsnitlige antal afkom $v = \int_0^\infty \alpha\gamma(x) < 1$. Følgende betragtes udelukkende Hawkes punktprocesser hvor intensitetsfunktionen $\mu(t)$ er konstant, altså stationære Hawkes punktprocesser.

Der findes en rumlig udgave af Hawkes punktprocessen, men i denne rapport betragtes udelukkende Hawkes punktprocesser på tidslinien. Det vil sige, at udfaldsrummet er $\mathbb{R}$ samt et krav om, at den rumlige punktproces skal have evolutionær karakter. Følgende defineres Hawkes punktproces på tidslinien som en klyngeproces.

Definition 3.16. – Hawkes punktproces

Hvis der gælder at:

1. Immigranterne $\{y\}$ "ankommer" efter en Poisson punktproces med intensitet μ på $\mathbb{R}$.
2. Hver immigrant $y_i \in \{y\}$ danner en klynge $\mathbf{C}_i$, der er uafhængig af de andre klynger. Hver klynge $\mathbf{C}_i$ består af én oprindelig immigrant y_i samt afkom af forskellige generationer. Generation $k + 1$ består af afkom fra generation k , ved at hvert punkt t_i i den k 'te generation, giver afkom efter en inhomogen Poisson punktproces på $\mathbb{R}$ med intensitetsfunktion $\alpha\gamma(t - t_i)$, for $k = 0, 1, 2, \dots$
3. $\mathbf{X}$ er en superposition af alle klyngerne, $\mathbf{C}_i$, for $i = 1, 2, \dots$

Så er $\mathbf{X}$ en Hawkes punktproces på $\mathbb{R}$. [3, s. 183-184]

Det bemærkes, at da Hawkes punktprocessen er defineret som en klyngeproces vil den eksistere jvf. lemma 2.15, hvis der gælder at

$$\sum_{y_i \in \mathbf{C}} p_A(y_i) < \infty.$$

Dette opfyldes hvis

$$v = \int_0^\infty \alpha\gamma(t)dt < 1, \quad (3.9)$$

idet, det er muligt at opfatte hver klynge som en forgræningsproces, og sætning 2.5 medfører at det gennemsnitlige antal punkter i hver klynge bliver endeligt, og da der er et endeligt antal klynger må den samlede mængde af punkter blive endelig. Der gælder endvidere,

ifølge [3, s. 184], at det gennemsnitlige antal punkter i klyngerne er givet ved $\frac{1}{1-\nu}$. Dermed er det gennemsnitlige antal punkter pr. tidsenhed givet ved produktet af μ og $\frac{1}{1-\nu}$, altså $\frac{\mu}{1-\nu}$.

En af de ting, der gør Hawkes punktprocessen interessant er, at det også er muligt at definere Hawkes punktprocessen ud fra dennes betingede intensitetsfunktion $\lambda^*(t)$.

Definition 3.17.

Hvis der gælder, at en punktproces $\mathbf{X}$ har følgende betingede intensitetsfunktion

$$\lambda^*(t) = \mu + \sum_{t_i < t} \alpha \gamma(t - t_i), \quad (3.10)$$

hvor $\mu > 0$, og $\alpha \gamma(t)$ er en tæthed på intervallet $(0, \infty)$, så siges $\mathbf{X}$ at være en Hawkes punktproces.

[3, s. 233-234]

Følgende ønskes det at vise, at de to definitioner på en Hawkes punktproces, 3.16 og 3.17, er ækvivalente. For at bevise dette vælges det at gå samme vej som Allan G. Hawkes anvendte i hans artikel fra 1974 [5, s. 493-503]. Til bevises anvendes to lemmaer, hvoraf det første vil blive bevist og det sidste blot præsenteret.

Lemma 3.18.

Hvis $\mu > 0$ og $\nu < 1$ så findes der en stationær Poisson klyngeproces med $\frac{\mu}{1-\nu}$ punkter per tidsenhed, der både opfylder definition 3.16 og 3.17.

[5, s. 495]

Bevis. Det ønskes først at vise, at der findes en Poisson klyngeproces, der opfylder definition 3.17. Der tages udgangspunkt i en fødsels-døds proces, og det vises at den kan opfylde ligning (3.10). Herefter vises det, at den samme fødsels-døds proces også er en Poisson klyngeproces med gennemsnitligt antal punkter $\frac{\mu}{1-\nu}$ pr. tidsenhed.

Antag at der haves en fødsels-døds proces, hvor dødsraten er 0 og fødselsraten er givet ved en funktion afhængigt af alderen af punktet $\alpha \gamma(t - t_i)$, for punktet t_i . Det bemærkes, at i en fødsels-døds proces, er både sandsynligheden for at dø og afkomsintensiteten, for de forskellige punkter, uafhængigt af hinanden. Dermed gælder der, at afkomsintensiteten i punktet t på tidslinien er givet ved summen af afkomsintensiteten for hvert af de eksisterende punkter $t_i < t$. Dermed kan intensitetsfunktionen for afkommet skrives som:

$$\sum_{t_i < t} \alpha \gamma(t - t_i).$$

Der er altså nu fundet intensitetsfunktionen for afkommet, men det ønskes også at tage højde for immigranter, der er ankommet med en intensitet på μ . Derfor bliver intensitetsfunktionen for punkterne i fødsels-døds punktprocessen følgende:

$$\lambda(t) = \mu + \sum_{t_i < t} \gamma(t - t_i). \quad (3.11)$$

Her er $\{t_i\}$ mængden af alle de punkter der er forekommet før tiden t , enten som immigranter eller som afkom af andre punkter. Det er nu vist, at der findes en fødsels-døds proces, der opfylder ligning (3.10) og dermed definition 3.17.

Følgende vil det blive vist at samme fødsels-døds proces fører til en punktproces defineret som 3.16.

Da de oprindelige immigranter er ankommet uafhængigt af hinanden med en intensitet på μ , og da disse er forfædre til alle punkterne, er det muligt at betragte disse som centerpunkter for klyngerne givet ved en homogen Poisson punktproces med intensitet μ . Hvis immigranternes afkom betragtes som første generation, og afkommet af første generation betragtes som anden generation osv. så kan punktprocessen betragtes som en forgrenende proces, jvf. definition 2.1. Der gælder, jvf. [5, s. 496] at den sandsynlighedsgenererende funktionen for klyngestørrelsen er givet ved:

$$\pi(z) = z \cdot \exp(v(\pi(z) - 1)).$$

Ydermere gælder der, jvf. sætning 2.5, at hvis $v < 1$, så er antallet af punkter i klyngerne endeligt.

Da $\pi(z)$ er en sandsynlighedsgenererende funktion, gælder der at det gennemsnitlige antal af punkter i klyngerne er givet ved $\pi'(1)$, jvf. sætning 2.4. Først findes den afledte af $\pi(z)$, som den afledte af produktet af to funktioner, hvoraf den ene er en sammensat funktion.

$$\pi'(z) = \exp(v(\pi(z) - 1)) + z \cdot \pi'(z) \cdot v \cdot \exp(v(\pi(z) - 1)).$$

Følgende anvendes det at $\pi(1) = 1$, jvf. afsnit 2.2.

$$\begin{aligned} \pi'(1) &= \exp(v(\pi(1) - 1)) + 1 \cdot \pi'(1) \cdot v \cdot \exp(v(\pi(1) - 1)) \\ \Downarrow \quad \pi'(1) &= \exp(0) + 1 \cdot \pi'(1) \cdot v \cdot \exp(0) \\ \Downarrow \quad \pi'(1) &= 1 + 1 \cdot \pi'(1) \cdot v \\ \Downarrow \quad \pi'(1)(v - 1) &= 1 \\ \Downarrow \quad \pi'(1) &= \frac{1}{v - 1}. \end{aligned}$$

For at finde det gennemsnitlige antal punkter per tidsenhed multipliceres intensiteten for immigranterne μ med den gennemsnitlige størrelse af en klynge $\frac{1}{1-v}$. Dermed bliver raten $\frac{\mu}{1-v}$.

Da det nu er vist, at da ovenstående fødsels-døds proces har samme opbygning og rate som definition 3.16, må den siges at være en Hawkes punktproces. $\square$

Dermed gælder der nu, at der er fundet en punktproces der stemmer med både definition 3.16 og 3.17. For at de to definitioner kan siges at være ækvivalente er det nødvendigt at de er entydige. Der gælder følgende.

Lemma 3.19.

Der findes højst en stationær punktproces der opfylder ligning (3.10) og $\int_0^\infty \alpha\gamma(t)dt < 1$.
[5, s. 497-498]

Hvilket ikke vil blive bevist her.

Det er nu muligt at præsentere følgende sætning, der er trivielt bevist ved lemma 3.18 og 3.19.

Sætning 3.20.

Hvis $\nu < 1$ og $\alpha\gamma(s)$ opfylder

$$0 < \nu = \int_0^\infty \alpha\gamma(s)ds < 1,$$

så findes der én punktproces der både opfylder definition 3.17 og definition 3.16.
[5, s. 498]

Nu er Hawkes punktprocessen bevist, og det er vist at den både kan defineres som en klyngeproces og via dens betingede intensitetsfunktion. Følgende betragtes nogle af dens egenskaber.

Da Hawkes punktprocessen er defineret ved den betingede intensitetsfunktion er det relativt simpelt at finde dens likelihood og log-likelihoodfunktion via sætning 3.11.

Sætning 3.21.

For en Hawkes punktproces gælder der, at dennes likelihoodfunktion er givet ved:

$$L(\theta) = \left(\prod_{i=1}^n \mu + \sum_{t_j < t_i} \alpha\gamma(t_i - t_j) \right) \cdot \exp \left(- \int_0^T \left(\mu + \sum_{t_j < s} \alpha\gamma(s - t_j) \right) ds \right), \quad (3.12)$$

hvor $\theta = \{\mu, \alpha, \gamma\}$.

Ligeledes er log-likelihoodfunktionen givet ved:

$$\mathcal{L}(\theta) = \left(\sum_{i=1}^n \log \left(\mu + \sum_{t_j < t_i} \alpha\gamma(t_i - t_j) \right) \right) - \int_0^T \left(\mu + \sum_{t_j < s} \alpha\gamma(s - t_j) \right) ds. \quad (3.13)$$

Bevis. For en Hawkes punktproces er den betingede intensitetsfunktion givet ved definition 3.17. Dette indsættes i sætning 3.11.

$$\begin{aligned} L(\theta) &= \left(\prod_{i=1}^n \lambda^*(t_i) \right) \exp \left(- \int_0^T \lambda^*(s) ds \right) \\ &= \left(\prod_{i=1}^n \mu + \sum_{t_j < t_i} \alpha\gamma(t_i - t_j) \right) \cdot \exp \left(- \int_0^T \mu + \sum_{t_j < s} \alpha\gamma(s - t_j) ds \right). \end{aligned}$$

For at finde $\mathcal{L}(\theta)$ tages logaritmen til $L(\theta)$.

$$\begin{aligned}\mathcal{L}(\theta) &= \log(L) \\ &= \log \left(\left(\prod_{i=1}^n \mu + \sum_{t_j < t_i} \alpha \gamma(t_i - t_j) \right) \cdot \exp \left(- \int_0^T \left(\mu + \sum_{t_j < s} \alpha \gamma(s - t_j) \right) ds \right) \right) \\ &= \left(\sum_{i=1}^n \log \left(\mu + \sum_{t_j < t_i} \alpha \gamma(t_i - t_j) \right) \right) - \int_0^T \left(\mu + \sum_{t_j < s} \alpha \gamma(s - t_j) \right) ds.\end{aligned}$$

□

[3, s. 183-185, 233-234][5, s. 493-503]

3.5 Simulering

Hvis der haves en model, ønskes det ofte at være i stand til at simulere data fra den. Hvis det ønskes at simulere punktmønstre fra en Poisson punktproces på tidslinien med intensitet λ , er det relativt simpelt via definition 3.15, idet hvert interval simuleres fra $Exp(\lambda)$. Hvis det ønskes, at simulere fra Hawkes punktprocessen anvendes dennes specielle struktur, idet den både kan defineres som en klyngeproces, og via dens betingede intensitetsfunktion.

En simuleringsalgoritme for en umærket Hawkes punktproces på tidslinien er relativ simpel at opskrive, idet der direkte står i definition 3.16, hvordan den genereres. Den består af tre forskellige trin, hvor det første trin simulerer center punkterne, som en homogen Poisson punktproces på tidslinien med intensitet μ . Det andet trin simulerer de forskellige generationer af punkter, hvor hvert punkt afføder nye punkter efter en inhomogene Poisson punktproces med intensitetsfunktion $\alpha \gamma(t - t_i)$. Det endelige tredje trin udgør superpositionen af samtlige punkter. Dette kan ses i algoritme 1.

Algoritme 1 Hawkes punktproces

En umærket Hawkes punktproces $\mathbf{X}$ på intervallet $[0, T]$ kan genereres på følgende måde. Sæt $k=0$

- Generer $\mathbf{X}_k$ som en umærket Poisson punktproces med intensitet μ på intervallet $[0, T]$.

while Antallet af punkter i $\mathbf{X}_k > 0$ **do**

- Lad $k = k + 1$ og generer $\mathbf{X}_k$ som en superposition af $\mathbf{Y}_i \sim \text{Poisson}((t_i, T], \alpha \cdot \gamma(t - t_i))$, hvor $t_i \in \mathbf{X}_{k-1}$ for $i = 1, 2, \dots$

end while

- Lad $\mathbf{X} = \bigcup_{i=0}^k \mathbf{X}_i$.
-

Det er muligt, at implementere algoritme 1 i R, men der er også andre måder at konstruere punktmønstre fra Hawkes punktprocesser. En mulighed er at udnytte den evolutionære

egenskab der opfyldes af punktprocesser på tidslinien. Dette kan gøres ved først at simulere immigranterne som en homogen Poisson punktproces med intensitetsfunktion μ . Herefter startes der ved den nedre grænse i tidsintervallet og hver gang man kommer til en hændelse (t_i) simuleres afkom af denne som en inhomogen Poisson punktproces med intensitetsfunktion $\alpha\gamma(t - t_i)$. Følgende præsenteres en algoritme bygget op på denne måde.

Algoritme 2 Hawkes punktproces

Punktmønstre for en Hawkes punktproces $\mathbf{X}$ intervallet $[0, T]$ kan genereres på følgende måde. Lad $i = 1$

1. Generer $\mathbf{x}$ som en Poisson punktproces med intensitetsfunktion μ på intervallet $[0, T]$.
 - while** $t_i < T$ **do**
 2. Generer $x \sim \text{Poisson}((t_i, T], \alpha\gamma(t - t_i))$.
 3. Lad $\mathbf{x} = \mathbf{x} \cup x$.
 4. Sorter $\mathbf{x}$ i forhold til placeringen på tidslinien.
 5. $i = i + 1$
 - end while**
-

Konstruktionen af immigranterne, punkt 1 i algoritmen, stemmer over ens med punkt 1 i definition 3.16. Hvert punkt afføder nye punkter efter en inhomogen Poisson punktproces med intensitetsfunktion $\alpha\gamma(t - t_i)$, hvilket stemmer overens med punkt 2 i definitionen. Endelig består $\mathbf{x}$ af alle immigranter og afkom i intervallet, hvilket svarer til punkt 3 i definitionen.

Hvis det ikke ønskes at arbejde med inhomogene Poisson punktprocesser i $\mathbb{R}$, er det muligt at simulere punktmønstre fra Hawkes punktprocessen på følgende måde. Hvis der haves en mængde af punkter $\{t_1, t_2, \dots, t_k\}$, gælder der jvf. beviset for lemma 3.4, at

$$\log(1 - F(t_{k+1} | \mathcal{H}_{t_{k+1}})) = - \int_{t_k}^{t_{k+1}} \lambda^*(t) dt.$$

Følgende indsættes den betingede intensitetsfunktion for Hawkes punktprocessen.

$$\log(1 - F(t_{k+1} | \mathcal{H}_{t_{k+1}})) = - \int_{t_k}^{t_{k+1}} \left(\mu + \alpha \sum_{i=1}^k \gamma(t - t_i) \right) dt.$$

Det er nu muligt at finde det næste punkt t_u ved at løse ligningen.

$$\log(U) + \int_{t_k}^{t_u} \left(\mu + \alpha \sum_{i=1}^k \gamma(t - t_i) \right) dt = 0. \quad (3.14)$$

Hvor $U \sim \text{unif}[0, 1]$ idet der gælder at $\log(1 - F(t_{k+1} | \mathcal{H}_{t_{k+1}}))$ er ens i intervallet, da der ikke er nogle punkter, mellem punkterne t_k og t_u , der ændrer på den betingede tæthedsfunktion.

Følgende forlades den generelle umærkede Hawkes punktproces, og tilfælde hvor $\gamma(\cdot) = \exp(-\beta \cdot)$ betragtes. Dermed forlades den generelle Hawkes punktproces og der betragtes følgende en delmængde af disse. Udtrykket for γ indsættes nu i ligning (3.14) og integralet

udregnes.

$$\begin{aligned}
& \log(U) + \int_{t_k}^{t_u} \left(\mu + \alpha \sum_{i=1}^k \exp(-\beta(t - t_i)) \right) dt = 0 \\
& \Leftrightarrow \log(U) + \int_{t_k}^{t_u} \mu dt - \frac{\alpha}{\beta} \int_{t_k}^{t_u} \left(\sum_{i=1}^k -\beta \exp(-\beta(t - t_i)) \right) dt = 0 \\
& \Leftrightarrow \log(U) + \mu(t_u - t_k) - \frac{\alpha}{\beta} \left(\sum_{i=1}^k \exp(-\beta(t_u - t_i)) - \sum_{i=1}^k \exp(-\beta(t_k - t_i)) \right) = 0. \quad (3.15)
\end{aligned}$$

Følgende defineres en funktion $S(k)$ rekursivt.

Definition 3.22.

For et punktmønster $\{t_1, \dots, t_k\}$ gælder der at:

$$S(k) = \begin{cases} 1, & \text{for } k = 1 \\ \exp(-\beta(t_k - t_{k-1}))S(k-1) + 1 & \text{for } k \geq 2 \end{cases}.$$

For $S(k)$ gælder følgende lemma.

Lemma 3.23.

Hvis der haves et punktmønster $\{t_1, \dots, t_k\}$, så gælder følgende

$$-S(k)(1 - \exp(-\beta(t_u - t_k))) = \sum_{i=1}^k \exp(-\beta(t_u - t_i)) - \sum_{i=1}^k \exp(-\beta(t_k - t_i)). \quad (3.16)$$

Hvor $S(k)$ er defineret som i definition 3.22.

Bevis. Lemma 3.23 vil blive bevist via induktion over k .

Basis: Lad $k = 1$. Der gælder da følgende

$$\begin{aligned}
-S(1)(1 - \exp(-\beta(t_2 - t_1))) &= \exp(-\beta(t_2 - t_1)) - 1 \\
&= \exp(-\beta(t_2 - t_1)) - \exp(-\beta(t_1 - t_1)) \\
&= \sum_{i=1}^1 \exp(-\beta(t_{k+1} - t_i)) - \sum_{i=1}^1 \exp(-\beta(t_k - t_i)).
\end{aligned}$$

Antagelse Antag at ligning (3.16) er sand for $k - 1$ skridt.

Induktion Det ønskes nu at vise at ligning (3.16) også er sand for k skridt.

$$\begin{aligned}
& -S(k)(1 - \exp(-\beta(t_u - t_k))) \\
&= (-\exp(-\beta(t_k - t_{k-1}))S(k-1) - 1)(1 - \exp(-\beta(t_u - t_k))) \\
&= -\exp(-\beta(t_k - t_{k-1}))S(k-1) \\
&\quad + \exp(-\beta(t_k - t_{k-1}))\exp(-\beta(t_u - t_k))S(k-1) \\
&\quad - 1 + \exp(-\beta(t_u - t_k)) \\
&= -\exp(-\beta(t_k - t_{k-1}))S(k-1) \\
&\quad + \exp(t_u - t_{k-1})S(k-1) \\
&\quad - 1 + \exp(-\beta(t_u - t_k)) \\
&= -\exp(-\beta(t_k - t_{k-1}))S(k-1) \\
&\quad + \exp(t_u - t_{k-1})S(k-1) - S(k-1) + S(k-1) \\
&\quad - 1 + \exp(-\beta(t_u - t_k)) \\
&= -\exp(-\beta(t_k - t_{k-1}))S(k-1) + S(k-1) \\
&\quad + (\exp(t_u - t_{k-1}) - 1)S(k-1) \\
&\quad - 1 + \exp(-\beta(t_u - t_k)) \\
&= -\sum_{i=1}^{k-1} \exp(-\beta(t_k - t_i)) + \sum_{i=1}^{k-1} \exp(-\beta(t_{k-1} - t_i)) \\
&\quad + \sum_{i=1}^{k-1} \exp(-\beta(t_u - t_i)) - \sum_{i=1}^{k-1} \exp(-\beta(t_{k-1} - t_i)) \\
&\quad - 1 + \exp(-\beta(t_u - t_k)) \\
&= -\sum_{i=1}^k \exp(-\beta(t_k - t_i)) + \sum_{i=1}^{k-1} \exp(-\beta(t_{k-1} - t_i)) \\
&\quad + \sum_{i=1}^k \exp(-\beta(t_u - t_i)) - \sum_{i=1}^{k-1} \exp(-\beta(t_{k-1} - t_i)) \\
&= \sum_{i=1}^k \exp(-\beta(t_u - t_i)) - \sum_{i=1}^k \exp(-\beta(t_k - t_i))
\end{aligned}$$

□

Det er muligt, at omskrive ligning (3.15) ved brug af $S(k)$ fra definition 3.22 og lemma 3.23.

$$\begin{aligned}
& \log(U) + \mu(t_u - t_k) - \frac{\alpha}{\beta} \left(\sum_{i=1}^k \exp(-\beta(t_u - t_i)) - \sum_{i=1}^k \exp(-\beta(t_k - t_i)) \right) = 0, \\
& \Leftrightarrow \log(U) + \mu(t_u - t_k) - \frac{\alpha}{\beta} S(k) (1 - \exp(-\beta(t_u - t_k))) = 0. \quad (3.17)
\end{aligned}$$

Dermed er det muligt, at opskrive en algoritme der genererer punktmønstre for en umærket Hawkes punktproces med $\gamma(\cdot) = \exp(-\beta\cdot)$.

Algoritme 3 Hawkes punktproces

En umærket Hawkes punktproces $\mathbf{X}$ på intervallet $[0, T]$ kan genereres på følgende måde.

$k=1$

- Lad $t_k \sim \text{Exp}(\mu)$.
 - **while** $t_k < T$ **do**
 - $k = k + 1$
 - Generer $U \sim \text{unif}[0, 1]$
 - Bestem u der løser ligning (3.17)
 - $t_k = u$
 - **end while**
-

Løsningen til ligning (3.17) findes med en 1-dimensionel Newton-Rhapson algoritme, jvf. afsnit 3.3.3. Begyndelsestilstanden vælges til det forrige punkt adderet med $\text{Exp}(\mu)$. [10, s. 148-149]

3.6 Modelkontrol

Når der er konstrueret en model, der kan lave punktmønstre, er det relevant at undersøge om denne model overhovedet passer til virkeligheden. Dette kan gøres ved at undersøge om den er i stand til at lave punktmønstre, der minder om de data der haves fra virkeligheden.

En af måderne, hvorpå der ofte udføres modelkontrol er ved at konstruere et histogram for de empiriske data, og sammenligne den med en kurve for den teoretiske fordeling for den konstruerede model. Det kan dog være vanskeligt at se problemer i et sådant histogram, og derfor bruges der ofte q-q plots. Et q-q plot er en punktplotning, hvor kvantilerne for de empiriske data plottes imod kvantilerne for den teoretisk fordeling for modellen.

Når der arbejdes med punktprocesser er det muligt at omforme de givne data, således at de kan holdes op imod en kendt teoretisk fordeling. Omformningen af de givne data foregår med følgende sætning.

Sætning 3.24.

Lad $\{t_i\}_{i \in \mathbb{Z}}$ være en punktproces med intensitet $\lambda^*(t_i)$, og lad $s_i = \Lambda^*(t_i)$. Så er $\{s_i\}_{i \in \mathbb{Z}}$ en Poisson punktproces med intensitet 1.
[3, s. 261]

Hvis modellen er valgt, og har en kendt betinget intensitetsfunktion, er det dermed muligt at omforme de givne data til en form, hvor de kan plottes imod den eksponentielle fordeling med intensitet 1. Hvis en sådan plotning følger identitetslinien, giver det en indikation af at modellen er valgt til at modellere data, der minder om det givne datasæt.

Der er mange andre måder at foretage modelkontrol på, men dette er den eneste, der vil blive anvendt i denne rapport.

[3, s. 261].

I forbindelse med dette afsnit er der lavet en del udregninger/billeder i R. Scripts for dette kan ses i appendiks 11.6.1.

Det ønskes at konstruere en model for jordskælvshændelser i Danmark. For at få en model der passer på virkeligheden, tages der udgangspunkt i et virkeligt datasæt for området omkring Danmark. Følgende afsnit præsenterer datasættet, samt enkelte indledende undersøgelser.

