

Controle van het minimale aantal fMRI-scans door simulatieonderzoek



Ammar Khawaja
Ammar Khawaja

BWI-Werkstuk
BWI-Werkstuk

Controle van het minimale aantal fMRI-scans door simulatieonderzoek

Ammar Khawaja

BWI-Werkstuk

Begeleider: dr. Fetsje Bijma

vrije Universiteit
Faculteit der Exacte Wetenschappen
Studierichting Bedrijfskunde & Informatica

De Boelelaan 1081a
1081 HV Amsterdam

September 2008

Voorwoord

Het schrijven van een BWI-Werkstuk is een verplicht onderdeel van de studie Bedrijfswiskunde & Informatica (BWI). Het doel van het werkstuk is het onderzoeken van een probleem en dit op een deskundige manier te beschrijven. Het werkstuk moet zowel een bedrijfsgericht aspect als een wiskunde of informatica aspect bevatten.

In dit werkstuk heb ik een simulatieonderzoek verricht om te controleren of een formule voor het bepalen van het minimale aantal fMRI-scans werkt in de praktijk. Deze formule is afgeleid door een power analyse waarbij een aantal benaderingen worden gemaakt.

Ik wil Fetsje Bijma bedanken voor haar begeleiding tijdens het schrijven van het werkstuk en haar suggesties bij het verrichten van het simulatieonderzoek. Ik wens de lezer heel veel leesplezier.

Ammar Khawaja
September 2008

Samenvatting

Functionele MRI (fMRI) is een relatief nieuwe toepassing van de MRI. Deze techniek wordt gebruikt voor hersenactiviteitonderzoeken. Hoewel de fMRI techniek een aantal voordelen kent ten opzichte van andere bestaande technieken, kan een lange scanduur belastend zijn voor de patiënt. De reden hiervoor is dat