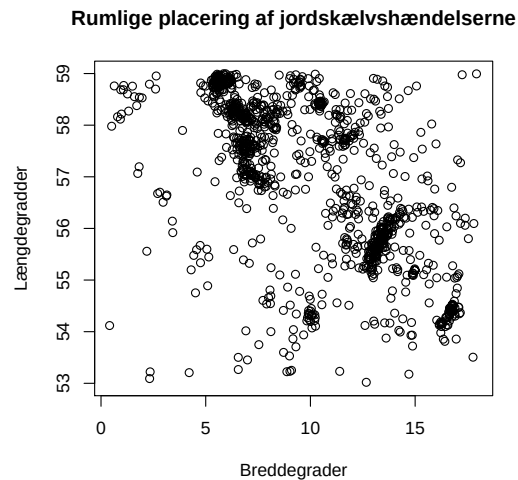
4.1 Datasæt

Det datasæt, der arbejdes med i dette projekt, er et datasæt over jordskælvshændelser i området omkring Danmark. Datasættet er hentet fra *De Nationale Geologiske Undersøgelser for Danmark og Grønland* (GEUS) [2, Registrerede jordskælv]. Datasættet består af alle registrerede jordskælv i et vindue omkring Danmark, der strækker sig fra 0,41 til 17,93 breddegrader og 53,02 til 58,997 længdegrader. Formen på vinduet er jeg ikke bekendt med, men den rumlige placering kan ses i figur 4.1. Da der udelukkende betragtes punktprocesser på tidslinien er formen af observationsvinduet alligevel irrelevant.

Datasættet indeholder følgende information om hændelserne.

- Tidspunktet for hændelsen, fordelt på årstallet, måneden, dagen, timen, minuttet, sekundet.
- Placeringen af hændelsen, bredde- og længdegrader.
- Dybden af hændelsen målt i kilometer.
- Styrken af hændelsen målt på Richter-skalaen.

Datasættet strækker sig fra år 1930 til og med år 2009, og består af 1187 registrerede jordskælv. Det kraftigste jordskælv er målt i 1981 til 5,2 på Richter-skalaen. Til sammenligning er det svageste målt til 0,3 i 2009. Et udpluk af datasættet kan ses i tabel 4.1. De generelle data for datasættet er fundet med R-script 11.4.



Figur 4.1: Den rumlige placering af jordskælvshændelserne. Billedet er konstrueret med R-script 11.3.

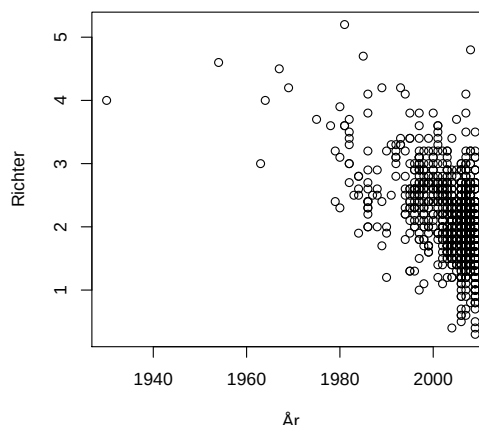
Tabel 4.1: Overblik over jordskælvsdata

#	Aar	Maaned	Dag	Time	Minut	Sekund	NORD	OEST	dybde	Richter
1	1930	10	31	23	16	38,9	55,665	12,505	62,3	4,0
2	1954	10	18	16	44	32,2	56,846	8,291	25,5	4,6
3	1960	1	19	0	5	0,0	58,500	6,500	15,0	NA
4	1963	6	19	17	0	19,8	56,270	11,740	0,0	3,0
5	1964	7	14	5	33	56,7	57,03	7,2	36	4,0
6	1967	8	21	13	41	49,3	57,092	4,593	15,0	4,5
7	1969	4	5	19	9	47,1	57,155	6,764	0,1	4,2
8	1975	11	12	00	0,6	19,3	57,102	7,141	40,0	3,7
9	1978	4	26	12	32	40,1	57,447	7,086	15,0	3,6
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
53	1989	11	23	14	12	37,7	58,123	7,702	14,9	1,7
54	1990	5	9	16	47	32,0	57,650	12,535	37,7	1,9
55	1990	5	24	9	51	56,9	56,478	11,931	10,0	3,2
56	1990	5	28	8	43	48,8	55,044	13,003	15,0	1,2
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

4.2 Udpluk af data

Ud fra datasættet bemærkes det, at antallet af jordskælv i observationsvinduet ser ud til at være drastigt forøget op igennem tiden, jvf. figur 4.2. Dette skyldes ikke, at der forekommer flere jordskælv i dag end tidligere, jvf. afsnit 2.1. Det skyldes derimod, at der er blevet opstillet flere målestationer, og disse er blevet væsentligt bedre til at måle hændelserne. Det giver derfor mening, at undlade en del af de tidligere data, for at undgå en ændring i intensiteten grundet flere og bedre målestationer.

Det vælges derfor kun, at betragte de jordskælv der er forekommet fra år 2006 til og med år 2009. Dermed er der fjernet 468 jordskælv, og der er således 719 tilbage. For yderligere at mindske risikoen for at forbedringen af opmålingerne afspejles i de senere analyser, sorteres alle jordskælv med en styrke på mindre end 2 på Richter-skalaen fra. Dermed vil



Figur 4.2: Tidspunktet (året) og styrken af de registrerede jordskælv i Danmark. Bemærk en hvis diskretisering idet det blot er året for hændelsen der er medtaget. Konstrueret med R-script 11.5

en eventuel forbedring af målestationerne, til opmåling af de svageste jordskælv, ikke blive afspejlet i den udvalgte data. I tidsperioden fra 2006 til og med 2009 er der registreret 354 jordskælv med en styrke på 2,0 eller mere på Richter-skalaen. Denne udvælgelse af datasættet foretages med R-script 11.6.

For at give en ide om, at der i det valgte tidsinterval, ikke optræder for mange jordskælvs-hændelser til sidst, grundet bedre målestationer, sammenlignes tælle-målet for de halvårslige intervaller. Dette kan ses i tabel 4.2. Det ses dermed, at der ikke umiddelbart er nogen grund til at tro, at der er nogen problemer med bedre udstyr i tidsperioden fra 2006 til og med 2009.

Tabel 4.2: Kontrol af de relevante jordskælvshændelser. I angiver det første halvår mens II angiver det andet. Konstrueret med R-script 11.8

	2006 I	2006 II	2007 I	2007 II	2008 I	2008 II	2009 I	2009 II
Tælle-mål	46	45	54	38	48	34	42	47

4.3 Fejl og mangler

I datasættet er der enkelte mangler, altså steder hvor der ikke er angivet nogen værdi, som f.eks. tredje række sidste søjle i tabel 4.1. Der er dog ikke nogen steder i den udvalgte del af datasættet hvor der mangler data. Yderligere vil der være en del afrunding, til f.eks. første decimal i søjlen for Richter værdien.

Der vil endvidere kunne være direkte fejl i datasættet. Sådanne fejl er oftest meget svære at finde, da der ofte ikke er nogen måde at kontrollere det modtagne datasæt. Det er kun muligt at lokalisere sådanne fejl, hvis den antagne værdi ligger uden for det mulige.

4.4 Fordelingen af punkterne

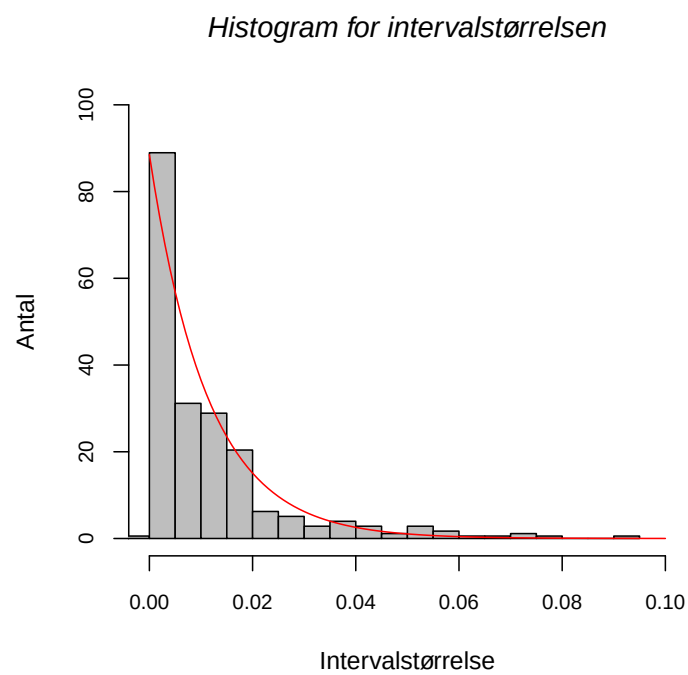
Inden det er muligt at vælge hvilken type af model det ønskes at tilpasse til datasættet, er det nødvendigt at gøre nogle indledende iagttagelser. Årsagen til dette er, at forskellige modeller er i stand til at simulere punktmønstre med forskellige tendenser. F.eks. er der nogle modeltyper, der ikke virker til at modellere punktmønstre, hvor fordelingen af punkterne er meget regulær. Derfor undersøges det om de udvalgte data har en tendens til at klumpe, eller om de er mere regulære end en Poisson punktproces.

Det vælges at betragte fordelingen af ventetiden. For at kunne gøre dette er det nødvendigt at have ét tal for hvornår hændelserne indtræffer. Da udplukket af datasættet starter den 1. Januar 2006 klokken 00.00.00 vælges dette som begyndelsestidspunktet ($t_0 = 0$), og sluttidspunktet bliver 31. December klokken 23.59.59 år 2009. Det er dermed muligt at konstruere en ny variabel, der består af antallet af år, der er gået siden starttidspunktet. Her er det væsentligt at huske at år 2008 er et skudår¹, og dette år er derfor en dag længere end de andre. Tidsintervallet bliver således $I = [0, T]$ hvor $T = 4$. Der regnes derfor gennemgående i projektet med en årslængde på 365,24219 dage.

Det er nu muligt at betragte fordelingen af ventetiderne. For en homogen Poisson punktproces vil fordelingen af ventetiden, mellem to hændelser, være givet ved en eksponentiel fordeling med en intensitet på $\mu = \frac{354}{T}$. På figur 4.3, er det muligt at se histogrammet for de empiriske tidsintervallerne, samt det teoretiske tæthed for eksponentialfordelingen med parametren μ . Det lader umiddelbart til at der blandt de empiriske data er en overvægt af korte intervaller, og en overvægt af de helt lange intervaller. Dette antyder at der er en større sammenklumpning end for en Poisson punktproces.

Det er dog værd at nævne, at der ikke er en voldsom afvigelse fra Poisson punktprocessen med intensitet μ .

¹Der er skudår alle de år der går op i 4, men ikke i 100. Hvis årstallet går op i 400 er det dog skudår alligevel.



Figur 4.3: Histogram for de empiriske kvantiler for intervalstørrelsen, samt den teoretiske tæthedsfunktion for den eksponentielle fordeling med parametren μ . Konstrueret med R-script 11.9.

KAPITEL 5

DEN UMÆRKEDE HAWKES MODEL

I forbindelse med dette afsnit er der lavet en del udregninger/billeder i R. Scripts for dette kan ses i appendiks 11.6.2.

Det ønskes, at finde en model for jordskælvshændelserne i Danmark. Modellen ønskes konstrueret på tidslinien, således at den giver information om hvor mange jordskælv der vil komme indenfor et givent tidsinterval. Der blev i afsnit 2.1 redegjort for at jordskælv har en mulighed for at danne efterskælv, hvor sandsynligheden for sådanne er størst lige efter hændelsen, og at der derfor vil være en hvis sammenklumpning af jordskælvshændelser. Hvilket passer godt med, at der i afsnit 4.4, blev der redegjort for at hændelserne i datasættet er mindre regulær end en Poisson punktproces på tidslinien. Ydermere blev der redegjort for at der også kan forekomme forskælv, dette er ikke umiddelbart noget jeg vælger at medtage i modellen. Yderligere vil jordskælv, der ikke er efterskælv, forekomme med en grundintensitet der ikke er påvirket af tidspunktet for hændelsen.

Dermed er det muligt, at anvende en Hawkes punktproces som model for jordskælvshændelserne, idet denne kan anvendes som model for punktmønstre, der danner klynger. Yderligere gælder der, jvf. definition 3.17, at den betingede intensitetsfunktion er givet ved ligning (3.10). Her vil μ være grundintensiteten af jordskælvshændelserne og intensiteten for efterskælv ved tidspunktet t , for hændelsen t_i , vil være givet ved $\alpha\gamma(t - t_i)$. Da det ønskes at intensitetsfunktionen for efterskælv skal være faldende når tidsintervallet øges, vælges det at $\gamma(\cdot) = \exp(-\beta(\cdot))$, således bliver:

$$\lambda^*(t) = \mu + \alpha \sum_{t_i < t} \exp(-\beta(t - t_i)). \quad (5.1)$$

Dermed er der tre ukendte parametre i den betingede intensitetsfunktion, nemlig μ , α og β . For at kunne konstruere modellen er det nødvendigt at bestemme disse tre parametre.

5.1 Likelihoodfunktionen for den umærkede model

Det ønskes at anvende MLE til at bestemme μ , α og β . Derfor bestemmes log-likelihood-funktionen for en Hawkes punktproces på intervallet $I = [0, T]$, med en betinget intensitetsfunktion som den i ligning (5.1). Den generelle log-likelihoodfunktion for en Hawkes punktproces er kendt fra sætning 3.21, og det er nok at indsætte $\exp(-\beta(\cdot))$ på $\gamma(\cdot)$ plads.

$$\begin{aligned}\mathcal{L}(\theta) &= \sum_{i=1}^n \left(\log \left(\mu + \sum_{t_j < t_i} \alpha \gamma(t_i - t_j) \right) \right) - \int_0^T \left(\mu + \sum_{t_j < s} \alpha \gamma(s - t_j) \right) ds \\ &= \sum_{i=1}^n \left(\log \left(\mu + \sum_{t_j < t_i} \alpha \exp(-\beta(t_i - t_j)) \right) \right) - \int_0^T \left(\mu + \sum_{t_j < s} \alpha \exp(-\beta(s - t_j)) \right) ds \\ &= \sum_{i=1}^n \log(\mu + \alpha A(i)) - \int_0^T \left(\mu + \sum_{t_j < s} \alpha \exp(-\beta(s - t_j)) \right) ds, \quad (5.2)\end{aligned}$$

hvor

$$A(i) = \begin{cases} 0, & \text{for } i = 1 \\ \left(\sum_{t_j < t_i} \exp(-\beta(t_i - t_j)) \right) & \text{for } i \geq 2 \end{cases}.$$

Der gælder endvidere:

$$\begin{aligned}\int_0^T \left(\mu + \left(\sum_{t_j < s} \alpha \exp(-\beta(s - t_j)) \right) \right) ds &= \int_0^T \mu ds + \int_0^T \left(\sum_{t_j < s} \alpha \exp(-\beta(s - t_j)) \right) ds \\ &= \mu T + \left(\sum_{i=1}^n \alpha \left(\frac{\exp(\beta(0))}{\beta} - \frac{\exp(-\beta(T - t_i))}{\beta} \right) \right) \\ &= \mu T + \left(\sum_{i=1}^n \alpha \frac{1 - \exp(-\beta(T - t_i))}{\beta} \right) \quad (5.3)\end{aligned}$$

Det er nu muligt at forsimple udtrykket for $\mathcal{L}(\theta)$ yderligere, hvis ligning (5.3) indsættes i ligning (5.2).

$$\mathcal{L}(\theta) = \left(\sum_{i=1}^n \log(\mu + \alpha A(i)) \right) - \mu T + \left(\sum_{i=1}^n \frac{\alpha}{\beta} (\exp(-\beta(T - t_i)) - 1) \right). \quad (5.4)$$

5.2 Parameter estimation

Da der er tre ukendte parametre der skal bestemmes, er det nødvendigt, at løse et lignings-system bestående af tre ligninger og tre ubekendte.

$$\frac{\partial}{\partial \mu} \mathcal{L}(\theta) = 0, \quad \frac{\partial}{\partial \alpha} \mathcal{L}(\theta) = 0, \quad \frac{\partial}{\partial \beta} \mathcal{L}(\theta) = 0,$$

hvor $\theta = \{\mu, \alpha, \beta\}$.

For at løse et sådant ligningssystem anvendes en tre-dimensionel Newton-Rapson.

$$\theta^{(n)} = \theta^{(n-1)} - H^{-1} \Delta(\theta^{(n-1)}).$$

Det er derfor nødvendigt, at bestemme vektoren bestående af de partielt afledte $\Delta(\mathcal{L}(\theta))$ samt Hessianmatricen, altså:

$$\Delta(\mathcal{L}(\theta)) = \begin{bmatrix} \frac{\partial}{\partial \mu} \mathcal{L}(\theta) \\ \frac{\partial}{\partial \alpha} \mathcal{L}(\theta) \\ \frac{\partial}{\partial \beta} \mathcal{L}(\theta) \end{bmatrix}, \quad H(\mathcal{L}(\theta)) = \begin{bmatrix} \frac{\partial^2 \mathcal{L}(\theta)}{\partial \mu^2} & \frac{\partial^2 \mathcal{L}(\theta)}{\partial \mu \partial \alpha} & \frac{\partial^2 \mathcal{L}(\theta)}{\partial \mu \partial \beta} \\ \frac{\partial^2 \mathcal{L}(\theta)}{\partial \alpha \partial \mu} & \frac{\partial^2 \mathcal{L}(\theta)}{\partial \alpha^2} & \frac{\partial^2 \mathcal{L}(\theta)}{\partial \alpha \partial \beta} \\ \frac{\partial^2 \mathcal{L}(\theta)}{\partial \beta \partial \mu} & \frac{\partial^2 \mathcal{L}(\theta)}{\partial \beta \partial \alpha} & \frac{\partial^2 \mathcal{L}(\theta)}{\partial \beta^2} \end{bmatrix}.$$

Først findes de partielt afledte af $\mathcal{L}(\theta)$. Udregningerne er ret lange og er derfor flyttet til appendiks 11.4.

$$\begin{aligned} \frac{\partial}{\partial \mu} \mathcal{L}(\theta) &= \left(\sum_{i=1}^n \frac{1}{\mu + \alpha A(i)} \right) - T. \\ \frac{\partial}{\partial \alpha} \mathcal{L}(\theta) &= \left(\sum_{i=1}^n \frac{A(i)}{\mu + \alpha A(i)} \right) + \left(\sum_{i=1}^n \frac{1}{\beta} (\exp(-\beta(T - t_i)) - 1) \right). \\ \frac{\partial}{\partial \beta} \mathcal{L}(\theta) &= \left(\sum_{i=1}^n \frac{-\alpha B(i)}{\mu + \alpha A(i)} \right) - \alpha \left(\sum_{i=1}^n \frac{\exp(-\beta(T - t_i)) - 1}{\beta^2} + \frac{(T - t_i) \exp(-\beta(T - t_i))}{\beta} \right), \end{aligned}$$

hvor

$$B(i) = \begin{cases} 0, & \text{for } i = 1 \\ \left(\sum_{t_j < t_i} (t_i - t_j) \exp(-\beta(t_i - t_j)) \right) & \text{for } i \geq 2 \end{cases}$$

Dermed er det muligt, at bestemme vektoren med de partielt afledte. For endvidere at bestemme Hessian-matricen, er det nødvendigt at kende de andenordens afledte. Derfor findes de forskellige andenordens partielt afledte af log-likelihoodfunktionen til Hawkes punktprocessen fra ligning (5.4). Først findes de dobbelt afledte af μ , α og β . Da udregningerne er ret lange er de derfor flyttet til appendiks 11.4.

$$\begin{aligned}
\frac{\partial}{\partial \mu^2} \mathcal{L}(\theta) &= \sum_{i=1}^n \frac{-1}{(\mu + \alpha A(i))^2} \\
\frac{\partial}{\partial \alpha^2} \mathcal{L}(\theta) &= \sum_{i=1}^n \frac{-A(i)^2}{(\mu + \alpha A(i))^2} \\
\frac{\partial}{\partial \beta^2} \mathcal{L}(\theta) &= \left(\sum_{i=1}^n \frac{\alpha C(i)}{\mu + \alpha A(i)} - \left(\frac{\alpha B(i)}{(\mu + \alpha A(i))} \right)^2 \right) + \alpha \left(\sum_{i=1}^n \frac{2(\exp(-\beta(T - t_i)) - 1)}{\beta^3} \right. \\
&\quad \left. + \frac{2(T - t_i) \exp(-\beta(T - t_i))}{\beta^2} + \frac{(T - t_i)^2 \exp(-\beta(T - t_i))}{\beta} \right),
\end{aligned}$$

hvor

$$C(i) = \begin{cases} 0, & \text{for } i = 1 \\ \left(\sum_{t_j < t_i} (t_i - t_j)^2 \exp(-\beta(t_i - t_j)) \right) & \text{for } i \geq 2 \end{cases}.$$

Da der gælder, at

$$\begin{aligned}
\frac{\partial}{\partial \alpha \partial \mu} \mathcal{L}(\theta) &= \frac{\partial}{\partial \mu \partial \alpha} \mathcal{L}(\theta), \\
\frac{\partial}{\partial \beta \partial \mu} \mathcal{L}(\theta) &= \frac{\partial}{\partial \mu \partial \beta} \mathcal{L}(\theta), \\
\frac{\partial}{\partial \beta \partial \alpha} \mathcal{L}(\theta) &= \frac{\partial}{\partial \alpha \partial \beta} \mathcal{L}(\theta),
\end{aligned}$$

er det nok at finde ligningerne til højre. Da udregningerne er ret lange er de derfor flyttet til appendiks 11.4.

$$\begin{aligned}
\frac{\partial}{\partial \mu \partial \alpha} \mathcal{L}(\theta) &= \sum_{i=1}^n \frac{-A(i)}{(\mu + \alpha A(i))^2} \\
\frac{\partial}{\partial \mu \partial \beta} \mathcal{L}(\theta) &= \sum_{i=1}^n \frac{\alpha B(i)}{(\mu + \alpha A(i))^2} \\
\frac{\partial}{\partial \alpha \partial \beta} \mathcal{L}(\theta) &= \left(\sum_{i=1}^n \frac{-B(i)}{\mu + \alpha A(i)} + \frac{\alpha A(i) B(i)}{(\mu + \alpha A(i))^2} \right) \\
&\quad - \left(\sum_{i=1}^n \frac{1}{\beta^2} (\exp(-\beta(T - t_i)) - 1) + \frac{1}{\beta} (T - t_i) \exp(-\beta(T - t_i)) \right).
\end{aligned}$$

Dermed er Hessian-matricen også bestemt.

Før det er muligt at finde den inverse til Hessian-matricen, er det nødvendigt at sikre sig at den findes. Det vil sige, at det er nødvendigt at kontrollere om determinanten er lig 0. Hvis den er dette, er matricen singulær og har dermed ingen invers.

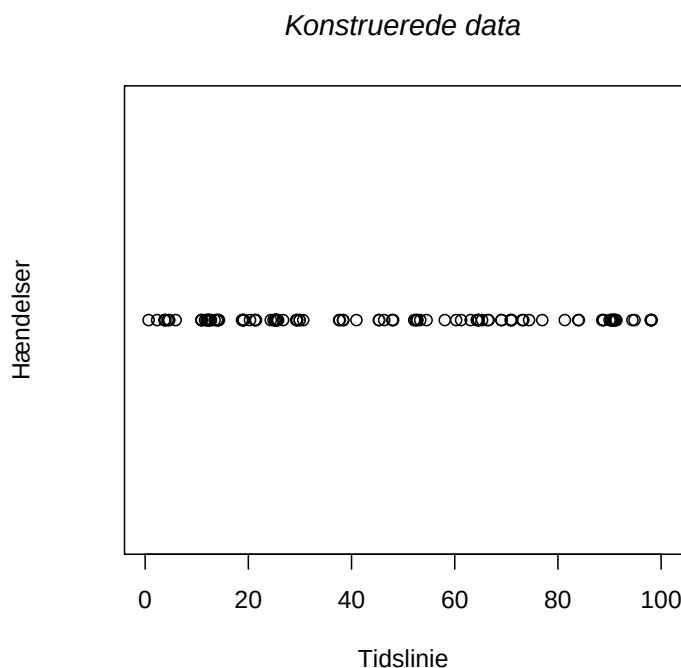
Da både $H(\mathcal{L}(\theta))$ og $\Delta(\theta)$ er kendt, er det nu muligt at approksimere de værdier der maksimerer log-likelihoodfunktionen og dermed likelihoodfunktionen med en tre-dimensionel Newton-Rhapon algoritme fra afsnit 3.3.3. Inden algoritmen anvendes til at approksimere parametrene for de udvalgte jordskælvs hændelserne i Danmark, vælges det at afprøve den

på et datasæt konstrueret med algoritme 3. Således at det er muligt, at se om algoritmen giver noget der tilnærmelsesvis er rigtigt. Det vælges at simulere på intervallet $[0, 100]$. I tabel 5.1 er det muligt, at se de valgte værdier for det konstruerede datasæt, samt de approksimerede værdier med den tre-dimensionelle Newton-Rhapon algoritme med $H(\mathcal{L}(\theta))$ og $\Delta(\theta)$ angivet som ovenfor. Det er nødvendigt med startværdier til algoritmen. I dette tilfælde er det valgt at tage udgangspunkt i det halve af de *korrekte* værdier der er brugt til konstruktionen af datasættet.

Tabel 5.1: Kontrol af tre-dimensionel Newton-Rhapon på et konstrueret datasæt. R-script for konstruktionen af det kunstige datasæt samt approksimationerne kan ses i R-script 11.12.

	μ	α	β
Korrekte parametre	0.5	5	10
Gæt	0.25	2.5	5
Approximation (1 skridt)	0,41	3,37	5,65
Approximation (5 skridt)	0,61	4,09	7.82
Approximation (10 skridt)	0,61	4,09	7.82
Approximation (20 skridt)	0,61	4,09	7.82

På figur 5.1, er det muligt at se det konstruerede datasæt, hvor der er 122 punkter. Det bemærkes at de estimerede parametre afviger lidt fra de *korrekte* parametre. Dette er ikke så overraskende, idet det simulerede datasæt for en Hawkes punktproces let kan være atypisk for de givne parametre, idet der er en stor variation af antallet af punkter i klyngerne, jvf. [12, s. 11].



Figur 5.1: Det konstruerede datasæt. R-script for plotningen kan ses i R-script 11.13

Det ønskes følgende at anvende Newton-Rhapson algoritmen til at approksimere parametrene for de relevante jordskælv. Derfor er det nødvendigt at bestemme nogle startværdier for parametrene. Hvis alle punkterne var tilfældige punkter, der var ankommet med en Poisson punktproces, vil μ være lig antallet af punkter delt med størrelsen på tidsintervallet, jvf. eksempel 3.12. Derfor vælges μ som $\frac{354}{T} = \frac{354}{4}$. Den reelle værdi af μ burde være mindre end dette, idet det er en rimelig antagelse, at nogle af punkterne vil være afkom af andre punkter. Startværdien af α og β bliver derimod valgt noget mere tilfældigt. Det ønskes at vælge en positiv startværdi. De vælges således at $\alpha = 200$ og $\beta = 800$. Umiddelbart burde gættet være mere eller mindre irrelevant, hvis bare der foretages nok iterationer med algoritmen, men dette er dog ikke helt tilfældet. Der kan opstå problemer, idet likelihoodfunktionen kan være meget flad. Derfor er det nødvendigt at komme med et gæt, der ligger tilstrækkelig tæt på $\hat{\theta}$ til at algoritmen konvergerer imod det rigtige. På tabel 5.2 er det muligt at se gættet på startværdierne, samt værdierne approksimeret med algoritmen ved 1, 5, 10 og 20 iterationer.

Tabel 5.2: Parameterestimation med tre-dimensionel Newton-Rhapson. R-script for parameterestimationen kan ses i R-script 11.14

	μ	α	β
Gæt	$\frac{345}{4}$	200	800
Approksimation (1 skridt)	63,28	159,94	738,57
Approksimation (5 skridt)	69,63	188,61	878,38
Approksimation (10 skridt)	69,63	188,61	878,38
Approksimation (20 skridt)	69,63	188,61	878,38

Det antages at $\{\mu^{(20)}, \alpha^{(20)}, \beta^{(20)}\} = \{\hat{\mu}, \hat{\alpha}, \hat{\beta}\} = \hat{\theta}$. Disse værdier kan indsættes i den betingede intensitetsfunktion fra ligning (5.1).

$$\lambda^*(t) = \hat{\mu} + \hat{\alpha} \sum_{t_i < t} \exp(-\hat{\beta}(t - t_i)).$$

Den første kontrol af de estimerede parametre, er at indsætte dem i de partielt afledte til log-likelihoodfunktionen for at se om de giver 0 eller noget tæt på. Script til udregningerne er R-script 11.15.

$$\Delta(\mathcal{L}(\hat{\theta})) = \begin{bmatrix} \frac{\partial}{\partial \mu} \mathcal{L}(\hat{\theta}) \\ \frac{\partial}{\partial \alpha} \mathcal{L}(\hat{\theta}) \\ \frac{\partial}{\partial \beta} \mathcal{L}(\hat{\theta}) \end{bmatrix} \approx \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Dermed gælder der, at de approksimerede parametre er meget tæt på at maksimere likelihoodfunktionen.

Hvis det nu ønskes, at finde det gennemsnitlige antal afkom, er det muligt på baggrund af de estimerede parametre. Betragt en hypotetisk punkt der starter i tidspunktet 0. Så vil det

gennemsnitlige antal afkom for dette punkt være givet ved integralet fra 0 til ∞ .

$$\begin{aligned}
 \int_0^{\infty} \hat{\alpha} \exp(-\hat{\beta}(s-0)) ds &= -\frac{\hat{\alpha}}{\hat{\beta}} \int_0^{\infty} -\hat{\beta} \exp(-\hat{\beta}(s)) ds \\
 &= -\frac{\hat{\alpha}}{\hat{\beta}} (\lim_{s \rightarrow \infty} (\exp(\hat{\beta}(s))) - \exp(-\hat{\beta} \cdot 0)) \\
 &= -\frac{\hat{\alpha}}{\hat{\beta}} (0 - 1) \\
 &= \frac{\hat{\alpha}}{\hat{\beta}} \\
 &\approx \frac{188,61}{878,38} \\
 &\approx 0,21.
 \end{aligned}$$

5.3 Modelkontrol

Selvom der er estimeret parametre, der maksimerer likelihoodfunktionen er det dog stadig ikke sikkert, at modellen kan simulere data, der minder om det givne datasæt for de udvalgte jordskælvshændelser omkring Danmark. Dette skyldes, at selve den valgte modeltype kan passe dårligt til virkeligheden. Hvis man f.eks. prøvede at tilpasse en model baseret på en Hawkes punktproces til et meget regulært datasæt, vil det stadig være muligt at estimere de bedste parametre, men modellen vil stadig ikke passe særlig godt på virkeligheden. Det er derfor nødvendigt at kontrollere om den valgte model passer på virkeligheden.

Modellen vil blive kontrolleret ved hjælp af sætning 3.24. Dermed gælder der, at hvis modellen er velvalgt vil $\{s_i\} = \Lambda^*(t_i)$ for $i = 1, \dots, 354$ være fordelt som en Poisson punktproces med intensitet 1, altså bør $\{s_i\}$ have interval længder trukket fra en den eksponentielle fordeling med parametren 1. Først konstrueres $\Lambda^*(t)$ ved hjælp af definition 3.9.

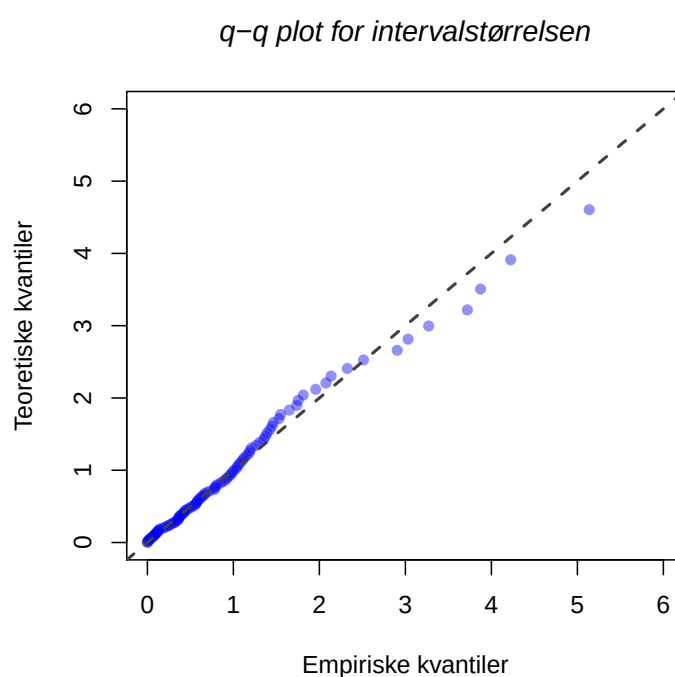
$$\begin{aligned}
 \Lambda^*(t) &= \int_0^t \left(\mu + \sum_{t_i < s} \alpha \exp(-\beta(s-t_i)) \right) ds \\
 &= \mu t - \frac{\alpha}{\beta} \int_0^t \left(\sum_{t_i < s} -\beta \exp(-\beta(s-t_i)) \right) ds \\
 &= \mu t - \frac{\alpha}{\beta} \left(\sum_{t_i < t} \exp(-\beta(t-t_i)) - 1 \right) \\
 &= \mu t + \frac{\alpha}{\beta} \left(\sum_{t_i < t} 1 - \exp(-\beta(t-t_i)) \right)
 \end{aligned}$$

Det bemærkes at ovenstående udregner er de samme som i ligning (5.3).

Følgende indsættes MLE.

$$\Lambda^*(t) = \hat{\mu}t - \frac{\hat{\alpha}}{\hat{\beta}} \sum_{t_i < t} \exp(-\hat{\beta}(t - t_i))\hat{\mu}t + \frac{\hat{\alpha}}{\hat{\beta}} \left(\sum_{t_i < t} 1 - \exp(-\hat{\beta}(t - t_i)) \right)$$

Mængden $\{s_j\} = \Lambda^*(t_j)$, for $j = 1, \dots, 354$ konstrueres og kvantilerne for intervallængderne sammenlignes med kvantilerne for intervallængderne for en Poisson punktproces med intensitet 1. Dette kan ses på figur 5.2, hvor det fremgår at fordelingen af de transformerede data ser ud til at passe nogenlunde på en Poisson punktproces med intensitet 1. Dermed er der grund til at mene, at modellen er en nogenlunde repræsentation af virkeligheden.



Figur 5.2: q-q plot for det transformerede datasæt. R-script for konstruktionen af s og q-q plottet kan ses i R-script 11.16.

KAPITEL 6

MÆRKEDE PUNKTPROCESSER PÅ TIDSLINIEN

Hvis der gives en punktproces $\mathbf{Y}$ kan det ofte være ønskværdigt, at tilføje en *ekstra* viden til punkterne. Dette kan gøres ved hjælp af en mærket punktproces. Dermed er den mærkede punktproces $\mathbf{X}$ givet ved $\mathbf{X} = \{(t_i, k_i) | t_i \in \mathbf{Y}\}$, hvor k_i er mærket for punktet t_i tilhørende mærkerummet M . For en mærket punktproces på tidslinien er grundintensiteten, $\lambda_g(t)$, den punktproces der genererer placeringen af punkterne på tidslinien. Mærkerne kan bruges til at tilføje viden om f.eks. styrken af et jordskælv, eller det kunne være en geometrisk figur, der illustrerer størrelsen af objekterne eller lignende.

Der er to forskellige måder at betragte mærkede punktprocesser på tidslinien på. Den første er at betragte dem som punktprocesser med udfaldsrummet $\mathbb{R} \times M$. Den anden er at betragte dem som en punktproces på $\mathbb{R}$, hvor der til hvert punkt t_i er knyttet en stokastisk variabel k_i . Begge metoder giver punktmønstre på formen $\{(t_i, k_i)\}$. Følgende defineres simple mærkede punktprocesser, samt stationaritet for mærkede punktprocesser.

Definition 6.1.

For en mærket punktproces $\mathbf{X}$ gælder følgende:

1. $\mathbf{X}$ er simpel hvis grundintensiteten er simpel.
2. $\mathbf{X}$ siges at være stationær, hvis sandsynlighedsstrukturen for punktprocessen er invariant under skift i $\mathbb{R}$.

[3, s. 195]

Der findes flere forskellige typer af mærker, der kan klassificeres alt efter hvad de afhænger af. Følgende præsenteres to forskellige typer.

Definition 6.2.

Der findes en mærket punktproces $\mathbf{X}$.

1. Hvis fordelingen af mærkerne er uafhængige af de tidligere punkter, og grundintensiteten er uafhængig af mærkerne, så siges mærkerne at være uafhængige.
2. Hvis fordelingen af mærkerne er uafhængige af de tidligere punkter, så siges mærkerne at være uforudsigelige.

[3, s. 195]

Dermed gælder der at de uforudsigelige mærker er i stand til at påvirke grundintensiteten. Følgende defineres den mærkede Poisson punktproces.

Definition 6.3. – Mærket Poisson punktproces

Hvis der findes en punktproces $\mathbf{X}$, hvor punkterne er fordelt efter en Poisson punktproces på tidslinien $\mathbf{Y}$, med en lokalt integrabel intensitetsfunktion μ , som har indbyrdes uafhængige mærker $\{k_i | t_i \in \mathbf{Y}\}$, så siges $\mathbf{X}$ at være en mærket Poisson punktproces.

[8, s. 26-27]

Hvis mærkerne alle er fordelt efter den samme fordeling kaldes denne fordeling for mærkefordelingen.

En måde at definere en punktproces på er ved dennes betingede intensitetsfunktion. Ligeledes kan de mærkede punktprocesser defineres på samme måde.

Definition 6.4.

Den betingede intensitetsfunktion for mærkede punktprocesser er givet ved

$$\lambda^*(t, k) = \lambda^*(t) f(k | \mathcal{H}_t),$$

hvor $\lambda^(t)$ er den betingede intensitetsfunktion for grundintensiteten, og hvor $k \in M$.*

[3, s. 249]

Når mærkede punktprocesser på tidslinien betragtes, er det muligt at bestemme likelihood-funktionen, ud fra den betingede intensitetsfunktion.

Sætning 6.5. – Likelihoodfunktion

Hvis der gives en mærket punktproces $\mathbf{X}$, samt et mærket punktmønster bestående af $((t_1, k_1), \dots, (t_n, k_n))$ på intervallet $[0, T] \times \mathbf{M}$ er dennes likelihoodfunktion givet ved:

$$L(\theta) = \left(\prod_{i=1}^n \lambda^*(t_i, k_i) \right) \exp \left(- \int_0^T \lambda^*(s) ds \right).$$

Hvor $\lambda^*(t, k)$ er den betingede intensitetsfunktion for den mærkede punktproces, og $\mathbf{M}$ er mærkerummet.

[3, s. 251]

Bevis. Sætning 6.5 bevises ligesom sætning 3.11

□

Følgende præsenteres en mærket udgave af Hawkes punktprocessen med uforudsigelige mærker. Det bemærkes at der i resten af rapporten udelukkende ses på mærkede punktprocesser hvor mærkerne er på denne form.

[3, s. 194-196, 249-251]

6.1 Den mærkede Hawkes punktproces

Det ønskes følgende at betragte mærkede Hawkes punktprocesser med uforudsigelige mærker. Ligesom tidligere, er det både muligt at definere den mærkede Hawkes punktproces som en klyngeproces, og ud fra dens betingede intensitetsfunktion. Det bemærkes at når mærkerne er uforudsigelige er $f(k|\mathcal{H}_t) = f(k)$ idet fordelingen af mærkerne er uafhængigt af de tidligere mærker samt tidspunktet for hændelsen.

Definition 6.6. – Mærket Hawkes punktproces

Hvis der gives en punktproces $\mathbf{X}$, med en betinget intensitetsfunktion på følgende form

$$\begin{aligned} \lambda^*(t, k) &= \lambda^*(t) f(k) \\ &= \left(\mu + \sum_{t_i < t} \alpha \gamma(t - t_i, k_i) \right) f(k), \end{aligned}$$

hvor $\lambda^*(t)$ er grundintensiteten, så siges $\mathbf{X}$ at være en mærket Hawkes punktproces.

[6, s. 630]

Definition 6.7.

Hvis der gælder at:

1. Immigranterne $\{y\}$ "ankommer" efter en Poisson punktproces med intensitet μ på $\mathbb{R}$.
2. Immigranternes mærker er i.i.d. fra mærkefordelingen.
3. Hver immigrant $y_i \in \{y\}$ danner en klynge C_i , der består af mærkede hændelser der er uafhængig af de andre klynger. Hver klynge C_i består af én oprindelig immigrant (y_i, k_i) samt afkom af forskellige generationer. Generation $k+1$ består af afkom fra generation k , ved at hvert punkt (t_i, k_i) i den k 'te generation, giver afkom, placeret på tidslinien efter en inhomogen Poisson punktproces på $\mathbb{R}$ med intensitetsfunktion $\alpha\gamma(t - t_i, k_i)$, hvor der til hvert afkom er tilknyttet et mærke trukket fra mærkefordelingen.
4. $\mathbf{X}$ er en superposition af alle klyngerne, C_i , for $i = 1, 2, \dots$

Så er $\mathbf{X}$ en Hawkes punktproces på $\mathbb{R} \times M$.

[6, s. 630]

Idet mærkerne er konstrueret på samme måde, som i.i.d. udtræk fra mærkefordelingen, er det muligt at vise at definition 6.6 og 6.7 er ens på samme måde som for sætning 3.20.

For en mærket Hawkes punktproces på tidslinien er det muligt at finde et udtryk for likelihoodfunktionen og log-likelihood funktionen.

Sætning 6.8.

For en mærket Hawkes punktproces gælder der, at dennes likelihoodfunktion er givet ved:

$$L(\theta) = \left(\prod_{i=1}^n \left(\mu + \sum_{t_j < t_i} \alpha\gamma(t_i - t_j, k_j) \right) f(k_i) \right) \cdot \exp \left(- \int_0^T \left(\mu + \sum_{t_j < s} \alpha\gamma(s - t_j, k_j) \right) ds \right).$$

Ligeledes er log-likelihoodfunktionen givet ved.

$$\mathcal{L}(\theta) = \left(\sum_{i=1}^n \log \left(\left(\mu + \sum_{t_j < t_i} \alpha\gamma(t_i - t_j, k_j) \right) f(k_i) \right) \right) - \int_0^T \left(\mu + \sum_{t_j < s} \alpha\gamma(s - t_j, k_j) \right) ds.$$

Bevis. For at finde likelihoodfunktionen indsættes definition 6.6, samt grundintensiteten i

lemma 6.5.

$$\begin{aligned} L(\theta) &= \left(\prod_{i=1}^n \lambda^*(t_i, k_i) \right) \exp \left(- \int_0^T \lambda^*(s) ds \right) \\ &= \left(\prod_{i=1}^n \left(\mu + \sum_{t_j < t_i} \alpha \gamma(t - t_j, k_j) \right) f(k_i) \right) \exp \left(- \int_0^T \left(\mu + \sum_{t_j < s} \alpha \gamma(s - t_j, k_j) \right) ds \right) \end{aligned}$$

Efterfølgende kan log-likelihoodfunktionen findes ved at tage logaritmen til likelihoodfunktionen.

$$\begin{aligned} \mathcal{L}(\theta) &= \log(L(\theta)) \\ &= \log \left(\left(\prod_{i=1}^n \left(\mu + \sum_{t_j < t_i} \alpha \gamma(t - t_j, k_j) \right) f(k_i) \right) \right. \\ &\quad \cdot \exp \left(- \int_0^T \left(\mu + \sum_{t_j < s} \alpha \gamma(s - t_j, k_j) \right) ds \right) \Bigg) \\ &= \left(\sum_{i=1}^n \log \left(\left(\mu + \sum_{t_j < t_i} \alpha \gamma(t - t_j, k_j) \right) f(k_i) \right) \right) - \int_0^T \left(\mu + \sum_{t_j < s} \alpha \gamma(s - t_j, k_j) \right) ds \end{aligned}$$

□

[3, s. 246-251][7, s. 53-55]

6.2 Simulering af den mærkede Hawkes punktproces

Hvis det ønskes at simulere punktmønstre, for en mærket Hawkes punktproces med uforudsigelige mærker, er det muligt, ligesom for det umærkede tilfælde, at opbygge algoritmen efter definitionen 6.7. Algoritmen består af fem forskellige trin, hvor det første trin simulerer immigranternes placering, som en homogen Poisson punktproces på tidslinien med intensitet μ . Det andet trin simulerer et mærke til hver af immigranterne fra mærkefordelingen. Det tredje trin er en mærket punktproces sammensat af placeringerne på tidslinien fra første trin og mærkerne fra andet trin. Det fjerde trin simulerer de forskellige generationer af afkom, hvor hver hændelse (t_i, k_i) generer afkom hvis placering på tidslinien er givet ved en inhomogen Poisson punktproces med intensitetsfunktion $\alpha \gamma(t - t_i, k_i)$. Mærkerne simuleres herefter fra mærkefordelingen. Det Endelige femte trin udgør superpositionen af samtlige punktermønstre. Dette kan ses i algoritme 4.

Algoritme 4 Mærket Hawkes punktproces med uforudsigelige mærker

En mærket Hawkes punktproces $\mathbf{X}$ med uforudsigelige mærker på intervallet $[0, T]$, kan genereres på følgende måde. Lad $k = 0$.

1. Generer $\mathbf{Y}$ som en Poisson punktproces med intensitetsfunktion μ på $[0, T]$.
2. For $t_i \in \mathbf{Y}$ generer k_i fra mærkefordelingen.
3. Lad $\mathbf{X}_k$ være den mærkede punktproces der består af $\{t_i, k_i\}$ for $i = 1, 2, \dots$

while Antallet af punkter i $\mathbf{X}_k > 0$ **do**

4. Generer $\mathbf{X}_k$ som en superposition af $\mathbf{Y}_i \sim \text{Poisson}([t_i, T], \alpha\gamma(t - t_i, k_i))$, hvor $(t_i, k_i) \in \mathbf{X}_{k-1}$ og mærkerne simuleres fra mærkefordelingen. Lad $k = k + 1$.

end while

5. Lad $\mathbf{X} = \bigcup_{i=0}^k \mathbf{X}_i$.

Det ses relativt simpelt, at algoritme 4 danne punktmønstre for mærkede Hawkes punktprocesser med uforudsigelige mærker, idet punkt 1-3 i algoritmen stemmer overens med punkt 1-2 i definition 6.7 og ligeledes stemmer 4-5 overens med 3-4 i definitionen.

En anden mulighed er at udnytte den evolutionære egenskab. Dette kan gøres, ved først at simulere immigranterne som en homogen Poisson punktproces med intensitetsfunktion μ , hvor der til hver immigrant simuleres et mærke fra mærkefordelingen. Herefter der startes ved den nedre grænse i tidsintervallet og hver gang man kommer til en hændelse (t_i, k_i) , simuleres afkom af denne hvor grundintensiteten er en inhomogen Poisson punktproces med intensitetsfunktion $\alpha\gamma(t - t_i, k_i)$, og mærkerne trækkes fra mærkefordelingen. Følgende præsenteres en algoritme bygget op på denne måde.

Algoritme 5 Mærket Hawkes punktproces med uforudsigelige mærker

En mærket Hawkes punktproces $\mathbf{X}$ med uforudsigelige mærker på intervallet $[0, T]$ kan genereres på følgende måde. Lad $i = 1$

1. Generer $\mathbf{Y}$ som en Poisson punktproces med intensitetsfunktion μ på intervallet $[0, T]$.
2. For $t_j \in \mathbf{Y}$ generer k_j fra mærkefordelingen.
3. Lad $\mathbf{X}$ være den mærkede punktproces bestående af $\{t_j, k_j\}$ for $j = 1, 2, \dots$ Hvor $\mathbf{x}$ er sorteret i forhold til punkternes placering på tidslinien.

while $t_i < T$ **do**

4. Generer $x \sim \text{Poisson}((t_i, T], \alpha\gamma(t - t_i, k_i))$.
5. Generer et y fra mærkefordelingen for hvert x .
6. Lad $\mathbf{X} = \mathbf{X} \bigcup_{j=1}^{\text{length}(x)} (x_j, y_j)$.
7. Sorter $\mathbf{X}$ i forhold til placeringen på tidslinien.
8. $i=i+1$

end while

Konstruktionen af immigranterne, punkt 1-3 i algoritmen, stemmer overens med punkt 1-2

i definition 6.7. Samtlige hændelser (t_i, k_i) afføder nye hændelser med en grundintensitet der følger en inhomogen Poisson punktproces med intensitetsfunktion $\alpha\gamma(t - t_i, k_i)$, og hvor mærkerne trækkes fra mærkefordelingen. Dermed gælder der, at da hver klynge består af forskellige afkom fra én immigrant og klyngerne er konstrueret i overensstemmelse med punkt 3. i definition 6.7, så genererer algoritme 5 punktmønstre fra den mærkede Hawkes punktproces med uforudsigelige mærker.

KAPITEL 7

DEN MÆRKEDE MODEL

I forbindelse med dette afsnit er der lavet en del udregninger/billeder i R. Scripts for dette kan ses i appendiks 11.6.3.

Det ønskes i det følgende afsnit, at konstruere en model for jordskælvshændelser i Danmark. Denne model skal være i stand til både at modellere tidspunkt, samt styrken af jordskælvs-hændelser. Ydermere skal den være i stand til, at afspejle flere forskellige tendenser for jordskælv. En af tendenserne er, at en jordskælvhændelse har en tendens til at have efterskælv, samt at et kraftigt jordskælv ofte vil have flere efterskælv end et mindre kraftigt jordskælv, jvf. afsnit 2.1. Ydermere vælges det at lade styrken af jordskælvshændelserne være uafhængigt af tidspunktet for hændelsen, da der ikke er nogen grund til at antage at tidspunktet vil påvirke styrken. Yderligere vælges det, at styrken af et jordskælv skal være uafhængigt af styrken af tidligere jordskælv, også selvom jordskælvet måske er et efterskælv af et andet jordskælv. Denne sidste antagelse bygger ikke på noget geografisk argument, men er medtaget af rent praktiske grunde for at mærkerne er uforudsigelige. En sidste ting der også bør medtages er, at sandsynligheden for et efterskælv af en jordskælvshændelse bør være størst lige efter hændelsen er indtruffet, og være monotont aftagende som funktion af tiden. Bemærk at modellen ikke konstrueres således at den er i stand til at medtage forskælv.

En model der opfylder de ovennævnte kriterier, er en mærket Hawkes punktproces med uforudsigelige mærker. Ifølge definition 6.6 har en mærket Hawkes punktproces med uforudsigelige mærker følgende betingede intensitetsfunktion.

$$\lambda^*(t, k) = \left(\mu + \sum_{t_i < t} \alpha \gamma(t - t_i, k_i) \right) f(k), \quad (7.1)$$

hvor $f(k)$ er tæthedsfunktionen for mærkefordelingen, μ og α er en ukendt konstant og $\gamma(\cdot)$ er en ukendt funktion.

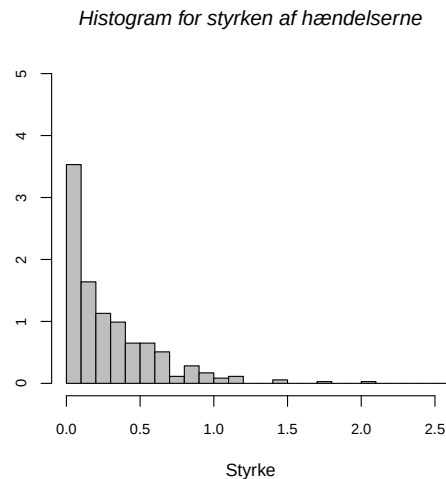
Det er nødvendigt at udvælge funktioner for γ og $f(k)$, før det er muligt at begynde at bestemme de forskellige parametre. Da Richter-skalaen er en logaritmisk skala, vælges det at antage, at antallet af efterskælv vokser eksponentielt, som en funktion af styrken af en

hændelse. Det vælges derfor at lade $\gamma(t, k) = \exp(k) \cdot \exp(-\beta \cdot (t))$, hvor β er en ukendt konstant. På denne måde bibeholdes den egenskab i modellen fra afsnit 5, der medfører at sandsynligheden, for et efterskælv af en hændelse, er en monotont aftagende funktion af tiden efter hændelsen, som ønsket. Følgende betragtes $f(k)$.

7.1 Fordelingen af mærkerne

Der er ingen grund til at tro, at fordelingen af styrken for jordskælvshændelser vil ændres i fremtiden, jvf. afsnit 2.1. Det ønskes derfor at finde en fordelingsfunktion for det givne datasæt. I de udvalgte data, der arbejdes med i denne rapport, er der medtaget alle hændelser med en styrke på to, eller mere på Richter-skalaen i tidsperioden fra 2006 til og med 2009.

For at gøre det simplere, at finde fordelingen af hændelser, vælges det at trække to fra alle de givne styrker, i det udvalgte datasæt. Det er derfor nødvendigt, at huske at addere med disse to igen under en eventuel simulation. Det ønskes umiddelbart, at finde en fordeling der opfylder, at der er mange relativt svage jordskælv, og få relativt stærke. På figur 7.1 ses et histogram for de redigerede styrke af de udvalgte jordskælvshændelserne.



Figur 7.1: Histogram for styrken af jordskælvshændelserne i det udvalgte datasæt. Bemærk at der er trukket to fra styrken. Histogrammet er konstrueret med R-script 11.19.

På baggrund af figur 7.1 vælges det at antage, at fordelingen af de redigerede styrker følger en eksponentiel fordeling. Dermed bliver tæthedsfunktionen for mærkefordelingen af redigerede styrker:

$$f(\tilde{k}) = \eta \cdot \exp(-\eta \cdot \tilde{k}), \quad \tilde{k} \geq 0.$$

Tæthedsfunktionen for de *rigtige* styrker bliver dermed:

$$f(k) = \eta \cdot \exp(-\eta \cdot (k - 2)), \quad k \geq 2. \quad (7.2)$$

Hvis det ønskes at generere uforudsigelige mærker $\{k_i\}$, til en model der simulere jordskælv med en styrke på 2 eller mere på richterskalaen, kan disse dermed simuleres som $2 + \tilde{k}_i$, hvor $\tilde{k}_i \sim \text{Exp}(\eta)$.

7.2 Likelihoodfunktionen for den mærkede model

Da det senere ønskes at udføre inferens, for at bestemme de ukendte parametre, findes log-likelihoodfunktionen for modellen. Først indsættes tætheden for mærkefordelingen, ligning (7.2) samt funktionerne for γ ind i den betingede intensitetsfunktion ligning (7.1).

$$\lambda^*(t, k) = \left(\mu + \left(\sum_{t_i < t} \exp(k_i) \cdot \alpha \cdot \exp(-\beta(t - t_i)) \right) \right) \eta \cdot \exp(-\eta \cdot (k - 2)). \quad (7.3)$$

Ydermere er den betingede intensitetsfunktion for grundintensiteten givet ved:

$$\lambda^*(t) = \mu + \left(\sum_{t_i < t} \exp(k_i) \cdot \alpha \cdot \exp(-\beta(t - t_i)) \right).$$

Da der nu haves en betinget intensitetsfunktion, der ikke indeholder ukendte funktioner, findes log-likelihoodfunktionen for modellen med sætning 6.8.

$$\begin{aligned} \mathcal{L}(\mu, \alpha, \beta) &= \left(\sum_{i=1}^n \log(\lambda^*(t_i, k_i)) \right) - \int_0^T \lambda^*(s) ds. \\ &= \left(\sum_{i=1}^n \log \left(\left(\mu + \left(\sum_{t_j < t_i} \exp(k_j) \cdot \alpha \cdot \exp(-\beta(t_i - t_j)) \right) \right) \cdot \eta \cdot \exp(-\eta \cdot (k_i - 2)) \right) \right) \\ &\quad - \int_0^T \left(\mu + \left(\sum_{t_j < s} \exp(k_j) \cdot \alpha \cdot \exp(-\beta(s - t_j)) \right) \right) ds. \end{aligned} \quad (7.4)$$

For overskuelighedens skyld betragtes de to led i ligning (7.4) hver for sig. Først betragtes den første del.

$$\begin{aligned} &\left(\sum_{i=1}^n \log \left(\left(\mu + \left(\sum_{t_j < t_i} \exp(k_j) \cdot \alpha \cdot \exp(-\beta(t_i - t_j)) \right) \right) \cdot \eta \cdot \exp(-\eta \cdot (k_i - 2)) \right) \right) \\ &= \left(\sum_{i=1}^n \log(\mu + \alpha A(i)) + \log(\eta) + (-\eta(k_i - 2)) \right) \\ &= \left(\sum_{i=1}^n \log(\eta) + (-\eta(k_i - 2)) \right) + \left(\sum_{i=1}^n \log(\mu + \alpha A(i)) \right), \end{aligned} \quad (7.5)$$

hvor

$$A(i) = \begin{cases} 0, & \text{for } i = 1 \\ \left(\sum_{t_j < t_i} \exp(k_j) \exp(-\beta(t_i - t_j)) \right) & \text{for } i \geq 2 \end{cases}.$$

Følgende betragtes det sidste led i ligning (7.4).

$$\begin{aligned}
 & \int_0^T \left(\mu + \left(\sum_{t_j < s} \exp(k_j) \cdot \alpha \cdot \exp(-\beta(s - t_j)) \right) \right) ds \\
 &= \int_0^T \mu ds + \int_0^T \left(\sum_{t_j < s} \exp(k_j) \cdot \alpha \cdot \exp(-\beta(s - t_j)) \right) ds \\
 &= \mu T - \frac{\alpha}{\beta} \int_0^T \left(\sum_{t_j < s} \exp(k_j) \cdot (-\beta) \exp(-\beta(s - t_j)) \right) ds \\
 &= \mu T - \frac{\alpha}{\beta} \left(\sum_{i=1}^n \exp(k_i) (\exp(-\beta(T - t_i)) - \exp(0)) \right) \\
 &= \mu T + \left(\sum_{i=1}^n \alpha \exp(k_i) \frac{1 - \exp(-\beta(T - t_i))}{\beta} \right). \tag{7.6}
 \end{aligned}$$

Nu indsættes ligning (7.5) og (7.6) i (7.4).

$$\begin{aligned}
 \mathcal{L}(\mu, \alpha, \beta, \eta) &= \left(\sum_{i=1}^n \log(\eta) + (-\eta(k_i - 2)) \right) + \left(\sum_{i=1}^n \log(\mu + \alpha A(i)) \right) \\
 &\quad - \mu T - \left(\sum_{i=1}^n \alpha \exp(k_i) \frac{1 - \exp(-\beta(T - t_i))}{\beta} \right). \tag{7.7}
 \end{aligned}$$

7.3 Parameter estimation for den mærkede model

Det ønskes at bestemme de ukendte parametre μ, α, β og η , således at modellen passer bedst til det givne datasæt. Det ønskes altså at finde de værdier af μ, α, β og η , der maksimerer likelihoodfunktionen, og i dette tilfælde log-likelihoodfunktionen. Da $\mathcal{L}(\mu, \alpha, \beta, \eta)$ er konkav, vil det entydige maksimum være når hældningen er lig 0. Det er derfor nødvendigt, at løse følgende ligningssystem.

$$\begin{aligned}
 \frac{\delta}{\delta \mu} \tilde{\mathcal{L}}(\theta) &= 0, \\
 \frac{\delta}{\delta \alpha} \tilde{\mathcal{L}}(\theta) &= 0, \\
 \frac{\delta}{\delta \beta} \tilde{\mathcal{L}}(\theta) &= 0, \\
 \frac{\delta}{\delta \eta} \tilde{\mathcal{L}}(\theta) &= 0,
 \end{aligned}$$

hvor $\theta = \{\mu, \alpha, \beta, \eta\}$. For at løse et sådant ligningssystem anvendes Newton-Rhapson metoden fra afsnit 3.3.3.

$$\theta^{(n)} = \theta^{(n-1)} - H^{-1}(\mathcal{L}(\theta^{(n-1)})) \Delta(\theta^{(n-1)}).$$

Det er derfor nødvendigt, at bestemme vektoren bestående af de partielt afledte $\Delta(\mathcal{L}(\theta))$ samt Hessian-matricen. Det er altså nødvendigt at bestemme:

$$\Delta(\mathcal{L}(\theta)) = \begin{bmatrix} \frac{\partial}{\partial \mu} \mathcal{L}(\theta) \\ \frac{\partial}{\partial \alpha} \mathcal{L}(\theta) \\ \frac{\partial}{\partial \beta} \mathcal{L}(\theta) \\ \frac{\partial}{\partial \eta} \mathcal{L}(\theta) \end{bmatrix}, \quad H(\mathcal{L}(\theta)) = \begin{bmatrix} \frac{\partial^2 \mathcal{L}(\theta)}{\partial \mu^2} & \frac{\partial^2 \mathcal{L}(\theta)}{\partial \mu \partial \alpha} & \frac{\partial^2 \mathcal{L}(\theta)}{\partial \mu \partial \beta} & \frac{\partial^2 \mathcal{L}(\theta)}{\partial \mu \partial \eta} \\ \frac{\partial^2 \mathcal{L}(\theta)}{\partial \alpha \partial \mu} & \frac{\partial^2 \mathcal{L}(\theta)}{\partial \alpha^2} & \frac{\partial^2 \mathcal{L}(\theta)}{\partial \alpha \partial \beta} & \frac{\partial^2 \mathcal{L}(\theta)}{\partial \alpha \partial \eta} \\ \frac{\partial^2 \mathcal{L}(\theta)}{\partial \beta \partial \mu} & \frac{\partial^2 \mathcal{L}(\theta)}{\partial \beta \partial \alpha} & \frac{\partial^2 \mathcal{L}(\theta)}{\partial \beta^2} & \frac{\partial^2 \mathcal{L}(\theta)}{\partial \beta \partial \eta} \\ \frac{\partial^2 \mathcal{L}(\theta)}{\partial \eta \partial \mu} & \frac{\partial^2 \mathcal{L}(\theta)}{\partial \eta \partial \alpha} & \frac{\partial^2 \mathcal{L}(\theta)}{\partial \eta \partial \beta} & \frac{\partial^2 \mathcal{L}(\theta)}{\partial \eta^2} \end{bmatrix}.$$

Alle udregninger af de forskellige partielt afledte kan findes i appendiks 11.5. Først findes de første ordens partielt afledte.

$$\begin{aligned} \frac{\partial}{\partial \mu} \mathcal{L}(\theta) &= \left(\sum_{i=1}^n \frac{1}{\mu + \alpha A(i)} \right) - T \\ \frac{\partial}{\partial \alpha} \mathcal{L}(\theta) &= \left(\sum_{i=1}^n \frac{A(i)}{\mu + \alpha A(i)} \right) - \left(\sum_{i=1}^n \exp(k_i) \frac{1 - \exp(-\beta(T - t_i))}{\beta} \right) \\ \frac{\partial}{\partial \beta} \mathcal{L}(\theta) &= \left(\sum_{i=1}^n \frac{-\alpha B(i)}{\mu + \alpha A(i)} \right) \\ &\quad - \alpha \left(\sum_{i=1}^n \exp(k_i) \left(\frac{\exp(-\beta(T - t_i)) - 1}{\beta^2} + \frac{(T - t_i) \exp(-\beta(T - t_i))}{\beta} \right) \right) \\ \frac{\partial}{\partial \eta} \mathcal{L}(\theta) &= \left(\sum_{i=1}^n \frac{1}{\eta} + (2 - k_i) \right), \end{aligned}$$

hvor

$$B(i) = \begin{cases} 0, & \text{for } i = 1 \\ \left(\sum_{t_j < t_i} \exp(k_j) (t_i - t_j) \exp(-\beta(t_i - t_j)) \right) & \text{for } i \geq 2 \end{cases}$$

Dermed er det muligt, at bestemme vektoren med de partielt afledte. For endvidere at bestemme Hessian-matricen er det nødvendigt, at kende de anden ordens partielt afledte. Derfor findes de forskellige partielt afledte af anden orden for log-likelihoodfunktionen til

Hawkes punktprocessen fra ligning (7.7). Først findes de dobbelt afledte af μ , α , β og η .

$$\begin{aligned}\frac{\partial}{\partial \mu^2} \mathcal{L}(\theta) &= \sum_{i=1}^n \frac{-1}{(\mu + \alpha A(i))^2} \\ \frac{\partial}{\partial \alpha^2} \mathcal{L}(\theta) &= \sum_{i=1}^n \frac{-A(i)^2}{(\mu + \alpha A(i))^2} \\ \frac{\partial}{\partial \beta^2} \mathcal{L}(\theta) &= \left(\sum_{i=1}^n \frac{\alpha C(i)}{\mu + \alpha A(i)} - \left(\frac{\alpha B(i)}{(\mu + \alpha A(i))} \right)^2 \right) \\ &\quad + \alpha \left(\sum_{i=1}^n \exp(k_i) \left(\frac{2(\exp(-\beta(T - t_i)) - 1)}{\beta^3} \right. \right. \\ &\quad \left. \left. + \frac{2(T - t_i) \exp(-\beta(T - t_i))}{\beta^2} + \frac{(T - t_i)^2 \exp(-\beta(T - t_i))}{\beta} \right) \right), \\ \frac{\partial}{\partial \eta^2} \mathcal{L}(\theta) &= \sum_{i=1}^n \frac{-1}{\eta^2},\end{aligned}$$

hvor

$$C(i) = \begin{cases} 0, & \text{for } i = 1 \\ \left(\sum_{t_j < t_i} \exp(k_j) (t_i - t_j)^2 \exp(-\beta(t_i - t_j)) \right) & \text{for } i \geq 2. \end{cases}$$

Da der gælder, at

$$\begin{aligned}\frac{\partial}{\partial \alpha \partial \mu} \mathcal{L}(\theta) &= \frac{\partial}{\partial \mu \partial \alpha} \mathcal{L}(\theta), & \frac{\partial}{\partial \beta \partial \alpha} \mathcal{L}(\theta) &= \frac{\partial}{\partial \alpha \partial \beta} \mathcal{L}(\theta), \\ \frac{\partial}{\partial \beta \partial \mu} \mathcal{L}(\theta) &= \frac{\partial}{\partial \mu \partial \beta} \mathcal{L}(\theta), & \frac{\partial}{\partial \eta \partial \alpha} \mathcal{L}(\theta) &= \frac{\partial}{\partial \alpha \partial \eta} \mathcal{L}(\theta), \\ \frac{\partial}{\partial \eta \partial \mu} \mathcal{L}(\theta) &= \frac{\partial}{\partial \mu \partial \eta} \mathcal{L}(\theta), & \frac{\partial}{\partial \eta \partial \beta} \mathcal{L}(\theta) &= \frac{\partial}{\partial \beta \partial \eta} \mathcal{L}(\theta),\end{aligned}$$

er det nok at finde seks af dem.

$$\frac{\partial}{\partial \mu \partial \alpha} \mathcal{L}(\theta) = \sum_{i=1}^n \frac{-A(i)}{(\mu + \alpha A(i))^2}.$$

$$\frac{\partial}{\partial \mu \partial \beta} \mathcal{L}(\theta) = \sum_{i=1}^n \frac{\alpha B(i)}{(\mu + \alpha A(i))^2}.$$

$$\frac{\partial}{\partial \mu \partial \eta} \mathcal{L}(\theta) = 0.$$

$$\begin{aligned} \frac{\partial}{\partial \alpha \partial \beta} \mathcal{L}(\theta) = & \left(\sum_{i=1}^n \frac{-B(i)}{\mu + \alpha A(i)} + \frac{\alpha A(i) B(i)}{(\mu + \alpha A(i))^2} \right) \\ & - \left(\sum_{i=1}^n \exp(k_i) \left(\frac{1}{\beta^2} (\exp(-\beta(T - t_i)) - 1) + \frac{1}{\beta} (T - t_i) \exp(-\beta(T - t_i)) \right) \right). \end{aligned}$$

$$\frac{\partial}{\partial \alpha \partial \eta} \mathcal{L}(\theta) = 0.$$

$$\frac{\partial}{\partial \beta \partial \eta} \mathcal{L}(\theta) = 0.$$

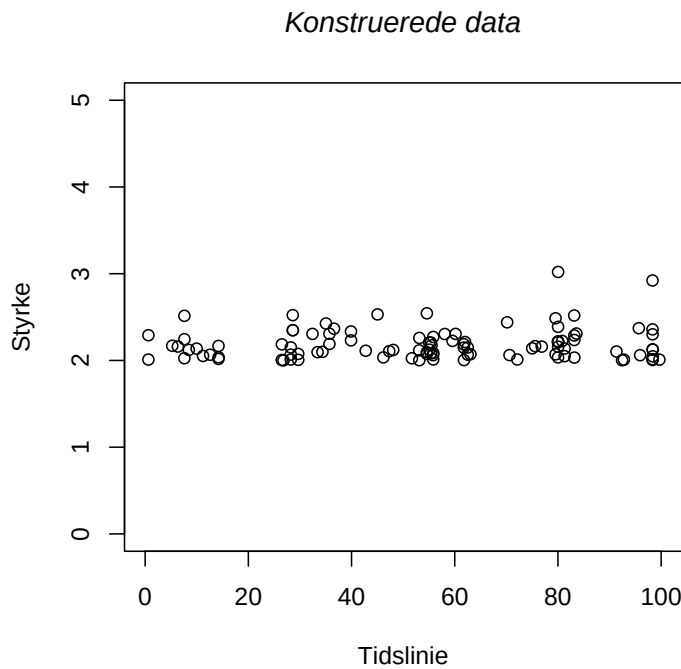
Dermed er Hessian-matricen bestemt. Bemærk, at for at det er muligt, at finde den inverse til Hessian-matricen er det nødvendigt, at matricen er singulær.

Newton-Rhapon algoritmen kan anvendes, jvf. afsnit 3.3.3. Inden algoritmen anvendes til at approksimere de parametrene, for de udvalgte jordskælvshændelserne i Danmark, vælges det at afprøve den på et datasæt konstrueret med algoritme 5, således er det muligt at se om algoritmen giver noget, der tilnærmelsesvis er rigtigt. Det vælges at simulere på intervallet $[0, 100]$. I tabel 7.1, er det muligt at se de valgte værdier for det konstruerede datasæt, samt de approksimerede værdier, med den fire-dimensionelle Newton-Rhapon algoritme, hvor $H(\mathcal{L}(\theta))$ og $\Delta(\theta)$ angivet som ovenfor. Det er nødvendigt med startværdier til algoritmen, i dette tilfælde er det valgt, at tage udgangspunkt i det halve af de *korrekte* værdier, der er brugt til konstruktionen af datasættet.

Tabel 7.1: Fire-dimensionel Newton-Rhapon anvendes på et konstrueret datasæt. Data-sættet samt de forskellige approksimationer er bestemt med R-script 11.20.

	μ	α	β	η
Korrekte parametre	0.5	4	100	4
Gæt	0.25	2	50	2
Approksimation (1 skridt)	0,40	3,15	76,57	3,25
Approksimation (5 skridt)	0,61	4,92	111,47	5,35
Approksimation (10 skridt)	0,61	4,92	111,47	5,35
Approksimation (20 skridt)	0,61	4,92	111,47	5,35

På figur 7.2 er det muligt, at se det konstruerede datasæt. Det bemærkes, at de estimerede parametre afviger lidt fra de *korrekte*. Dette er ikke så overraskende, idet det simulerede datasæt for en mærket Hawkes punktproces, let kan være atypisk for de givne parametre, idet der er en stor variation af antallet af punkter i klyngerne, jvf. [12, s. 11]. Det bemærkes at der er 102 hændelser i det konstruerede datasæt.



Figur 7.2: Det konstruerede datasæt. Plottet med R-script 11.20

Det ønskes følgende, at anvende Newton-Rhapon algoritmen til at approksimere parametrene for de udvalgte jordskælvshændelser. Derfor er det nødvendigt, at bestemme nogle startværdier. Intensiteten af immigranterne vælges, som i det umærkede tilfælde, til $\mu = \frac{354}{T} = \frac{354}{4}$. Den reelle værdi af μ burde være mindre end dette gæt, idet det er en rimelig antagelse, at nogle af punkterne vil være afkom af andre punkter. Startværdien af α, β og η bliver derimod valgt noget mere tilfældigt. Det ønskes selvfølgelig, at vælge en positiv startværdi. De vælges således $\alpha = 25$, $\beta = 2000$ og $\eta = 2$. Umiddelbart burde gættet være mere eller mindre irrelevant, hvis bare der foretages nok iterationer med algoritmen, men dette er dog ikke helt tilfældet. Der kan opstå problemer, idet likelihoodfunktionen kan være meget flad, og det er derfor ønskværdigt at gættet ligger relativt tæt på den *rigtige* værdi. På tabel 7.2 er det muligt at se gættet på startværdierne, samt værdierne approksimeret med algoritmen ved 1, 5, 10 og 20 iterationer.

Tabel 7.2: Den tre-dimensionel Newton-Rhapon algoritme anvendt på de udvalgte jordskælvshændelser. Udregnet med R-script 11.21.

	μ	α	β	η
Gæt	$\frac{354}{4}$	25	2000	2
Approksimation (1 skridt)	73,88	31,87	2502,26	2,67
Approksimation (5 skridt)	76,40	33,20	2718,41	3,01
Approksimation (10 skridt)	76,40	33,20	2718,41	3,01
Approksimation (20 skridt)	76,40	33,20	2718,41	3,01

Det antages at $\{\mu^{(20)}, \alpha^{(20)}, \beta^{(20)}, \eta^{(20)}\} = \{\hat{\mu}, \hat{\alpha}, \hat{\beta}, \hat{\eta}\} = \hat{\theta}$. Disse værdier kan indsættes i

den betingede intensitetsfunktion fra ligning (7.3).

$$\lambda^*(t, k) = \left(\hat{\mu} + \sum_{t_i < t} \left(\exp(k_i) \cdot \hat{\alpha} \cdot \exp(-\hat{\beta}(t - t_i)) \right) \right) \hat{\eta} \cdot \exp(-\hat{\eta} \cdot (k - 2)).$$

Den første kontrol af de estimerede parametre er at indsætte dem i de partielt afledte til log-likelihoodfunktionen, for at se om de giver 0 eller noget tæt på dette. De afledte udregnes med R-script 11.22 til følgende:

$$\Delta(\mathcal{L}(\hat{\theta})) = \begin{bmatrix} \frac{\partial}{\partial \mu} \mathcal{L}(\hat{\theta}) \\ \frac{\partial}{\partial \alpha} \mathcal{L}(\hat{\theta}) \\ \frac{\partial}{\partial \beta} \mathcal{L}(\hat{\theta}) \\ \frac{\partial}{\partial \eta} \mathcal{L}(\hat{\theta}) \end{bmatrix} \approx \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Dermed gælder der, at de approksimerede parametre er meget tæt på at maksimere likelihoodfunktionen.

Hvis det nu ønskes, at finde det gennemsnitlige antal afkom, er det muligt på baggrund af de estimerede parametre. Betragt et hypotetisk punkt der starter i tidspunktet 0 og med et mærke der svarer til middelværdien af mærkerne. Så vil det gennemsnitlige antal afkom for dette punkt være givet ved integralet fra 0 til ∞ . Først findes middelværdien af mærkerne.

$$\begin{aligned} E[2 + \exp(\hat{\eta})] &= 2 + E[\exp(\hat{\eta})] \\ &= 2 + \frac{1}{\hat{\eta}} \\ &\approx 2 + \frac{1}{3,01} \\ &\approx 2,33 \end{aligned}$$

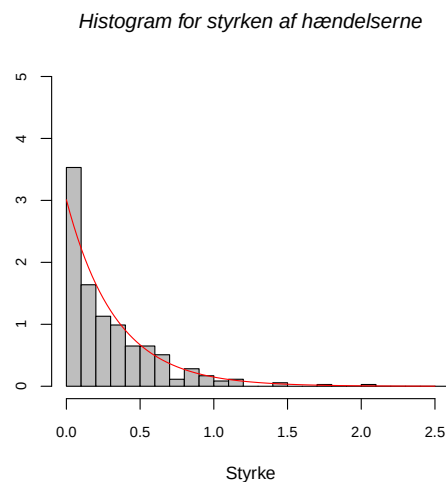
Følgende er det muligt at finde det gennemsnitlige antal afkom.

$$\begin{aligned} &\int_0^\infty \hat{\alpha} \exp(E[2 + \exp(\hat{\eta})]) \exp(-\hat{\beta}(s - 0)) ds \\ &= -\frac{\hat{\alpha}}{\hat{\beta}} E[2 + \exp(\hat{\eta})] \int_0^\infty -\hat{\beta} \exp(-\hat{\beta}(s)) ds \\ &= -\frac{\hat{\alpha}}{\hat{\beta}} \exp(E[2 + \exp(\hat{\eta})]) (\lim_{s \rightarrow \infty} (\exp(-\hat{\beta}(s))) - \exp(-\hat{\beta} \cdot 0)) \\ &= -\frac{\hat{\alpha}}{\hat{\beta}} \exp(E[2 + \exp(\hat{\eta})]) (0 - 1) \\ &= \frac{\hat{\alpha}}{\hat{\beta}} \exp(E[2 + \exp(\hat{\eta})]) \\ &\approx \frac{33,20}{2718,41} \exp(2,33) \\ &\approx 0,13. \end{aligned}$$

7.4 Modelkontrol af den mærkede model

Der er nu konstrueret en model, hvor parametrene er fundet med inferens via et udvalg af et givent datasæt. Dette medfører dog ikke nødvendigvis, at modellen er i stand til at konstruere punktmønstre, der ligner de udvalgte data fra det givne datasæt. Det ønskes derfor at udføre modelkontrol. Først betragtes mærkefordelingen.

I den valgte model antages det at mærkerne er fordelt efter $2 + \tilde{k}_i$, hvor $k_i \sim \text{Exp}(\hat{\eta})$ og $\hat{\eta} = 3,01$. Hvis dette stemmer overens med virkeligheden, vil der gælde at histogrammet, for styrken af de redigerede mærker¹, bør følge fordelingsfunktionen for en eksponentiel fordeling med parameteren $\hat{\eta}$.



Figur 7.3: Histogram for den redigerede styrke af jordskælvshændelserne i det udvalgte datasæt. Den røde kurve er fordelingsfunktionen for en eksponentiel fordeling med parameteren $\lambda = 3,01$. Bemærk at der er trukket 2 fra styrken. Konstrueret ud fra R-script 11.23.

På figur 7.3 ses histogrammet for de redigerede data, samt fordelingsfunktionen for den eksponentielle fordeling med parameteren $\hat{\eta}$. Yderligere kan der på figur 7.4, ses et q-q plot af de empiriske data for de redigerede styrker imod de teoretiske værdier.

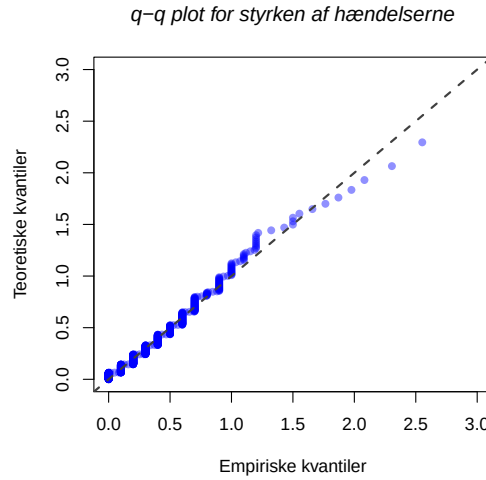
Det første der bemærkes er den markante *trappeform* af q-q plottet, hvilket antyder en eller anden form for diskretisering i datasættet. Dette skyldes, at de data der er til rådighed, for styrken af jordskælvhændelserne, ikke er kontinuerte, men derimod er afrundet til første decimal.

Det ses dermed, på baggrund af figur 7.3 og 7.4, at fordelingen af mærkerne i modellen nogenlunde følger fordelingen af mærkerne i virkeligheden.

Da mærkefordelingen er undersøgt, ønskes det følgende at undersøge grundintensiteten.

$$\lambda^*(t) = \left(\mu + \sum_{t_i < t} \exp(k_i) \cdot \alpha \cdot \exp(-\beta(t - t_i)) \right).$$

¹Der hvor der er trukket to fra styrken.



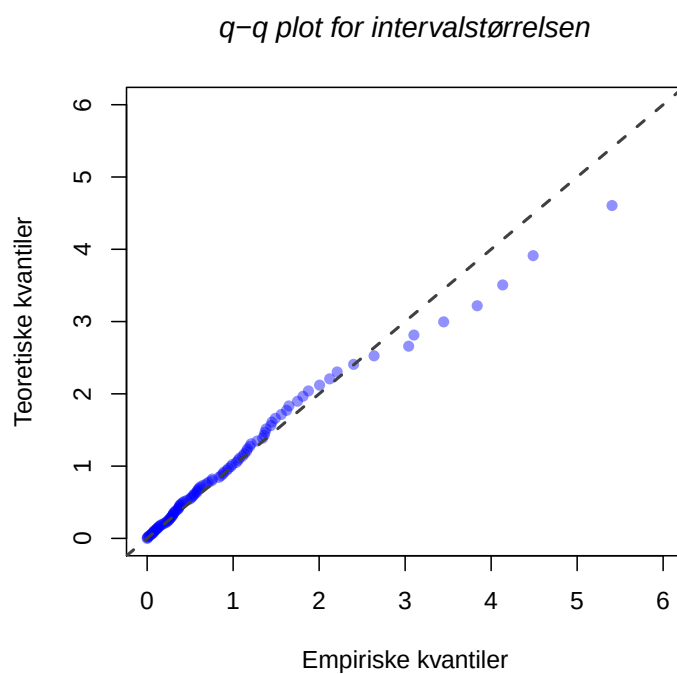
Figur 7.4: q-q plot for de empiriske værdier af de redigerede styrker imod de teoretiske værdier. Konstrueret ud fra R-script 11.23.

For at undersøge om funktionen for grundintensiteten passer til det givne datasæt, udregnes $\Lambda^*(t)$ ved hjælp af definition 3.9.

$$\begin{aligned}
 \Lambda^*(t) &= \int_0^t \left(\mu + \left(\sum_{t_i < s} \exp(k_i) \alpha \exp(-\beta(s - t_i)) \right) \right) ds \\
 &= \mu t - \frac{\alpha}{\beta} \int_0^t \left(\sum_{t_i < s} -\exp(k_i) \beta \exp(-\beta(s - t_i)) \right) ds \\
 &= \mu t - \frac{\alpha}{\beta} \left(\sum_{t_i < t} \exp(k_i) (\exp(-\beta(t - t_i)) - 1) \right) \\
 &= \mu t + \frac{\alpha}{\beta} \left(\sum_{t_i < t} \exp(k_i) (1 - \exp(-\beta(t - t_i))) \right)
 \end{aligned}$$

Følgende gælder der, jvf. sætning 3.24, at $\{s_i\} = \Lambda^*(t_i)$ for $i = 1, 2, \dots, 354$ vil være fordelt, som en Poisson punktproces med intensitet 1. Dermed bør intervallerne mellem hændelserne i $\{s_i\}$, hver især være trukket fra en eksponentiel fordeling med parameteren 1. På figur 7.5 er det muligt at se et q-q plot af den empiriske fordeling af intervallængderne i $\{s_i\}$, imod de teoretiske kvantiler for den eksponentielle fordeling med intensitet 1.

På baggrund af figur 7.5 vurderes det, at modellen er nogenlunde velvalgt.



Figur 7.5: q-q plot for det transformerede datasæt. Konstrueret ud fra R-script 11.24.

Der er nu fundet to forskellige modeller for jordskælvshændelser. Der gælder dog at når det ønskes at simulere punktmønstre på tidslinien, er det relevant at være opmærksom på eventuelle kantproblematikker. Problemerne skyldes, at hvis det ønskes at simulere punktmønstre på intervallet $[0, T]$, så er der ikke nogen viden om hændelserne før tiden 0. Sådanne hændelser, burde ellers have mulighed for at påvirke af antallet af punkter i intervallet. Der gælder, for Hawkes punktprocessen, at eventuelt afkom før tiden 0 er i stand til at have afkom i tidsintervallet $[0, T]$. De algoritmer, til simulering af mærkede Hawkes punktprocessen, der blev præsenteret i afsnit 3.5 og 6.2, medtager ikke afkom fra hændelser før tiden 0. Der simuleres altså ikke punktmønstre fra en stationær mærket Hawkes punktproces, men en approksimation til denne. En mulighed vil være at simulere i et større tidsinterval, og så kun anvende et delinterval af dette. Det der fravælges, i starten af de simulerede punktmønstre, vil være den såkaldte *burn-in* periode. En anden mulighed er at konstruere en simuleringsalgoritme, der direkte er i stand til at simulere stationære punktmønstre for mærkede Hawkes punktprocesser, uden at være nødsaget til at simulere på et større tidsinterval. Dette kaldes for *perfekt* simulering. I dette afsnit vil den sidste mulighed blive betragtet.

Da simuleringsalgoritmen er afhængig af hvilken form for mærker der er tilknyttet hændelserne, vælges det blot at betragte de uforudsigelige mærker.

For Hawkes punktprocessen på tidslinien opstår det eneste kantproblem ved tiden 0. For at konstruere en algoritme, der er i stand til at lave *perfekt* simulering, er det nødvendigt på en eller anden måde at være i stand til, at bestemme afkommet fra tidligere punkter, der falder i det givne tidsinterval. Hvis der haves et punkt t før tiden 0, må sandsynligheden for at dette punkt har mindst et afkom efter tiden 0 være givet ved $(1 - F(-t))$, hvor F er fordelingsfunktionen for længden af klyngerne. Det ønskes derfor, at bestemme fordelingen af længden af klyngerne.

[6, s. 629-630][8, s. 216-217]

8.1 Fordelingen af længden af en klynge

Længden af en klynge er længde af intervallet imellem det først og det sidste punkt i klyngen. Det er altså det størst mulige interval mellem to punkter i klyngen. Lad længden af en klynge være givet ved L . For en stationær mærket Hawkes punktproces med uforudsigelige mærker, er det muligt, at betragte længderne for de forskellige klynger, som udtræk fra én og samme tæthedsfunktion. Antag at 1_n er fordelingsfunktionen for klyngelængderne, når det kun er de første n generationer af den mærkede Hawkes punktproces, der betragtes. Der gælder da, at $1_n \rightarrow F$ når $n \rightarrow \infty$, hvor F er den *rigtige* fordelingsfunktion for klyngelængderne.

Sætning 8.1.

Hvis 1_n er defineret som ovenfor, og

$$\phi(f)(t) = E \left[\exp \left(-v + \int_0^T f(t-s) \alpha \gamma(s, k) ds \right) \right],$$

hvor $\phi(f)$ tilhører mængden af borelfunktioner^a der afbilder $[0, \infty) \rightarrow [0, 1]$. Så gælder der at 1_n er givet rekursivt ved

$$1_n = \phi(1_{n-1}), \quad n \in \mathbb{N}, \quad (8.1)$$

$$F = \phi(F). \quad (8.2)$$

Hvor $1_0(t) = 1$ for $t \geq 0$ er basisskridtet.

[6, s. 637]

^aEn borelfunktion er en funktion der afbilder et borelrum over i et andet borelrum.

Bevis. Lad $T > 0$ og lad $n \in \mathbb{N}$. Betragt en klynge, der starter med en hændelse i tidspunktet 0, med tilhørende mærke k . Lad $\Phi(0)$ være en Hawkes punktproces, der genererer punkterne i klyngen. Split punktprocessen $\Phi(0)$ op i tre nye punktprocesser Φ_1 , Φ_2 og Φ_3 , hvor der gælder, at:

- Punktprocessen Φ_1 består af alle de første generations afkom $t_i \in \Phi(0) \cap [0, t)$, der ikke selv genererer afkom af $n-1$ generationer, eller mindre på intervallet $[t, \infty)$.
- Punktprocessen Φ_2 består af alle de første generations afkom i intervallet $[0, t)$, der ikke er med i Φ_1 .
- Punktprocessen Φ_3 består af de første generations afkom, der er i intervallet $[t, \infty)$.

Bemærk, at Φ_1 , Φ_2 og Φ_3 alle er første generations afkom af hændelsen i tidspunktet 0 med mærket k . Den betingede intensitetsfunktion for en mærket Hawkes punktproces med uforudsigelige mærker er givet ved definition 6.6, og det er muligt, at betragte Φ_1 , Φ_2 og Φ_3 som tre inhomogene Poisson punktprocesser. For at finde intensitetsfunktionen for Φ_1 bemærkes det, at punktet $(0, k)$, i en mærket Hawkes punktproces med uforudsigelige mærker, danner afkom efter en inhomogen Poisson punktproces med intensitetsfunktion $\alpha \gamma(s - 0, k)$. For kun at medtage de afkom, der ikke selv genererer afkom af $n-1$, eller færre generationer i intervallet $[t, \infty)$ multipliceres med fordelingsfunktionen for klyngelængderne $1_{n-1}(t-s)$.

Derved fås at sandsynligheden for de klynger, der genereres af afkommet af $(0, k)$, vil være så korte at de ikke strækker sig ind over tiden t . Dermed bliver intensitetsfunktionen for Φ_1 :

$$\lambda_1(s, k) = \alpha\gamma(s-0, k)1_{n-1}(t-s) = \alpha\gamma(s, k)1_{n-1}(t-s), \quad \text{på} \quad [0, t).$$

Ligeledes er det muligt at finde intensitetsfunktionen for Φ_2 , hvor der multipliceres med $(1 - 1_{n-1}(t-s))$ i stedet for $1_{n-1}(t-s)$, idet det netop ønskes at få de første generations hændelser i $[0, t]$, der ikke er med i Φ_1 . Intensitetsfunktionen for Φ_3 fremkommer direkte fra definitionen, på den betingede intensitetsfunktion for den mærkede Hawkes punktproces, med uforudsigelige mærker.

$$\begin{aligned} \lambda_2(s, k) &= \alpha\gamma(s-0, k)(1 - 1_{n-1}(t-s)) = \alpha\gamma(s, k)(1 - 1_{n-1}(t-s)), & \text{på} & \quad [0, t). \\ \lambda_3(s, k) &= \alpha\gamma(s-0, k) = \alpha\gamma(s, k), & \text{på} & \quad [t, \infty). \end{aligned}$$

Der gælder at:

$$\begin{aligned} 1_n(t) &= E[Pr\{\text{Alle punkter er i } \Phi_1\}] \\ &= E[Pr\{\Phi_2 = \emptyset\}Pr\{\Phi_3 = \emptyset\}] \end{aligned} \quad (8.3)$$

Følgende anvendes det, at sandsynligheden for ikke at have nogle punkter i intervallet $[0, x]$ er givet ved én minus fordelingsfunktionen. Da både Φ_2 og Φ_3 er Poisson punktprocesser er deres fordelingsfunktioner hhv. $1 - \exp(-\int_0^t \lambda_2(s, k)ds)$ og $1 - \exp(-\int_t^\infty \lambda_3(s, k)ds)$.

$$\begin{aligned} 1_n(t) &= E \left[\left(1 - \left(1 - \exp \left(- \int_0^t \lambda_2(s, k)ds \right) \right) \right) \left(1 - \left(1 - \exp \left(- \int_t^\infty \lambda_3(s, k)ds \right) \right) \right) \right] \\ &= E \left[\exp \left(- \int_0^t \lambda_2(s, k)ds \right) \exp \left(- \int_t^\infty \lambda_3(s, k)ds \right) \right] \\ &= E \left[\exp \left(- \int_0^t \alpha\gamma(s, k)(1 - 1_{n-1}(t-s))ds - \int_t^\infty \alpha\gamma(s, k)ds \right) \right] \\ &= E \left[\exp \left(\int_0^t 1_{n-1}(t-s)\alpha\gamma(s, k)ds - \int_0^t \alpha\gamma(s, k)ds - \int_t^\infty \alpha\gamma(s, k)ds \right) \right] \\ &= E \left[\exp \left(- \int_0^\infty \alpha\gamma(s, k)ds + \int_0^t 1_{n-1}(t-s)\alpha\gamma(s, k)ds \right) \right] \\ &= E \left[\exp \left(-v + \int_0^t 1_{n-1}(t-s)\alpha\gamma(s, k)ds \right) \right] \\ &= \phi(1_{n-1}(t)). \end{aligned}$$

Hvilket stemmer ovenes med ligning 8.1. For at vise ligning 8.2 lades $n \rightarrow \infty$.

$$\begin{aligned} \lim_{n \rightarrow \infty} 1_n &= \lim_{n \rightarrow \infty} \phi(1_{n-1}) \\ \Leftrightarrow \lim_{n \rightarrow \infty} 1_n &= \phi(\lim_{n \rightarrow \infty} 1_{n-1}) \\ \Leftrightarrow F &= \phi(F). \end{aligned}$$



Ifølge [6, s. 638] er det ofte svært/umuligt at bestemme en lukket form for F .

[6, s. 637-638]

8.2 Perfekt simulerings alortime

På baggrund af fordelingen af længderne af klyngerne, og algoritmerne fra afsnit 6.2, er det nu muligt at konstruere en algoritme, der laver *perfekt* simulering af en mærket Hawkes punktproces med uforudsigelige mærker.

Algoritme 6 Perfect simulering af Hawkes punktproces

Det ønskes, at generere et punktmønster for en mærket Hawkes punktproces med uforudsigelige mærker på intervallet $[0, T]$.

1. Simuler immigranterne I_1 , som en Poisson punktproces med intensitetsfunktion μ på $[0, T]$. Ydermere simuleres m_i fra mærkefordelingen for hvert $t_i \in I_1$.
 2. Simuler afkommet af immigranterne af forskellige generationer som inhomogene Poisson punktprocesser med intensitetsfunktion $\alpha\gamma(t - t_i, k_i)$ på $(t_i, T]$ for $i = 1, 2, \dots$
 3. Simuler I_2 som en inhomogen Poisson punktproces med intensitetsfunktion $\lambda_2 = (1 - F(-t))\mu$ på $(-\infty, 0)$. Ydermere simuleres m_i for hvert $t_i \in I_2$.
 4. Hvert punkt (t_i, m_i) i I_2 afføder en klynge C_i . Simuler $\{(t_j, m_j) \in C_i | t_j \in [0, T]\}$ betinget med at der for hvert punkt i I_2 er mindst ét afkom af en eller anden generation i $[0, T]$.
 5. Lad $\mathbf{X}$ være en superposition af de mærkede punktmønstre fra trin 1, 2 og 4.
-

Algoritme 6 er i stand til direkte at generere punktmønstre for stationære mærkede Hawkes punktprocesser med uforudsigelige mærker, uden nogen form for *burn-in* tid.

Sætning 8.2.

Punktmønstre konstrueret med algoritme 6 er udtræk af fordelingen af mærkede Hawkes punktprocesser.
[6, s. 636]

Bevis. Betragt først immigranterne på $[0, T]$, disse simuleres via en Poisson punktproces med intensitetsfunktion μ , og de tilhørende mærker trækkes fra mærkefordelingen, hvilket stemmer overens med skridt 1 i definition 6.7. Hvis der endvidere haves en Poisson punktproces på $(-\infty, 0)$ med intensitetsfunktion μ , og denne udtyndes med sandsynlighed for baving $(1 - F(-t))$, hvor $F(t)$ er fordelingsfunktionen for klyngelængden, giver dette en inhomogen Poisson punktproces med intensitetsfunktion $(1 - F(-t))\mu$, (I_2), der danner de punkter i intervallet $(-\infty, 0)$, der har afkom i $[0, T]$. Udtyndingen medfører at de punkter, der ikke har afkom, af en eller anden generation, i $[0, T]$ ikke medtages i I_2 . I punkt 2 og

4 simuleres klyngerne i intervallet $[0, T]$, hvilket stemmer overens med punkt 3 i definition 6.7. Punkt 5 er en superposition af skridt 1, 2 og 4, hvilket stemmer overens med skridt 4 i definition 6.7. Dermed giver algoritme 6 punktmønstre fra en mærket Hawkes punktproces. $\square$

Der er et par ting, der er væsentligt at være opmærksom på, hvis det ønskes at anvende algoritme 6 til perfekt simulering. Skridt 1 og 2 er relativt simple ved brug af f.eks. algoritme 5, men i skridt 4 er det nødvendigt, at være sikker på at mindst én hændelse er efter $t = 0$. Dette kan løses ved gentagen simulering af punktmønstrene, hvor hver simulering forkastes, hvis den ikke har en hændelse i intervallet $[0, T]$. Skridt 5 i algoritmen er en trivial superposition, og er derfor ikke problematisk. Det vanskelige er skridt 3, der afhænger af fordelingsfunktionen af længden af klyngerne, der ikke nødvendigvis har en lukket form. Følgende præsenteres en algoritme til simulering af I_2 .

[6, s. 635-636]

8.3 Generering af I_2

Det ønskes at konstruere en algoritme, der er i stand til at generere I_2 . Ideen bag algoritmen er, at der benyttes to forskellige Poisson punktprocesser L_n og U_n på intervallet $(-\infty, 0)$, hvoraf intensitetsfunktionen for L_n er mindre end intensitetsfunktionen for I_2 , og intensitetsfunktionen for U_n er større, men at de begge konvergerer monotont imod intensitetsfunktionen for I_2 når $n \rightarrow \infty$. Først er det dog nødvendigt at sikre, at I_2 er endelig og derfor antages følgende

$$\int_{-\infty}^0 (1 - F(-t))\mu < \infty.$$

Dermed opnås det, at I_2 klemmes inde mellem L_n og U_n for $n = 1, \dots, \infty$ på følgende måde

$$\emptyset = L_0 \subseteq L_1 \subseteq \dots \subseteq I_2 \subseteq \dots \subseteq U_1 \subseteq U_0. \quad (8.4)$$

Det er så muligt at bestemme I_2 ved at finde U_n eller L_n for et tilstrækkeligt stort n . For at undgå, at skulle vurdere hvornår et n er tilstrækkeligt stort, findes U_n og L_n for $n = 1, 2, \dots$ indtil $U_k = L_k$ for $k \in \mathbb{N}$.

Det er nødvendigt at finde intensitetsfunktioner for L_n og U_n , således at ovenstående er opfyldt. Det er dermed nødvendigt at finde to funktioner, der begge konvergerer monotont imod $F(t)$, hvor den ene skal være voksende og den anden skal være aftagende. Der gælder at 1_n , funktionen fra sætning 8.1, er aftagende når n stiger og konvergerer imod $F(t)$, idet der tages flere og flere generationer med, og klyngelængderne derfor kun vil vokse. Dermed gælder der, at

$$\lambda_n^l(t) = (1 - 1_n(-t))\mu,$$

er voksende som funktion af n og konvergere imod $(1 - F(-t))\mu$.

For at finde intensitetsfunktionen for U_n anvendes en funktion f , hvor der gælder at f_n er monotont voksende og konvergerer imod $F(t)$ som funktion af n . Dermed gælder der, at

$$\lambda_n^u(t) = (1 - f_n(-t))\mu,$$

er aftagende som funktion af n og konvergerer imod $(1 - F(-t))\mu$. I denne rapport vil der ikke blive fundet en eksplicit funktion for f .

Følgende præsenteres en algoritme, der er i stand til at simulere I_2 . Ideen er at simulere et mærket punktmønster med for stor intensitet, og så udtynde dette indtil det har opnået den ønskede intensitet.

Algoritme 7 Simulering af I_2

1. Simuler et mærket punktmønster $\{(t_1, k_1), (t_2, k_2), \dots, (t_s, k_s)\}$ for U_0 .
 2. Hvis $U_0 = \emptyset$ så er $I_2 = \emptyset$, ellers generer $W_i \sim \text{unif}[0, 1]$ for $i = 1, 2, \dots, s$ og lad $n=1$.
 - while** $L_n \neq U_n$ **do**
 3. For $j = 1, \dots, s$ tilføj (t_j, k_j) til L_n , hvis $W_j \lambda_0^u(t_j) \leq \lambda_n^l(t_j)$, og til U_n hvis $W_j \lambda_0^u(t_j) \leq \lambda_n^u(t_j)$
 4. Lad $n = n + 1$.
 - end while**
 5. $I_2 = L_n$
-

Der gælder, at algoritmen er korrekt, og at den ikke fortsætter uendeligt længe.

Sætning 8.3.

Algoritme 7 simulerer I_2 og den terminerer.

[6, s. 642]

Bevis. Givet punkterne $t_1, t_2, \dots, t_s$ er sandsynligheden for at $L_n \neq U_n$ givet ved:

$$\begin{aligned} &Pr\{\lim_{n \rightarrow \infty} L_n \neq U_n\} \\ &\leq \sum_{j=1}^s \left(Pr\{\lim_{n \rightarrow \infty} (t_j \in U_n, t_j \notin L_n)\} + Pr\{\lim_{n \rightarrow \infty} (t_j \notin U_n, t_j \in L_n)\} \right) \end{aligned} \quad (8.5)$$

$$= \sum_{j=1}^s \left(Pr\{\lim_{n \rightarrow \infty} (t_j \in U_n, t_j \notin L_n)\} \right) \quad (8.6)$$

$$= \sum_{j=1}^s \left(Pr\{\lim_{n \rightarrow \infty} (W_j \lambda_0^u(t_j) \leq \lambda_n^u(t_j), W_j \lambda_0^u(t_j) > \lambda_n^l(t_j))\} \right)$$

$$= \sum_{j=1}^s \left(Pr\{\lambda_2(t_j) < W_j \lambda_0^u(t_j) \leq \lambda_2(t_j)\} \right)$$

$$= 0.$$

For at få uligheden i ligning (8.5) er det anvendt at fællesmængden ikke er trukket fra. For at få (8.6) er det anvendt, at der gælder at $Pr\{t_j \notin U_n, t_j \in L_n\} = 0$ da det jvf. (8.4) ikke er muligt for L_n at indeholde hændelser der ikke også er i U_n .

Dermed terminerer algoritmen indenfor endelig tid.

På grund af indlejringsegenskaben (8.4), gælder der endvidere at når $L_n = U_n$ er de begge lig I_2 . $\square$

Dermed haves en algoritme, der er istand til at simulere I_2 , og dette muliggør at algoritme 6 kan anvendes til at lave *perfekt* simulering af Hawkes punktprocessen.

[6, s. 642-644]

Der er igennem rapporten blevet konstrueret to forskellige modeller for jordskælvshændelser. Begge disse modeller er blevet holdt op imod de virkeligheden. Dette betyder dog ikke, at der ikke kan findes bedre modeller for hændelserne. Antallet og placeringen af efterskælv er i stor grad afhængig af funktionen $\gamma(\cdot)$, og hvis denne funktion ikke afspejler virkelighedens tendenser mht. efterskælv, vil modellen heller ikke gøre det. Der er i projektet blevet valgt to forskellige funktioner for $\gamma(\cdot)$, en for den umærkede model og en for den mærkede. Disse to funktioner er blevet valgt på baggrund af lidt grundlæggende teori omkring jordskælvshændelser, og på en sådan måde, at de er relativt lette at arbejde med i databehandlingen.

En anden problematik der optræder, er i forbindelse med parameterestimationen. Problemet skyldes, at *relativt* kraftige jordskælv måske kun indtræffer med et middeltidsinterval på 100 eller 1000 år, alt afhængig af styrken. Men med et udvalg af data på et tidsinterval på 4 år, er der således ikke stor sandsynlighed for, at de *relativt* kraftige jordskælv forefindes i denne tidsperiode.

Den sidste problematik, der vil blive nævnt her, er i forhold til den måde, der bliver foretaget modelkontrol på. Jvf. afsnit 4.4 ses det, at fordelingen af punkterne ikke afviger voldsomt fra en Poisson punktproces med intensitet $\mu = 354/4$. Det er derfor værd at overveje, om transformationerne til punktmønstret bestående af $\{s_i\}$, i afsnit 5.3 og 7.4, ikke ofte vil minde om punktmønstre fra en Poisson punktproces med intensitet 1, med mindre modellen var helt hen i vejret.

Når det så er sagt, er der konstrueret to modeller, der ser ud til at passe med det empiriske datasæt. Det er derfor relevant at overveje, hvad en sådan model kan bruges til. Et eksempel på en anvendelsesmulighed er inden for risikovurdering eller konstruktion. Hvis det f.eks. ønskes at bygge en dæmning, der skal være i stand til at klare et 1000 års jordskælv¹, er det muligt at lave et antal simuleringer, og på baggrund af dette bestemme hvilken styrke af jordskælv, der har et middeltidsinterval på 1000 år. En anden anvendelsesmulighed er inden for risikovurdering. Når en sådan skal laves betegnes risiko ofte som sandsynlighed

¹Et jordskælv der er så kraftigt, at det har et middeltidsinterval på 1000 år.

multipliseret med konsekvenserne for en hændelse. Modellen kan anvendes til at bestemme sandsynligheder for jordskælvshændelser af forskellige styrker.

En ting modellerne ikke kan bruges til, er at forudsige jordskælv på bestemte tidspunkter. Det er ikke muligt at køre en simulering, på baggrund af modellen, og så sige at der vil være et jordskælv 17 Januar 2013, bare fordi der er det i en enkelt simulation. Derudover er det værd at huske, at simuleringsalgoritmerne fra afsnit 3.5 og 6.2, ikke simulerer stationære punktmønstre, og derfor tilrådes det altid, enten at simulere med en *burn-in* periode, eller anvende perfekt simulering, som præsenteret i afsnit 8.

Det er dog vigtigt at være opmærksom på, at hvis det ønskes at anvende en model, så er konklusionerne aldrig bedre end modellen. Hvis denne ikke passer til virkeligheden vil analyserne baseret på modellen sandsynligvis heller ikke gøre det.

Formålet med dette projekt var at udvide min viden omkring Hawkes punktprocessen på tidslinien, samt hvorledes denne kan bruges til at konstruere modeller af virkelige hændelser. Der er blevet præsenteret lidt grundlæggende teori, der blandt andet ligger til grund for konstruktionen af modellen.

Efterfølgende præsenteres de rumlige punktprocesser på tidslinien, og nogle specielle egenskaber ved disse. Den betingede intensitetsfunktion bliver introduceret, og det bliver bevist hvorledes likelihoodfunktionen kan findes med denne. Der bliver lagt et stykke arbejde i, hvorledes det er muligt, at lave inferens med en Newton-Rhapson algoritme. Hawkes punktprocessen på tidslinien præsenteres, og et par algoritmer til simulation introduceres.

Der laves nogle indledende undersøgelser af et datasæt for jordskælvshændelser, og efterfølgende konstrueres en model for jordskælvshændelserne, ved brug af en umærket Hawkes punktproces.

For at udvide modellen til at medtage styrkerne af hændelserne, præsenteres de mærkede punktprocesser på tidslinien, og de mærkede Hawkes punktprocesser med uforudsigelige mærker defineres. To algoritmer, til simulation af mærkede Hawkes punktprocesser med uforudsigelige mærker, præsenteres. Derefter konstrueres en mærket model for jordskælvs-hændelserne, baseret på en Hawkes punktproces med uforudsigelige mærker, og parametrene i dennes estimeres.

Dernæst præsenteres det hvorledes det er muligt, at lave perfekt simulering fra mærkede Hawkes punktprocesser med uforudsigelige mærker. Således er det muligt at simulere stationære punktmønstre, i stedet for at skulle simulere punktmønstre med en *burn-in* periode i starten af intervallet.

- [1] Stephen C. Porter og Jeffrey Park Brian J. Skinner. *Dynamic Earth: An Introduction to Physical Geology*. Wiley-Interscience, femte edition, 2004. ISBN: 978-0-471-15228-6.
- [2] GEUS. Registrerede jordskælv. <http://www.geus.dk/>, besøgt d. 13/3-2011 kl. 15.21.
- [3] D. J. Daley og D. Vera-Jones. *An introduction to the Theory of Point Processes: Volume I: Elementary Theory and Methods*. Springer, 2002. ISBN: 0-387-95541-0.
- [4] C. Henry Edwards og David E. Penny. *Calculus*. Prentice Hall, 2002. ISBN: 0-13-095006-8.
- [5] Alan G. Hawkes og David Oakes. A cluster process representation of a self-exciting process. *Journal of Applied Probability*, 11(3), 1974.
- [6] Jesper Møller og Jakob G. Rasmussen. Perfect simulation of hawkes processes. *Ad. Appl. Prob.*, 37(side 629-646), 2005.
- [7] Jesper Møller og Jakob G. Rasmussen. Approximate simulation of hawkes processes. *methodol. Comput. Appl. Probab.*, 8(side 53-64), 2006.
- [8] Jesper Møller og Rasmus Plenge Waagepetersen. *Statistical Inference and Simulation for Spatial Point Processes*. Chapman and Hall/CRC, 2004. ISBN: 1-58488-265-4.
- [9] Peter Olofsson. *Probability, Statistics, and Stochastic Processes*. Wiley-Interscience, 2005. ISBN: 0-471-67969-0.
- [10] T. Ozaki. Maximum likelihood estimation of hawkes' self-exciting point processes. *Ann. Inst. Statist. Math.*, 31(side 145-155), 1979.
- [11] J. Larsen S. Kit og R. Graungaard R. S. Christophersen, C. Bisgaard. Rumlige punktprocesser. Skrevet på mat4 semestret 2010.
- [12] Jakob Gulddahl Rasmussen. Bayesian inference for hawkes processes. *Intern revision* <http://www.math.aau.dk/index.php?id=244/>.

- [13] Jakob Gulddahl Rasmussen. Kursus slides. <http://people.math.aau.dk/~jgr/teaching/punktproc10/punktproc10.html>, besøgt d. 23/5-2011 kl. 15.24.
- [14] Jakob Gulddahl Rasmussen. Kursus slides. <http://people.math.aau.dk/~jgr/teaching/punktproc10/punktproc10.html>, besøgt d. 23/5-2011 kl. 15.24.
- [15] Peter R. Turner. *Guide 2 scientific computing*. MACMILLAN PRESS LTD, 2000. ISBN: 978-0-333-79450-0.
- [16] William R. Wade. *An Introduction to Analysis*. Pearson Prentice Hall, 2004. ISBN: 0-13-14-5333-5.
- [17] Wikipedia. [www.wikipedia.org. http://en.wikipedia.org/wiki/Borel%E2%80%9393Cantelli_lemma](http://en.wikipedia.org/wiki/Borel%E2%80%9393Cantelli_lemma), besøgt d. 6/5-2011 kl. 14.03.
- [18] Wikipedia. [www.wikipedia.org. http://en.wikipedia.org/wiki/Hessian_matrix](http://en.wikipedia.org/wiki/Hessian_matrix), besøgt d. 29/5-2011 kl. 12.49.

Følgende præsenteres de forskellige bilag.

11.1 Borelmængde

Følgende defineres en borelmængde.

Definition 11.1. – Borelmængde

Hvis mængden $\mathcal{B}$ er mængden af alle Borelmængder, så er en Borelmængde $B \subseteq S$ enhver mængde, der kan konstrueres fra åbne delmængder af S ved brug af følgende to punkter:

- Hvis $B \in \mathcal{B}$, så gælder der, at $S \setminus B \in \mathcal{B}$.
- Hvis $B_i \in \mathcal{B}$ for $i = 1, 2, \dots$, så er $\bigcup_{i=1}^{\infty} B_i \in \mathcal{B}$.

Mængden $\mathcal{B}_0$ er mængden af alle begrænsede Borelmængder.

[14, Kursusgang 1 slide 4]

11.2 Additivitet

Følgende defineres additivitet, og et enkelt lemma præsenteres og bevises.

Definition 11.2. – Additivitet

Hvis der haves en funktion $f(x)$, hvor $0 < x < \infty$ så kaldes den:

- Additiv hvis $f(x+y) = f(x) + f(y)$.
- Subadditiv hvis $f(x+y) \leq f(x) + f(y)$.
- Superadditiv hvis $f(x+y) \geq f(x) + f(y)$.

[3, s. 64-65]

For en subadditiv funktion gælder der følgende.

Lemma 11.3.

Hvis der findes en subadditiv funktion $f(x)$, hvor x er begrænset af intervallet $[0, a]$ hvor $a > 0$, og hvor $f(x) \rightarrow 0$ når $x \rightarrow 0$, så gælder der at $\lambda \equiv \sup_{x>0} \frac{f(x)}{x}$ er endelig eller ∞ , og

$$\frac{f(x)}{x} \rightarrow \lambda, \quad \text{når } x \rightarrow 0.$$

[3, s. 64-65]

Lignende lemmaer kan konstrueres for de superadditive og de additive funktioner, men de præsenteres ikke, da de ikke anvendes i rapporten.

11.3 Tayler Polynomier

Hvis der findes en funktionsværdi i et punkt x_0 og det ønskes at finde en approksimeret værdi i et andet punkt x så er det muligt at anvende et Taylor polynomium. Først præsenteres en 1-dimensionel N 'te ordenstaylor polynomium.

Definition 11.4. – N 'te ordens Tayler polynomium

Taylor polynomiet af N 'te grad er givet ved

$$f(x) \approx f(x_0) + (x - x_0)f'(x_0) + \dots + \frac{(x - x_0)^N}{N!} f^{(N)}(x_0),$$

hvor x_0 er startpunktet.

Der findes også en n -dimensionel udgave, der er dog kun taylor polynomiet af første orden der vil blive præsenteret her.

Definition 11.5. – n -dimensionel 1. ordens Tayler polynomium

Det n -dimensionelle Taylor polynomium af 1. orden er givet ved

$$f(x+h) \approx f(x) + Jh,$$

hvor x er startpunktet og J er jakobimatricen for f i x .

[15, s. 45, 77]

11.4 Partielt afledte af den umærkede model

I det følgende findes de partielt afledte af log-likelihoodfunktionen (5.4), der anvendes i afsnit 5.

$$\begin{aligned}
 \frac{\partial}{\partial \mu} \mathcal{L}(\mu, \alpha, \beta) &= \frac{\partial}{\partial \mu} \left(\left(\sum_{i=1}^n \log(\mu + \alpha A(i)) \right) - \mu T + \left(\sum_{i=1}^n \frac{\alpha}{\beta} (\exp(-\beta(T - t_i)) - 1) \right) \right) \\
 &= \frac{\partial}{\partial \mu} \left(\sum_{i=1}^n \log(\mu + \alpha A(i)) \right) - \frac{\partial}{\partial \mu} \mu T + \frac{\partial}{\partial \mu} \left(\sum_{i=1}^n \frac{\alpha}{\beta} (\exp(-\beta(T - t_i)) - 1) \right) \\
 &= \left(\sum_{i=1}^n \frac{\partial}{\partial \mu} \log(\mu + \alpha A(i)) \right) - T \\
 &= \left(\sum_{i=1}^n \frac{1}{\mu + \alpha A(i)} \right) - T,
 \end{aligned}$$

hvor

$$A(i) = \begin{cases} 0, & \text{for } i = 1 \\ \sum_{t_j < t_i} \exp(-\beta(t_i - t_j)) & \text{for } i \geq 2 \end{cases}.$$

$$\begin{aligned}
 \frac{\partial}{\partial \alpha} \mathcal{L}(\mu, \alpha, \beta) &= \frac{\partial}{\partial \alpha} \left(\left(\sum_{i=1}^n \log(\mu + \alpha A(i)) \right) - \mu T + \left(\sum_{i=1}^n \frac{\alpha}{\beta} (\exp(-\beta(T - t_i)) - 1) \right) \right) \\
 &= \frac{\partial}{\partial \alpha} \left(\sum_{i=1}^n \log(\mu + \alpha A(i)) \right) - \frac{\partial}{\partial \alpha} \mu T + \frac{\partial}{\partial \alpha} \left(\sum_{i=1}^n \frac{\alpha}{\beta} (\exp(-\beta(T - t_i)) - 1) \right) \\
 &= \left(\sum_{i=1}^n \frac{\partial}{\partial \alpha} \log(\mu + \alpha A(i)) \right) + \left(\sum_{i=1}^n \frac{\partial}{\partial \alpha} \frac{\alpha}{\beta} (\exp(-\beta(T - t_i)) - 1) \right) \\
 &= \left(\sum_{i=1}^n \frac{A(i)}{\mu + \alpha A(i)} \right) + \left(\sum_{i=1}^n \frac{1}{\beta} (\exp(-\beta(T - t_i)) - 1) \right)
 \end{aligned}$$

$$\begin{aligned}
\frac{\partial}{\partial \beta} \mathcal{L}(\mu, \alpha, \beta) &= \frac{\partial}{\partial \beta} \left(\left(\sum_{i=1}^n \log(\mu + \alpha A(i)) \right) - \mu T + \left(\sum_{i=1}^n \frac{\alpha}{\beta} (\exp(-\beta(T - t_i)) - 1) \right) \right) \\
&= \frac{\partial}{\partial \beta} \left(\sum_{i=1}^n \log(\mu + \alpha A(i)) \right) - \frac{\partial}{\partial \beta} \mu T + \frac{\partial}{\partial \beta} \left(\sum_{i=1}^n \frac{\alpha}{\beta} (\exp(-\beta(T - t_i)) - 1) \right) \\
&= \left(\sum_{i=1}^n \frac{\partial}{\partial \beta} \log(\mu + \alpha A(i)) \right) + \alpha \left(\sum_{i=1}^n \frac{\partial}{\partial \beta} \frac{\exp(-\beta(T - t_i)) - 1}{\beta} \right) \\
&= \left(\sum_{i=1}^n \frac{-\alpha B(i)}{\mu + \alpha A(i)} \right) + \alpha \left(\sum_{i=1}^n \frac{-(T - t_i) \exp(-\beta(T - t_i)) \beta - (\exp(-\beta(T - t_i)) - 1)}{\beta^2} \right) \\
&= \left(\sum_{i=1}^n \frac{-\alpha B(i)}{\mu + \alpha A(i)} \right) - \alpha \left(\sum_{i=1}^n \frac{(T - t_i) \exp(-\beta(T - t_i))}{\beta} + \frac{(\exp(-\beta(T - t_i)) - 1)}{\beta^2} \right)
\end{aligned}$$

hvor

$$-B(i) = \frac{\partial}{\partial \beta} A(i) = \begin{cases} 0, & \text{for } i = 1 \\ \sum_{t_j < t_i} \exp(k_j)(t_i - t_j) \exp(-\beta(t_i - t_j)) & \text{for } i \geq 2 \end{cases}$$

$$\begin{aligned}
\frac{\partial^2}{\partial \mu^2} \mathcal{L}(\mu, \alpha, \beta) &= \frac{\partial}{\partial \mu} \frac{\partial}{\partial \mu} \mathcal{L}(\mu, \alpha, \beta) \\
&= \frac{\partial}{\partial \mu} \left(\left(\sum_{i=1}^n \frac{1}{\mu + \alpha A(i)} \right) - T \right) \\
&= \frac{\partial}{\partial \mu} \left(\sum_{i=1}^n \frac{1}{\mu + \alpha A(i)} \right) - \frac{\partial}{\partial \mu} T \\
&= \left(\sum_{i=1}^n \frac{\partial}{\partial \mu} \frac{1}{\mu + \alpha A(i)} \right) \\
&= \sum_{i=1}^n \frac{-1}{(\mu + \alpha A(i))^2}.
\end{aligned}$$

$$\begin{aligned}
\frac{\partial^2}{\partial \alpha^2} \mathcal{L}(\mu, \alpha, \beta) &= \frac{\partial}{\partial \alpha} \frac{\partial}{\partial \alpha} \mathcal{L}(\mu, \alpha, \beta) \\
&= \frac{\partial}{\partial \alpha} \left(\left(\sum_{i=1}^n \frac{A(i)}{\mu + \alpha A(i)} \right) + \left(\sum_{i=1}^n \frac{1}{\beta} (\exp(-\beta(T - t_i)) - 1) \right) \right) \\
&= \frac{\partial}{\partial \alpha} \left(\sum_{i=1}^n \frac{A(i)}{\mu + \alpha A(i)} \right) + \frac{\partial}{\partial \alpha} \left(\sum_{i=1}^n \frac{1}{\beta} (\exp(-\beta(T - t_i)) - 1) \right) \\
&= \left(\sum_{i=1}^n \frac{\partial}{\partial \alpha} \frac{A(i)}{\mu + \alpha A(i)} \right) \\
&= \sum_{i=1}^n \frac{-A(i)^2}{(\mu + \alpha A(i))^2}.
\end{aligned}$$

$$\begin{aligned}
\frac{\partial^2}{\partial \beta^2} \mathcal{L}(\mu, \alpha, \beta) &= \frac{\partial}{\partial \beta} \frac{\partial}{\partial \beta} \mathcal{L}(\mu, \alpha, \beta) \\
&= \frac{\partial}{\partial \beta} \left(\left(\sum_{i=1}^n \frac{-\alpha B(i)}{\mu + \alpha A(i)} \right) \right. \\
&\quad \left. - \alpha \left(\sum_{i=1}^n \frac{(T - t_i) \exp(-\beta(T - t_i))}{\beta} + \frac{(\exp(-\beta(T - t_i)) - 1)}{\beta^2} \right) \right) \\
&= \frac{\partial}{\partial \beta} \left(\sum_{i=1}^n \frac{-\alpha B(i)}{\mu + \alpha A(i)} \right) \\
&\quad - \frac{\partial}{\partial \beta} \alpha \left(\sum_{i=1}^n \frac{(T - t_i) \exp(-\beta(T - t_i))}{\beta} + \frac{\exp(-\beta(T - t_i)) - 1}{\beta^2} \right) \\
&= \left(\sum_{i=1}^n \frac{\partial}{\partial \beta} \frac{-\alpha B(i)}{\mu + \alpha A(i)} \right) \\
&\quad - \alpha \left(\sum_{i=1}^n \frac{\partial}{\partial \beta} \left(\frac{(T - t_i) \exp(-\beta(T - t_i))}{\beta} + \frac{\exp(-\beta(T - t_i)) - 1}{\beta^2} \right) \right) \\
&= \left(\sum_{i=1}^n \frac{-\frac{\partial}{\partial \beta} (\alpha B(i)) \cdot (\mu + \alpha A(i)) + B(i) \cdot \alpha^2 \frac{\partial}{\partial \beta} A(i)}{(\mu + \alpha A(i))^2} \right) \\
&\quad - \alpha \left(\sum_{i=1}^n \frac{\partial}{\partial \beta} \frac{(T - t_i) \exp(-\beta(T - t_i))}{\beta} + \frac{\partial}{\partial \beta} \frac{\exp(-\beta(T - t_i)) - 1}{\beta^2} \right) \\
&= \left(\sum_{i=1}^n \frac{\alpha C(i) \cdot (\mu + \alpha A(i)) - \alpha^2 B(i)^2}{(\mu + \alpha A(i))^2} \right) \\
&\quad - \alpha \left(\sum_{i=1}^n \frac{-(T - t_i)^2 \exp(-\beta(T - t_i)) \beta - (T - t_i) \exp(-\beta(T - t_i))}{\beta^2} \right. \\
&\quad \left. + \left(\frac{-(T - t_i) \exp(-\beta(T - t_i)) \beta^2 - (\exp(-\beta(T - t_i)) - 1) 2\beta}{\beta^4} \right) \right) \\
&= \left(\sum_{i=1}^n \frac{\alpha C(i)}{\mu + \alpha A(i)} - \frac{\alpha^2 B(i)^2}{(\mu + \alpha A(i))^2} \right) \\
&\quad - \alpha \left(\sum_{i=1}^n \frac{-(T - t_i)^2 \exp(-\beta(T - t_i))}{\beta} - \frac{(T - t_i) \exp(-\beta(T - t_i))}{\beta^2} \right. \\
&\quad \left. + \left(\frac{-(T - t_i) \exp(-\beta(T - t_i))}{\beta^2} - \frac{(\exp(-\beta(T - t_i)) - 1) 2}{\beta^3} \right) \right) \\
&= \left(\sum_{i=1}^n \frac{\alpha C(i)}{\mu + \alpha A(i)} - \frac{\alpha^2 B(i)^2}{(\mu + \alpha A(i))^2} \right) \\
&\quad + \alpha \left(\sum_{i=1}^n \frac{(T - t_i)^2 \exp(-\beta(T - t_i))}{\beta} + \frac{2(T - t_i) \exp(-\beta(T - t_i))}{\beta^2} \right. \\
&\quad \left. + \frac{2(\exp(-\beta(T - t_i)) - 1)}{\beta^3} \right),
\end{aligned}$$

hvor

$$-C(i) = \frac{\partial}{\partial \beta} B(i) = \begin{cases} 0, & \text{for } i = 1 \\ \sum_{t_j < t_i} \exp(k_j)(t_i - t_j)^2 \exp(-\beta(t_i - t_j)) & \text{for } i \geq 2 \end{cases}$$

$$\begin{aligned} \frac{\partial^2}{\partial^2 \mu \alpha} \mathcal{L}(\mu, \alpha, \beta) &= \frac{\partial}{\partial \mu} \frac{\partial}{\partial \alpha} \mathcal{L}(\mu, \alpha, \beta) \\ &= \frac{\partial}{\partial \mu} \left(\left(\sum_{i=1}^n \frac{A(i)}{\mu + \alpha A(i)} \right) + \left(\sum_{i=1}^n \frac{1}{\beta} (\exp(-\beta(T - t_i)) - 1) \right) \right) \\ &= \frac{\partial}{\partial \mu} \left(\sum_{i=1}^n \frac{A(i)}{\mu + \alpha A(i)} \right) + \frac{\partial}{\partial \mu} \left(\sum_{i=1}^n \frac{1}{\beta} (\exp(-\beta(T - t_i)) - 1) \right) \\ &= \sum_{i=1}^n \frac{\partial}{\partial \mu} \frac{A(i)}{\mu + \alpha A(i)} \\ &= \sum_{i=1}^n \frac{-A(i)}{(\mu + \alpha A(i))^2}. \end{aligned}$$

$$\begin{aligned} \frac{\partial^2}{\partial^2 \mu \beta} \mathcal{L}(\mu, \alpha, \beta) &= \frac{\partial}{\partial \beta} \frac{\partial}{\partial \mu} \mathcal{L}(\mu, \alpha, \beta) \\ &= \frac{\partial}{\partial \beta} \left(\left(\sum_{i=1}^n \frac{1}{\mu + \alpha A(i)} \right) - T \right) \\ &= \frac{\partial}{\partial \beta} \left(\sum_{i=1}^n \frac{1}{\mu + \alpha A(i)} \right) - \frac{\partial}{\partial \beta} T \\ &= \left(\sum_{i=1}^n \frac{\partial}{\partial \beta} \frac{1}{\mu + \alpha A(i)} \right) \\ &= \sum_{i=1}^n \frac{-\frac{\partial}{\partial \beta} (\mu + \alpha A(i))}{(\mu + \alpha A(i))^2} \\ &= \sum_{i=1}^n \frac{-\left(\frac{\partial}{\partial \beta} \mu + \frac{\partial}{\partial \beta} \alpha A(i) \right)}{(\mu + \alpha A(i))^2} \\ &= \sum_{i=1}^n \frac{-\alpha \frac{\partial}{\partial \beta} A(i)}{(\mu + \alpha A(i))^2} \\ &= \sum_{i=1}^n \frac{\alpha B(i)}{(\mu + \alpha A(i))^2}. \end{aligned}$$

$$\begin{aligned}
\frac{\partial^2}{\partial^2 \alpha \beta} \mathcal{L}(\mu, \alpha, \beta) &= \frac{\partial}{\partial \beta} \frac{\partial}{\partial \alpha} \mathcal{L}(\mu, \alpha, \beta) \\
&= \frac{\partial}{\partial \beta} \left(\left(\sum_{i=1}^n \frac{A(i)}{\mu + \alpha A(i)} \right) + \left(\sum_{i=1}^n \frac{1}{\beta} (\exp(-\beta(T - t_i)) - 1) \right) \right) \\
&= \frac{\partial}{\partial \beta} \left(\sum_{i=1}^n \frac{A(i)}{\mu + \alpha A(i)} \right) + \frac{\partial}{\partial \beta} \left(\sum_{i=1}^n \frac{1}{\beta} (\exp(-\beta(T - t_i)) - 1) \right) \\
&= \left(\sum_{i=1}^n \frac{\partial}{\partial \beta} \frac{A(i)}{\mu + \alpha A(i)} \right) + \left(\sum_{i=1}^n \frac{\partial}{\partial \beta} \frac{1}{\beta} (\exp(-\beta(T - t_i)) - 1) \right) \\
&= \left(\sum_{i=1}^n \frac{\frac{\partial}{\partial \beta} A(i) \cdot (\mu + \alpha A(i)) - A(i) \frac{\partial}{\partial \beta} (\mu + \alpha A(i))}{(\mu + \alpha A(i))^2} \right) \\
&\quad + \left(\sum_{i=1}^n \frac{-(T - t_i) \exp(-\beta(T - t_i)) \beta - (\exp(-\beta(T - t_i)) - 1)}{\beta^2} \right) \\
&= \left(\sum_{i=1}^n \frac{-B(i)(\mu + \alpha A(i)) - A(i) \left(\frac{\partial}{\partial \beta} \mu + \frac{\partial}{\partial \beta} \alpha A(i) \right)}{(\mu + \alpha A(i))^2} \right) \\
&\quad + \left(\sum_{i=1}^n \frac{-(T - t_i) \exp(-\beta(T - t_i))}{\beta} - \frac{\exp(-\beta(T - t_i)) - 1}{\beta^2} \right) \\
&= \left(\sum_{i=1}^n \frac{-B(i)}{(\mu + \alpha A(i))} + \frac{A(i) \alpha B(i)}{(\mu + \alpha A(i))^2} \right) \\
&\quad - \left(\sum_{i=1}^n \frac{-(T - t_i) \exp(-\beta(T - t_i))}{\beta} + \frac{\exp(-\beta(T - t_i)) - 1}{\beta^2} \right).
\end{aligned}$$

11.5 Pertielt afledte af den mærkede model

I det følgende findes de afledte af log-likelihoodfunktionen, (7.7), der anvendes i afsnit 7.3. Bemærk at en del af udgrigningerne minder om dem fra appendiks 11.4.

$$\begin{aligned}
\frac{\partial}{\partial \mu} \mathcal{L}(\mu, \alpha, \beta, \eta) &= \frac{\partial}{\partial \mu} \left(\left(\sum_{i=1}^n \log(\eta) + (-\eta(k_i - 2)) \right) + \left(\sum_{i=1}^n \log(\mu + \alpha A(i)) \right) \right. \\
&\quad \left. - \mu T - \left(\sum_{i=1}^n \alpha \exp(k_i) \frac{1 - \exp(-\beta(T - t_i))}{\beta} \right) \right) \\
&= \frac{\partial}{\partial \mu} \left(\sum_{i=1}^n \log(\eta) + (-\eta(k_i - 2)) \right) + \frac{\partial}{\partial \mu} \left(\sum_{i=1}^n \log(\mu + \alpha A(i)) \right) \\
&\quad - \frac{\partial}{\partial \mu} \mu T - \frac{\partial}{\partial \mu} \left(\sum_{i=1}^n \alpha \exp(k_i) \frac{1 - \exp(-\beta(T - t_i))}{\beta} \right) \\
&= \left(\sum_{i=1}^n \frac{\partial}{\partial \mu} \log(\mu + \alpha A(i)) \right) - T \\
&= \left(\sum_{i=1}^n \frac{1}{\mu + \alpha A(i)} \right) - T,
\end{aligned}$$

hvor

$$A(i) = \begin{cases} 0, & \text{for } i = 1 \\ \left(\sum_{t_j < t_i} \exp(k_j) \exp(-\beta(t_i - t_j)) \right) & \text{for } i \geq 2 \end{cases}.$$

$$\begin{aligned}
\frac{\partial}{\partial \alpha} \mathcal{L}(\mu, \alpha, \beta, \eta) &= \frac{\partial}{\partial \alpha} \left(\left(\sum_{i=1}^n \log(\eta) + (-\eta(k_i - 2)) \right) + \left(\sum_{i=1}^n \log(\mu + \alpha A(i)) \right) \right. \\
&\quad \left. - \mu T - \left(\sum_{i=1}^n \alpha \exp(k_i) \frac{1 - \exp(-\beta(T - t_i))}{\beta} \right) \right) \\
&= \frac{\partial}{\partial \alpha} \left(\sum_{i=1}^n \log(\eta) + (-\eta(k_i - 2)) \right) + \frac{\partial}{\partial \alpha} \left(\sum_{i=1}^n \log(\mu + \alpha A(i)) \right) \\
&\quad - \frac{\partial}{\partial \alpha} \mu T - \frac{\partial}{\partial \alpha} \left(\sum_{i=1}^n \alpha \exp(k_i) \frac{1 - \exp(-\beta(T - t_i))}{\beta} \right) \\
&= \left(\sum_{i=1}^n \frac{\partial}{\partial \alpha} \log(\mu + \alpha A(i)) \right) - \frac{\partial}{\partial \alpha} \alpha \left(\sum_{i=1}^n \exp(k_i) \frac{1 - \exp(-\beta(T - t_i))}{\beta} \right) \\
&= \left(\sum_{i=1}^n \frac{A(i)}{\mu + \alpha A(i)} \right) - \left(\sum_{i=1}^n \exp(k_i) \frac{1 - \exp(-\beta(T - t_i))}{\beta} \right).
\end{aligned}$$

$$\begin{aligned}
\frac{\partial}{\partial \beta} \mathcal{L}(\mu, \alpha, \beta, \eta) &= \frac{\partial}{\partial \beta} \left(\left(\sum_{i=1}^n \log(\eta) + (-\eta(k_i - 2)) \right) + \left(\sum_{i=1}^n \log(\mu + \alpha A(i)) \right) \right. \\
&\quad \left. - \mu T - \left(\sum_{i=1}^n \alpha \exp(k_i) \frac{1 - \exp(-\beta(T - t_i))}{\beta} \right) \right) \\
&= \frac{\partial}{\partial \beta} \left(\sum_{i=1}^n \log(\eta) + (-\eta(k_i - 2)) \right) + \frac{\partial}{\partial \beta} \left(\sum_{i=1}^n \log(\mu + \alpha A(i)) \right) \\
&\quad - \frac{\partial}{\partial \beta} \mu T - \frac{\partial}{\partial \beta} \left(\sum_{i=1}^n \alpha \exp(k_i) \frac{1 - \exp(-\beta(T - t_i))}{\beta} \right) \\
&= \left(\sum_{i=1}^n \frac{\partial}{\partial \beta} \log(\mu + \alpha A(i)) \right) - \left(\sum_{i=1}^n \frac{\partial}{\partial \beta} \alpha \exp(k_i) \frac{1 - \exp(-\beta(T - t_i))}{\beta} \right) \\
&= \left(\sum_{i=1}^n \frac{1}{\mu + \alpha A(i)} \frac{\partial}{\partial \beta} (\mu + \alpha A(i)) \right) - \alpha \left(\sum_{i=1}^n \exp(k_i) \frac{\partial}{\partial \beta} \frac{1 - \exp(-\beta(T - t_i))}{\beta} \right) \\
&= \left(\sum_{i=1}^n \frac{1}{\mu + \alpha A(i)} (-\alpha B(i)) \right) \\
&\quad - \alpha \left(\sum_{i=1}^n \exp(k_i) \left(\frac{(T - t_i) \exp(-\beta(T - t_i)) \beta - 1 + \exp(-\beta(T - t_i))}{\beta^2} \right) \right) \\
&= \left(\sum_{i=1}^n \frac{-\alpha B(i)}{\mu + \alpha A(i)} \right) \\
&\quad - \alpha \left(\sum_{i=1}^n \exp(k_i) \left(\frac{(T - t_i) \exp(-\beta(T - t_i))}{\beta} + \frac{\exp(-\beta(T - t_i)) - 1}{\beta^2} \right) \right),
\end{aligned}$$

hvor

$$-B(i) = \frac{\partial}{\partial \beta} A(i) = \begin{cases} 0, & \text{for } i = 1 \\ \left(\sum_{t_j < t_i} \exp(k_j) (t_i - t_j) \exp(-\beta(t_i - t_j)) \right) & \text{for } i \geq 2 \end{cases}$$

$$\begin{aligned}
\frac{\partial}{\partial \eta} \mathcal{L}(\mu, \alpha, \beta, \eta) &= \frac{\partial}{\partial \eta} \left(\left(\sum_{i=1}^n \log(\eta) + (-\eta(k_i - 2)) \right) + \left(\sum_{i=1}^n \log(\mu + \alpha A(i)) \right) \right. \\
&\quad \left. - \mu T - \left(\sum_{i=1}^n \alpha \exp(k_i) \frac{1 - \exp(-\beta(T - t_i))}{\beta} \right) \right) \\
&= \frac{\partial}{\partial \eta} \left(\sum_{i=1}^n \log(\eta) + (-\eta(k_i - 2)) \right) + \frac{\partial}{\partial \eta} \left(\sum_{i=1}^n \log(\mu + \alpha A(i)) \right) \\
&\quad - \frac{\partial}{\partial \eta} \mu T - \frac{\partial}{\partial \eta} \left(\sum_{i=1}^n \alpha \exp(k_i) \frac{1 - \exp(-\beta(T - t_i))}{\beta} \right) \\
&= \left(\sum_{i=1}^n \frac{\partial}{\partial \eta} \log(\eta) + \frac{\partial}{\partial \eta} (-\eta(k_i - 2)) \right) \\
&= \left(\sum_{i=1}^n \frac{1}{\eta} + (2 - k_i) \right).
\end{aligned}$$

$$\begin{aligned}
\frac{\partial^2}{\partial \mu^2} \mathcal{L}(\mu, \alpha, \beta, \eta) &= \frac{\partial}{\partial \mu} \frac{\partial}{\partial \mu} \mathcal{L}(\mu, \alpha, \beta, \eta) \\
&= \frac{\partial}{\partial \mu} \left(\left(\sum_{i=1}^n \frac{1}{\mu + \alpha A(i)} \right) - T \right) \\
&= \frac{\partial}{\partial \mu} \left(\sum_{i=1}^n \frac{1}{\mu + \alpha A(i)} \right) - \frac{\partial}{\partial \mu} T \\
&= \left(\sum_{i=1}^n \frac{\partial}{\partial \mu} \frac{1}{\mu + \alpha A(i)} \right) \\
&= \left(\sum_{i=1}^n \frac{-1}{(\mu + \alpha A(i))^2} \right).
\end{aligned}$$

$$\begin{aligned}
\frac{\partial^2}{\partial \alpha^2} \mathcal{L}(\mu, \alpha, \beta, \eta) &= \frac{\partial}{\partial \alpha} \frac{\partial}{\partial \alpha} \mathcal{L}(\mu, \alpha, \beta, \eta) \\
&= \frac{\partial}{\partial \alpha} \left(\left(\sum_{i=1}^n \frac{A(i)}{\mu + \alpha A(i)} \right) - \left(\sum_{i=1}^n \exp(k_i) \frac{1 - \exp(-\beta(T - t_i))}{\beta} \right) \right) \\
&= \frac{\partial}{\partial \alpha} \left(\sum_{i=1}^n \frac{A(i)}{\mu + \alpha A(i)} \right) - \frac{\partial}{\partial \alpha} \left(\sum_{i=1}^n \exp(k_i) \frac{1 - \exp(-\beta(T - t_i))}{\beta} \right) \\
&= \left(\sum_{i=1}^n \frac{\partial}{\partial \alpha} \frac{A(i)}{\mu + \alpha A(i)} \right) \\
&= \left(\sum_{i=1}^n \frac{-A(i)^2}{(\mu + \alpha A(i))^2} \right).
\end{aligned}$$

$$\begin{aligned}
\frac{\partial^2}{\partial \beta^2} \mathcal{L}(\mu, \alpha, \beta, \eta) &= \frac{\partial}{\partial \beta} \frac{\partial}{\partial \beta} \mathcal{L}(\mu, \alpha, \beta, \eta) \\
&= \frac{\partial}{\partial \beta} \left(\left(\sum_{i=1}^n \frac{-\alpha B(i)}{\mu + \alpha A(i)} \right) \right. \\
&\quad \left. - \alpha \left(\sum_{i=1}^n \exp(k_i) \left(\frac{(T - t_i) \exp(-\beta(T - t_i))}{\beta} + \frac{\exp(-\beta(T - t_i)) - 1}{\beta^2} \right) \right) \right) \\
&= \frac{\partial}{\partial \beta} \left(\sum_{i=1}^n \frac{-\alpha B(i)}{\mu + \alpha A(i)} \right) \\
&\quad - \frac{\partial}{\partial \beta} \alpha \left(\sum_{i=1}^n \exp(k_i) \left(\frac{(T - t_i) \exp(-\beta(T - t_i))}{\beta} + \frac{\exp(-\beta(T - t_i)) - 1}{\beta^2} \right) \right) \\
&= \left(\sum_{i=1}^n \frac{\partial}{\partial \beta} \frac{-\alpha B(i)}{\mu + \alpha A(i)} \right) \\
&\quad - \alpha \left(\sum_{i=1}^n \exp(k_i) \frac{\partial}{\partial \beta} \left(\frac{(T - t_i) \exp(-\beta(T - t_i))}{\beta} + \frac{\exp(-\beta(T - t_i)) - 1}{\beta^2} \right) \right) \\
&= \left(\sum_{i=1}^n \frac{-\frac{\partial}{\partial \beta} (\alpha B(i)) \cdot (\mu + \alpha A(i)) + B(i) \cdot \alpha^2 \frac{\partial}{\partial \beta} A(i)}{(\mu + \alpha A(i))^2} \right) \\
&\quad - \alpha \left(\sum_{i=1}^n \exp(k_i) \left(\frac{\partial}{\partial \beta} \frac{(T - t_i) \exp(-\beta(T - t_i))}{\beta} + \frac{\partial}{\partial \beta} \frac{\exp(-\beta(T - t_i)) - 1}{\beta^2} \right) \right) \\
&= \left(\sum_{i=1}^n \frac{\alpha C(i) \cdot (\mu + \alpha A(i)) - \alpha^2 B(i)^2}{(\mu + \alpha A(i))^2} \right) \\
&\quad - \alpha \left(\sum_{i=1}^n \exp(k_i) \left(\frac{-(T - t_i)^2 \exp(-\beta(T - t_i)) \beta - (T - t_i) \exp(-\beta(T - t_i))}{\beta^2} \right. \right. \\
&\quad \left. \left. + \left(\frac{-(T - t_i) \exp(-\beta(T - t_i)) \beta^2 - (\exp(-\beta(T - t_i)) - 1) 2\beta}{\beta^4} \right) \right) \right) \\
&= \left(\sum_{i=1}^n \frac{\alpha C(i)}{\mu + \alpha A(i)} - \frac{\alpha^2 B(i)^2}{(\mu + \alpha A(i))^2} \right) \\
&\quad - \alpha \left(\sum_{i=1}^n \exp(k_i) \left(\frac{-(T - t_i)^2 \exp(-\beta(T - t_i))}{\beta} - \frac{(T - t_i) \exp(-\beta(T - t_i))}{\beta^2} \right. \right. \\
&\quad \left. \left. + \left(\frac{-(T - t_i) \exp(-\beta(T - t_i))}{\beta^2} - \frac{(\exp(-\beta(T - t_i)) - 1) 2}{\beta^3} \right) \right) \right) \\
&= \left(\sum_{i=1}^n \frac{\alpha C(i)}{\mu + \alpha A(i)} - \frac{\alpha^2 B(i)^2}{(\mu + \alpha A(i))^2} \right) \\
&\quad + \alpha \left(\sum_{i=1}^n \exp(k_i) \left(\frac{(T - t_i)^2 \exp(-\beta(T - t_i))}{\beta} + \frac{2(T - t_i) \exp(-\beta(T - t_i))}{\beta^2} \right. \right. \\
&\quad \left. \left. + \frac{2(\exp(-\beta(T - t_i)) - 1)}{\beta^3} \right) \right),
\end{aligned}$$

$$-C(i) = \frac{\partial}{\partial \beta} B(i) = \begin{cases} 0, & \text{for } i = 1 \\ \left(\sum_{t_j < t_i} \exp(k_j) (t_i - t_j)^2 \exp(-\beta(t_i - t_j)) \right) & \text{for } i \geq 2 \end{cases}$$

$$\begin{aligned} \frac{\partial^2}{\partial \eta^2} \mathcal{L}(\mu, \alpha, \beta, \eta) &= \frac{\partial}{\partial \eta} \frac{\partial}{\partial \eta} \mathcal{L}(\mu, \alpha, \beta, \eta) \\ &= \frac{\partial}{\partial \eta} \left(\sum_{i=1}^n \frac{1}{\eta} + (2 - k_i) \right) \\ &= \left(\sum_{i=1}^n \frac{\partial}{\partial \eta} \frac{1}{\eta} + \frac{\partial}{\partial \eta} (2 - k_i) \right) \\ &= \sum_{i=1}^n \frac{-1}{\eta^2}. \end{aligned}$$

$$\begin{aligned} \frac{\partial^2}{\partial^2 \mu \alpha} \mathcal{L}(\mu, \alpha, \beta, \eta) &= \frac{\partial}{\partial \mu} \frac{\partial}{\partial \alpha} \mathcal{L}(\mu, \alpha, \beta, \eta) \\ &= \frac{\partial}{\partial \mu} \left(\left(\sum_{i=1}^n \frac{A(i)}{\mu + \alpha A(i)} \right) - \left(\sum_{i=1}^n \exp(k_i) \frac{1 - \exp(-\beta(T - t_i))}{\beta} \right) \right) \\ &= \frac{\partial}{\partial \mu} \left(\sum_{i=1}^n \frac{A(i)}{\mu + \alpha A(i)} \right) - \frac{\partial}{\partial \mu} \left(\sum_{i=1}^n \exp(k_i) \frac{1 - \exp(-\beta(T - t_i))}{\beta} \right) \\ &= \left(\sum_{i=1}^n \frac{\partial}{\partial \mu} \frac{A(i)}{\mu + \alpha A(i)} \right) \\ &= \sum_{i=1}^n \frac{-A(i)}{(\mu + \alpha A(i))^2}. \end{aligned}$$

$$\begin{aligned}\frac{\partial^2}{\partial^2 \mu \beta} \mathcal{L}(\mu, \alpha, \beta, \eta) &= \frac{\partial}{\partial \beta} \frac{\partial}{\partial \mu} \mathcal{L}(\mu, \alpha, \beta, \eta) \\&= \frac{\partial}{\partial \beta} \left(\left(\sum_{i=1}^n \frac{1}{\mu + \alpha A(i)} \right) - T \right) \\&= \frac{\partial}{\partial \beta} \left(\sum_{i=1}^n \frac{1}{\mu + \alpha A(i)} \right) - \frac{\partial}{\partial \beta} T \\&= \left(\sum_{i=1}^n \frac{\partial}{\partial \beta} \frac{1}{\mu + \alpha A(i)} \right) \\&= \left(\sum_{i=1}^n \frac{-\frac{\partial}{\partial \beta} (\mu + \alpha A(i))}{(\mu + \alpha A(i))^2} \right) \\&= \left(\sum_{i=1}^n \frac{-\left(\frac{\partial}{\partial \beta} \mu + \frac{\partial}{\partial \beta} \alpha A(i) \right)}{(\mu + \alpha A(i))^2} \right) \\&= \sum_{i=1}^n \frac{-\alpha \frac{\partial}{\partial \beta} A(i)}{(\mu + \alpha A(i))^2} \\&= \sum_{i=1}^n \frac{\alpha B(i)}{(\mu + \alpha A(i))^2}.\end{aligned}$$

$$\begin{aligned}\frac{\partial^2}{\partial^2 \mu \eta} \mathcal{L}(\mu, \alpha, \beta, \eta) &= \frac{\partial}{\partial \mu} \frac{\partial}{\partial \eta} \mathcal{L}(\mu, \alpha, \beta, \eta) \\&= \frac{\partial}{\partial \mu} \left(\sum_{i=1}^n \frac{1}{\eta} + (2 - k_i) \right) \\&= 0.\end{aligned}$$

$$\begin{aligned}
\frac{\partial^2}{\partial^2 \alpha \beta} \mathcal{L}(\mu, \alpha, \beta, \eta) &= \frac{\partial}{\partial \beta} \frac{\partial}{\partial \alpha} \mathcal{L}(\mu, \alpha, \beta, \eta) \\
&= \frac{\partial}{\partial \beta} \left(\left(\sum_{i=1}^n \frac{A(i)}{\mu + \alpha A(i)} \right) - \left(\sum_{i=1}^n \exp(k_i) \frac{1 - \exp(-\beta(T - t_i))}{\beta} \right) \right) \\
&= \frac{\partial}{\partial \beta} \left(\sum_{i=1}^n \frac{A(i)}{\mu + \alpha A(i)} \right) - \frac{\partial}{\partial \beta} \left(\sum_{i=1}^n \exp(k_i) \frac{1 - \exp(-\beta(T - t_i))}{\beta} \right) \\
&= \left(\sum_{i=1}^n \frac{\partial}{\partial \beta} \frac{A(i)}{\mu + \alpha A(i)} \right) - \left(\sum_{i=1}^n \exp(k_i) \frac{\partial}{\partial \beta} \frac{1 - \exp(-\beta(T - t_i))}{\beta} \right) \\
&= \left(\sum_{i=1}^n \frac{\frac{\partial}{\partial \beta} A(i) \cdot (\mu + \alpha A(i)) - A(i) \frac{\partial}{\partial \beta} (\mu + \alpha A(i))}{(\mu + \alpha A(i))^2} \right) \\
&\quad - \left(\sum_{i=1}^n \exp(k_i) \frac{(T - t_i) \exp(-\beta(T - t_i)) \beta + (\exp(-\beta(T - t_i)) - 1)}{\beta^2} \right) \\
&= \left(\sum_{i=1}^n \frac{-B(i)(\mu + \alpha A(i)) - A(i) \frac{\partial}{\partial \beta} \mu + \frac{\partial}{\partial \beta} \alpha A(i)}{(\mu + \alpha A(i))^2} \right) \\
&\quad - \left(\sum_{i=1}^n \exp(k_i) \frac{(T - t_i) \exp(-\beta(T - t_i)) \beta + (\exp(-\beta(T - t_i)) - 1)}{\beta^2} \right) \\
&= \left(\sum_{i=1}^n \frac{-B(i)}{(\mu + \alpha A(i))} + \frac{A(i) \alpha B(i)}{(\mu + \alpha A(i))^2} \right) \\
&\quad - \left(\sum_{i=1}^n \exp(k_i) \left(\frac{(T - t_i) \exp(-\beta(T - t_i))}{\beta} + \frac{(\exp(-\beta(T - t_i)) - 1)}{\beta^2} \right) \right).
\end{aligned}$$

$$\begin{aligned}
\frac{\partial^2}{\partial^2 \alpha \eta} \mathcal{L}(\mu, \alpha, \beta, \eta) &= \frac{\partial}{\partial \alpha} \frac{\partial}{\partial \eta} \mathcal{L}(\mu, \alpha, \beta, \eta) \\
&= \frac{\partial}{\partial \alpha} \left(\sum_{i=1}^n \frac{1}{\eta} + (2 - k_i) \right) \\
&= 0.
\end{aligned}$$

$$\begin{aligned}
\frac{\partial^2}{\partial^2 \beta \eta} \mathcal{L}(\mu, \alpha, \beta, \eta) &= \frac{\partial}{\partial \beta} \frac{\partial}{\partial \eta} \mathcal{L}(\mu, \alpha, \beta, \eta) \\
&= \frac{\partial}{\partial \beta} \left(\sum_{i=1}^n \frac{1}{\eta} + (2 - k_i) \right) \\
&= 0.
\end{aligned}$$

11.6 R-kode

Bemærk, at hvis det ønskes at implementere noget af R-koden anbefales det at køre dem kronologisk. Da det er nødvendigt at anvende pakken *spatstat*, følger et R-script der indlæser denne.

```
1 library(spatstat)
```

Listing 11.1: Indlæser pakke i R.

Ydermere præsenteres et script der indlæser datasættet.

```
1 Data<-read.table("C:/Users/Rune/Desktop/P10 Speciale/Data/Jordskælvsdata.
txt", sep = ",", fill=T, header=T)
```

Listing 11.2: Indlæser datasættet. Hvis det ønskes at køre scriptet skal stien for datasættets placering angives.

11.6.1 Datapræsentation

I dette afsnit præsenteres de R-scripts der anvendes i afsnit 4.

```
1 plot(c(0,0), main="Rumlige placering af jordskælvshændelserne", ylim=c
(53,59), xlim=c(0.40,18), xlab="Breddegrader", ylab="Længdegradder",
pch=16)
points(Data[,8],Data[,7])
```

Listing 11.3: Plotter den rumlige placering af jordskælvshændelserne i datasættet.

```
Nmin<-Data[1,7];Nmax<-Data[1,7]
2 Emin<-Data[1,8];Emax<-Data[1,8]
Rmin<-Data[1,10];Rmax<-Data[1,10]
4 for(i in 2:1187){
ifelse(Nmin>=Data[i,7],Nmin<-Data[i,7],NA)
6 ifelse(Nmax<=Data[i,7],Nmax<-Data[i,7],NA)
ifelse(Emin>=Data[i,8],Emin<-Data[i,8],NA)
8 ifelse(Emax<=Data[i,8],Emax<-Data[i,8],NA)
ifelse(Rmin>=Data[i,10],Rmin<-Data[i,10],NA)
10 ifelse(Rmax<=Data[i,10],Rmax<-Data[i,10],NA)
}
12 Nmin;Nmax;Emin;Emax;Rmin;Rmax

14 minår<-c(0);maxår<-c(0)
for(i in 2:1187){
16 ifelse(Data[i,10]==Rmin,minår<-Data[i,1],NA)
ifelse(Data[i,10]==Rmax,maxår<-Data[i,1],NA)
18 }
minår;maxår
```

Listing 11.4: Finder frem til den kraftigste og svageste hændelse, hvilket år de er fremkommet, samt ekstremerne mht. breddegrader og længdegradder.

```
1 plot(Data[,1],Data[,10], ylab="Richter",xlab="År")
```

Listing 11.5: Plotter jordskælvshændelserne i forhold til det år hændelsen forekommer og styrken.

```

1 #Fjerner alle hændelser før 2006
  k<-0
3 for(i in 1:1187){
  ifelse (Data[i-k,1] < 2006, Data<-Data[-(i-k)], NA )
5 ifelse (Data[i-k,1] < 2006, k<-k+1, NA )
  }
7
  #Fjerner alle hændelser med en værdi <2 på richterskalaen
9 k<-0
  for(i in 1:length(Data[,1])){
11 ifelse (Data[(i-k),10] < 2, Data<-Data[-(i-k)], NA )
    ifelse (Data[(i-k),10] < 2, k<-k+1, NA )
13 }

```

Listing 11.6: Udvælger jordskælvshændelserne i perioden fra 2006-2009 med en styrke på 2 eller mere på Richter-skalaen.

```

1 #Sætter 2006=t(0) og udregner hvor mange dage der er gået siden årsskifte
  for(i in 1:length(Data[,1])){
3 ifelse (Data[i,2] == 1, Data[i,11]<-Data[i,3]+(Data[i,4]+(Data[i,5]+Data[i,6]/60)/60)/24, NA)
    ifelse (Data[i,2] == 2, Data[i,11]<-31+Data[i,3]+(Data[i,4]+(Data[i,5]+Data[i,6]/60)/60)/24, NA)
5 ifelse (Data[i,2] == 3, Data[i,11]<-59+Data[i,3]+(Data[i,4]+(Data[i,5]+Data[i,6]/60)/60)/24, NA)
    ifelse (Data[i,2] == 4, Data[i,11]<-90+Data[i,3]+(Data[i,4]+(Data[i,5]+Data[i,6]/60)/60)/24, NA)
7 ifelse (Data[i,2] == 5, Data[i,11]<-120+Data[i,3]+(Data[i,4]+(Data[i,5]+Data[i,6]/60)/60)/24, NA)
    ifelse (Data[i,2] == 6, Data[i,11]<-151+Data[i,3]+(Data[i,4]+(Data[i,5]+Data[i,6]/60)/60)/24, NA)
9 ifelse (Data[i,2] == 7, Data[i,11]<-181+Data[i,3]+(Data[i,4]+(Data[i,5]+Data[i,6]/60)/60)/24, NA)
    ifelse (Data[i,2] == 8, Data[i,11]<-212+Data[i,3]+(Data[i,4]+(Data[i,5]+Data[i,6]/60)/60)/24, NA)
11 ifelse (Data[i,2] == 9, Data[i,11]<-243+Data[i,3]+(Data[i,4]+(Data[i,5]+Data[i,6]/60)/60)/24, NA)
    ifelse (Data[i,2] == 10, Data[i,11]<-273+Data[i,3]+(Data[i,4]+(Data[i,5]+Data[i,6]/60)/60)/24, NA)
13 ifelse (Data[i,2] == 11, Data[i,11]<-304+Data[i,3]+(Data[i,4]+(Data[i,5]+Data[i,6]/60)/60)/24, NA)
    ifelse (Data[i,2] == 12, Data[i,11]<-334+Data[i,3]+(Data[i,4]+(Data[i,5]+Data[i,6]/60)/60)/24, NA)
15 }

17 #Udregner antal dage siden t(0)
  for(j in 2006:2009){
19 for(i in 1:length(Data[,1])){
    ifelse (Data[i,1] == j, Data[i,11]<-Data[i,11]+365 $\Lambda$ (j-2006), NA)
21 }}

23 #Tillægger skudår
  for(i in 1:length(Data[,1])){
25 ifelse (Data[i,1] >= 2008, Data[i,11]<-Data[i,11]+1, NA)
    }
27

```

```

#udregner i total år altså 365,24219 dage pr. år
29 for(i in 1:length(Data[,1])){
  Data[i,12]<-Data[i,11]/365.24219
31 }

```

Listing 11.7: Tilføjer to søjler til datasættet. En med antal dage siden starten af 2006 og en med antal år.

```

1 #Først konstrueres et tællemål
  Mål=function(x,y=Data[,12]){
2   i<-1
  while(x>y[i]){
3    while(x>=4,i<-length(y)+1,i<-i+1)
4    ifelse(x>=4,break,NA)
5  }
  i-1
6 }

11 #Tæller antallet af hændelser pr. halvår.
  ant<-c(0)
13 for(i in 1:8){
  ant[i]<-Mål(i/2)-Mål(i/2-0.5)
15 }
  ant

```

Listing 11.8: Kontrol af om intensiteten ser ud til at være uafhængig af tiden.

```

#Først findes intervallængderne
2 int<-c(0)
  int[1]<-Data[1,12]
4 for(i in 2:(length(Data[,1])-1)){
  int[i]<-Data[i+1,12]-Data[i,12]
6 }

8 #Udrenger mu
  mu<-354/4

10 #Plotning af histogram
12 hist(int, cex.axis=.8,col="grey",freq=FALSE,ylab="Antal",xlab="
  Intervalstørrelse",main="",xlim=c(0,0.1),ylim=c(0,100),breaks=15)
  curve(dexp(x,mu),col="Red",add=TRUE)
14 title(main="Histogram for intervalstørrelsen",font.main=3)

```

Listing 11.9: Plotter histogram for intervallængderne mellem hændelserne for de udvalgte data.

11.6.2 Den umærkede Hawkes model

I dette afsnit præsenteres de R-scripts der anvendes i afsnit 5.

```

Hawkes=function(mu, alpha , beta , g){
2 S<-c(0); S[1]<-1

4 U<-runif(1,min=0,max=1)
  t<-c(0)
6 t[1]<-rexp(1,mu)

8 i<-1
  while(t[i]<g){
10 i<-i+1
    U<-runif(1,min=0,max=1)
12 u<-t[i-1]+rexp(1,mu)

14 for(j in 1:200){
    f<-log(U)+mu*Lambda(u-t[i-1])+alpha/beta*LambdaS[i-1]*Lambda(1-exp(-beta*Lambda(u-t[i-1])))
16 df<-mu+alpha*LambdaS[i-1]*Lambdaexp(-beta*Lambda(u-t[i-1]))
    u<-u-f/df
18 }
    t[i]<-u
20 S[i]<-exp(-beta*Lambda(t[i]-t[i-1]))*LambdaS[i-1]+1
  }
22 t<-t[-(length(t))]
  t
24 }

```

Listing 11.10: Script for algoritme 3.

```

NR=function(n,mu0,alpha0,beta0,x){
2 f<-c(0); vec<-c(mu0,alpha0,beta0); T=x[length(x)]

4 for(k in 1:n){
  A<-c(0); B<-c(0); C<-c(0)
6 A[1]<-0; B[1]<-0; C[1]<-0
  for(i in 2:length(x)){
8 A[i]<-sum(exp(-vec[3]*Lambda(x[i]-x[1:(i-1)])))
  B[i]<-sum((x[i]-x[1:(i-1)])*Lambdaexp(-vec[3]*Lambda(x[i]-x[1:(i-1)])))
10 C[i]<-sum((x[i]-x[1:(i-1)])^2*Lambdaexp(-vec[3]*Lambda(x[i]-x[1:(i-1)])))
  }
12
  #Konstruktion af de først ordens afledte
14 f[1]<-(-1)*T+sum(1/(vec[1]+vec[2]*Lambda))
  f[2]<-sum(A/(vec[1]+vec[2]*Lambda))+sum(1/vec[3]*Lambda(exp(-vec[3]*Lambda(T-x))-1))
16 f[3]<-(-1)*Lambda*sum((T-x)*Lambdaexp(-vec[3]*Lambda(T-x))/vec[3]+(exp(-vec[3]*Lambda(T-x))-1)/vec[3]^2)+(-1)*Lambda*sum(B/(vec[1]+vec[2]*Lambda))

18 df1dx<-(-1)*Lambda*sum(1/(vec[1]+vec[2]*Lambda)^2)
  df1dy<-(-1)*Lambda*sum(A/(vec[1]+vec[2]*Lambda)^2)
20 df1dz<-sum(vec[2]*Lambda*B/(vec[1]+vec[2]*Lambda)^2)
  df2dx<-df1dy
22 df2dy<-(-1)*Lambda*sum((A/(vec[1]+vec[2]*Lambda))^2)
  df2dz<-sum(1/vec[3]*Lambda(T-x)*Lambdaexp(-vec[3]*Lambda(T-x))+1/vec[3]^2*Lambda(exp(-vec[3]*Lambda(T-x))-1))+sum(-B/(vec[1]+vec[2]*Lambda)+vec[2]*Lambda*Lambda*B/(vec[1]+vec[2]*Lambda)^2)

```

```

24 df3dx<-df1dz
   df3dy<-df2dz
26 df3dz<-vec[2]*sum(1/vec[3]*Lambda(T-x)^2*Lambdaexp(-vec[3]*Lambda(T-x))+2/vec[3]^2*Lambda(T-x)*
   Lambdaexp(-vec[3]*Lambda(T-x))+2/vec[3]^3*Lambda(exp(-vec[3]*Lambda(T-x))-1))+sum(vec[2]*LambdaC/(
   vec[1]+vec[2]*Lambda)-(vec[2]*LambdaB/(vec[1]+vec[2]*Lambda))^2)

28 J<-matrix(data=NA,nrow=3,ncol=3)
   J[1,1]<-df1dx;J[1,2]<-df1dy;J[1,3]<-df1dz
30 J[2,1]<-df2dx;J[2,2]<-df2dy;J[2,3]<-df2dz
   J[3,1]<-df3dx;J[3,2]<-df3dy;J[3,3]<-df3dz
32
   vec<-vec-solve(J)%A%f
34 }
   vec
36 }

```

Listing 11.11: Script for den tre-dimensionel Newton-Rhapson for modellen.

```

#Konstruktion af data
2 t<-c(0)
   t<-Hawkes(mu=0.5,alpha=5,beta=10,g=100)
4 length(t)

6 #Newton-Rhapson approksimation af parametrene.
   NR(n=1,mu0=0.25,alpha0=4,beta0=5,x=t)
8 NR(n=5,mu0=0.25,alpha0=4,beta0=5,x=t)
   NR(n=10,mu0=0.25,alpha0=4,beta0=5,x=t)
10 NR(n=20,mu0=0.25,alpha0=4,beta0=5,x=t)

```

Listing 11.12: Konstruktion af kunstigt datasæt samt forskellige Newton-Rhapson approksimation af de valgte parametre.

```

plot(c(0,0), type="n", xlab="Tidslinie", ylab="Hændelser",xlim=range
   (0,100),ylim=range(0,2), yaxt="n")
2 for(i in 1:length(t)){
   points(t[i],1)}
4 title(main = "Konstruerede data", font.main = 3)

```

Listing 11.13: Plotning af konstrueret data.

```

NR(n=1,mu0=354/4,alpha0=200,beta0=800,x=Data[,12])
2 NR(n=5,mu0=354/4,alpha0=200,beta0=800,x=Data[,12])
   NR(n=10,mu0=354/4,alpha0=200,beta0=800,x=Data[,12])
4 NR(n=20,mu0=354/4,alpha0=200,beta0=800,x=Data[,12])

```

Listing 11.14: Parameterestimation for modellen.

```

vec<-NR(n=20,mu0=354/4,alpha0=200,beta0=800,x=Data[,12])
2 x=Data[,12]
   A<-c(0);B<-c(0);C<-c(0)
4 A[1]<-0;B[1]<-0;C[1]<-0
   for(i in 2:length(Data[,12])){
6 A[i]<-sum(exp(-vec[3]*Lambda(x[i]-x[1:(i-1)])))
   B[i]<-sum((x[i]-x[1:(i-1)])*Lambdaexp(-vec[3]*Lambda(x[i]-x[1:(i-1)])))
8 C[i]<-sum((x[i]-x[1:(i-1)])^2*Lambdaexp(-vec[3]*Lambda(x[i]-x[1:(i-1)])))
   }
10

```

```

T=Data[ length (Data[,12]),12]
12 f<-c(0); T<-4
   f[1]<-(-1)ΔT+sum(1/(vec[1]+vec[2]ΔA))
14 f[2]<-sum(A/(vec[1]+vec[2]ΔA))+sum(1/vec[3]Δ(exp(-vec[3]Δ(T-x))-1))
   f[3]<-(-1)Δvec[2]Δsum((T-x)Δexp(-vec[3]Δ(T-x))/vec[3]+(exp(-vec[3]Δ(T-x))-1)/vec[3]^2)+(-1)Δvec[2]Δsum(B/(vec[1]+vec[2]ΔA))
16 f

```

Listing 11.15: Kontrol af om de approksimerede parametre giver nul i de partielt afledte.

```

t<-Data[,12]
2
#Konstruktion af s.
4 s<-c(0)
   for(i in 1:354){
6 s[i]<-vec[1]Δt[i]+vec[2]/vec[3]Δsum(1-exp(-vec[3]Δ(t[i]-t[1:(i-1)])))
   }
8
#Finder intervaller mellem hændelserne.
10 int<-c(0)
   int[1]<-s[1]
12 for(i in 1:(length(s)-1)){
   int[i]<-s[i+1]-s[i]
14 }

16 #Genererer kvantiler
   a<-seq(0,1,by=0.01)
18 INTkvantiler<- quantile(int, a)

20 #Tegner grafen
   plot(INTkvantiler, qexp(a,1), col="#0000FF70", main="", xlim=c(0,6), ylim=
       =c(0,6), xlab="Empiriske kvantiler", ylab="Teoretiske kvantiler", pch
       =16)
22 abline(0,1,col="#404040FF",lty=2,lwd=2)
   title(main = "q-q plot for intervalstørrelsen", font.main = 3)

```

Listing 11.16: Modelkontrol.

11.6.3 Den mærkede Hawkes model

I dette afsnit præsenteres de R-scripts der anvendes i afsnit 7.

```

1 MerkHawkes=function(mu, alpha, beta, eta, T){
  t<-c(0); k<-c(0)
3 i<-1
  t[i]<-rexp(1, mu); k[i]<-rexp(1, eta)+2
5 while(t[length(t)]<T){
  i<-i+1
7 t[i]<-rexp(1, mu)+t[i-1]
  k[i]<-rexp(1, eta)+2
9 }
  t<-t[-length(t)]
11 k<-k[-length(k)]

13 Matrix<-matrix(data=c(t, k), nrow=length(t), ncol=2)

15 gamma=function(x, y){ ifelse(x<=Matrix[i, 1], 0, exp(Matrix[i, 2])^alpha*exp(-
  beta*(x-Matrix[i, 1]))) }

17 i<-1
  while(Matrix[i, 1]<T){
19 tid<-c(0); z<-c(0)
  K<-rpoispp(gamma, lmax=exp(Matrix[i, 2])^alpha, win=owin(c(0, T), c(0, 1)))
21 ifelse(length(K$x)<1, NA, tid<-K$x)
  ifelse(length(K$x)<1, NA, z<-rexp(length(tid), eta)+2)
23 ifelse(length(K$x)<1, NA, Matrix<-matrix(data=c(Matrix[, 1], tid, Matrix[, 2], z),
  nrow=(length(Matrix[, 1])+length(tid)), ncol=2))
  ifelse(length(K$x)<1, NA, Matrix<-Matrix[order(Matrix[, 1]), ])
25 i<-i+1
  ifelse(i==length(Matrix[, 1]), break, NA)
27 }
  Matrix
29 }

```

Listing 11.17: Script for algoritme 5.

```

1 NR4=function(n, mu0, alpha0, beta0, eta0, x, m, T){
  f<-c(0); vec<-c(mu0, alpha0, beta0, eta0)
3
  for(k in 1:n){
5 A<-c(0); B<-c(0); C<-c(0)
  A[1]<-0; B[1]<-0; C[1]<-0
7 for(i in 2:length(x)){
  A[i]<-sum(exp(m[1:(i-1)])^alpha*exp(-vec[3]*Lambda(x[i]-x[1:(i-1)])))
9 B[i]<-sum(exp(m[1:(i-1)])^alpha*exp(-vec[3]*Lambda(x[i]-x[1:(i-1)]))
  )
  C[i]<-sum(exp(m[1:(i-1)])^alpha*exp(-vec[3]*Lambda(x[i]-x[1:(i-1)]))
  )
11 }

13 #Konstruktion af de først ordens afledte
  f[1]<-(-1)*T+sum(1/(vec[1]+vec[2]*Lambda))
15 f[2]<-sum(A/(vec[1]+vec[2]*Lambda))+sum(exp(m)/vec[3]*Lambda(exp(-vec[3]*Lambda(T-x))-1))

```

```

f[3]<-( -1)*vec[2]*sum( exp(m)*Lambda((T-x)*Lambdaexp(-vec[3]*Lambda(T-x)))/vec[3]+( exp(-vec
[3]*Lambda(T-x))-1)/vec[3]^2))+(-1)*vec[2]*sum(B/(vec[1]+vec[2]*Lambda))
17 f[4]<-length(m)/vec[4]+sum(2-m)

19 df1dx<-( -1)*sum(1/(vec[1]+vec[2]*Lambda)^2)
df1dy<-( -1)*sum(A/(vec[1]+vec[2]*Lambda)^2)
21 df1dz<-sum(vec[2]*Lambda/(vec[1]+vec[2]*Lambda)^2)
df1de<-0
23 df2dx<-df1dy
df2dy<-( -1)*sum((A/(vec[1]+vec[2]*Lambda))^2)
25 df2dz<-sum( exp(m)*Lambda(1/vec[3]*Lambda(T-x)*Lambdaexp(-vec[3]*Lambda(T-x))+1/vec[3]^2*Lambda(-
vec[3]*Lambda(T-x))-1))+sum(-B/(vec[1]+vec[2]*Lambda)+vec[2]*Lambda*B/(vec[1]+vec[2]*Lambda)^2)
df2de<-0
27 df3dx<-df1dz
df3dy<-df2dz
29 df3dz<-vec[2]*sum( exp(m)*Lambda(1/vec[3]*Lambda(T-x)^2*Lambdaexp(-vec[3]*Lambda(T-x))+2/vec[3]^2*Lambda
(T-x)*Lambdaexp(-vec[3]*Lambda(T-x))+2/vec[3]^3*Lambda( exp(-vec[3]*Lambda(T-x))-1))+sum( vec
[2]*Lambda/(vec[1]+vec[2]*Lambda)-(vec[2]*Lambda/(vec[1]+vec[2]*Lambda))^2)
df3de<-0
31 df4dx<-df1de
df4dy<-df2de
33 df4dz<-df3de
df4de<-( -1)*Lambda(length(m))/(vec[4]^2)

35 J<-matrix( data=NA, nrow=4, ncol=4)
37 J[1,1]<-df1dx; J[1,2]<-df1dy; J[1,3]<-df1dz; J[1,4]<-df1de
J[2,1]<-df2dx; J[2,2]<-df2dy; J[2,3]<-df2dz; J[2,4]<-df2de
39 J[3,1]<-df3dx; J[3,2]<-df3dy; J[3,3]<-df3dz; J[3,4]<-df3de
J[4,1]<-df4dx; J[4,2]<-df4dy; J[4,3]<-df4dz; J[4,4]<-df4de
41
vec<-vec-solve(J)%Lambda%f
43 }
vec
45 }

```

Listing 11.18: Script for den fire-dimensionel Newton-Rhapson for modellen.

```

1 #Trækker 2 fra richter tallet
for(i in 1:length(Data[,1])){
3 Data[i,13]<-Data[i,10]-2
}
5
#Plotter histogram
7 hist(Data[,13], cex.axis=.8,col="grey", freq=FALSE, ylab="", xlab="
Styrke", main="", xlim=c(0,2.5), ylim=c(0,5), breaks=30)
title(main="Histogram for styrken af hændelserne", font.main=3)

```

Listing 11.19: Konstruktion af histogram for styrken af de udvalgte jordskælvs hændelser.
Bemærk at der er trukket to fra styrkerne.

```

#Konstruktion af prøvedata
2 sim<-c(0)
sim<-MerkHawkes(0.5,4,100,4,100)
4 length(sim[,1])

```

```

6 #Plotning af test
  plot(c(0,0), type="n", xlab="Tidslinie", ylab="Styrke", xlim=range(0,100),
       ylim=range(0,5))
8 for(i in 1:length(sim[,1])){
  points(sim[i,1],sim[i,2])}
10 title(main = "Konstruerede data", font.main = 3)

12 #Parameterbestemmelse af konstrueret data
  NR4(n=1,mu0=0.25,alpha0=2,beta0=50,eta0=2,x=sim[,1],m=sim[,2],T=100)
14 NR4(n=5,mu0=0.25,alpha0=2,beta0=50,eta0=2,x=sim[,1],m=sim[,2],T=100)
  NR4(n=10,mu0=0.25,alpha0=2,beta0=50,eta0=2,x=sim[,1],m=sim[,2],T=100)
16 NR4(n=20,mu0=0.25,alpha0=2,beta0=50,eta0=2,x=sim[,1],m=sim[,2],T=100)

```

Listing 11.20: Konstruktion af kunstigt datasæt samt forskellige Newton-Rhapson approksimation af de valgte parametre.

```

  NR4(n=1,mu0=354/4,alpha0=25,beta0=2000,eta0=2,x=Data[,12],m=Data[,10],T
    =4)
2 NR4(n=5,mu0=354/4,alpha0=25,beta0=2000,eta0=2,x=Data[,12],m=Data[,10],T
    =4)
  NR4(n=10,mu0=354/4,alpha0=25,beta0=2000,eta0=2,x=Data[,12],m=Data[,10],T
    =4)
4 NR4(n=20,mu0=354/4,alpha0=25,beta0=2000,eta0=2,x=Data[,12],m=Data[,10],T
    =4)

```

Listing 11.21: Parameterbestemmelse for modellen.

```

  vec<-NR4(n=20,mu0=354/4,alpha0=25,beta0=2000,eta0=2,x=Data[,12],m=Data
    [,10],T=4)
2
  x<-Data[,12];m<-Data[,10]
4 A<-c(0);B<-c(0);C<-c(0)
  A[1]<-0;B[1]<-0;C[1]<-0
6 for(i in 2:length(x)){
  A[i]<-sum(exp(m[1:(i-1)])*exp(-vec[3]*Lambda(x[i]-x[1:(i-1)])))
8 B[i]<-sum(exp(m[1:(i-1)])*Lambda(x[i]-x[1:(i-1)])*exp(-vec[3]*Lambda(x[i]-x[1:(i-1)]))
    ))
  C[i]<-sum(exp(m[1:(i-1)])*Lambda(x[i]-x[1:(i-1)])^2*exp(-vec[3]*Lambda(x[i]-x[1:(i-1)]))
    ))
10 }

12 #Konstruktion af første ordens afledte
  f<-c(0);T<-4
14 f[1]<-(-1)*T+sum(1/(vec[1]+vec[2]*Lambda))
  f[2]<-sum(A/(vec[1]+vec[2]*Lambda))+sum(exp(m)/vec[3]*Lambda(exp(-vec[3]*Lambda(T-x))-1))
16 f[3]<-(-1)*vec[2]*sum(exp(m)*Lambda((T-x)*exp(-vec[3]*Lambda(T-x))/vec[3]+(exp(-vec
    [3]*Lambda(T-x))-1)/vec[3]^2))+(-1)*vec[2]*sum(B/(vec[1]+vec[2]*Lambda))
  f[4]<-length(m)/vec[4]+sum(2-m)
18 f

```

Listing 11.22: Kontrollerer om de estimerede parametre medfører at de første ordens afledte er nul.

```

  lambda<-3.01
2
  #plotter histogram med fordelingsfunktion

```

```

4 hist(Data[,13], cex.axis=.8,col="grey", freq = FALSE,ylab="", xlab="
  Styrke", main="", xlim=c(0,2.5), ylim=c(0,5),breaks = 20)
  curve(dexp(x,lambda), col="Red", add = TRUE)
6 title(main = "Histogram for styrken af hændelserne", font.main = 3)

8 a<-seq(0.001,0.999,by=0.001)
  RichterKvantiler <- quantile(Data[,13], a)
10 #q-q plot
12 plot(RichterKvantiler, qexp(a,lambda), col="#0000FF70", main="", xlim=c
  (0,3), ylim=c(0,3), xlab="Empiriske kvantiler", ylab="Teoretiske
  kvantiler", pch=16)
  abline(0,1,col="#404040FF",lty=2,lwd=2)
14 title(main = "q-q plot for styrken af hændelserne", font.main = 3)

```

Listing 11.23: Undersøger om fordelingen af mærkerne i modellen passer til virkeligheden.

```

t<-Data[,12]; m<-Data[,13]
2
  #Konstruktion af s
4 s<-c(0)
  for(i in 1:354){
6 s[i]<-vec[1]^t[i]+vec[2]/vec[3]^sum(exp(m[1:(i-1)])^(1-exp(-vec[3]^(t[i]-
  t[1:(i-1)]))))))
  }
8
  s<-sort(s)
10
  #Konstruerer intervallængder
12 int<-c(0)
  int[1]<-s[1]
14 for(i in 2:(length(s))){
  int[i]<-s[i]-s[i-1]
16 }

18 #Genererer kvantiler
  a<-seq(0,1,by=0.01)
20 INTkvantiler<- quantile(int, a)

22 #Tegner grafen
  plot(INTkvantiler, qexp(a,1), col="#0000FF70", main="", xlim=c(0,6), ylim
    =c(0,6), xlab="Empiriske kvantiler", ylab="Teoretiske kvantiler", pch
    =16)
24 abline(0,1,col="#404040FF",lty=2,lwd=2)
  title(main = "q-q plot for intervalstørrelsen", font.main = 3)

```

Listing 11.24: Undersøger om grundintensiteten i modellen passer til virkeligheden.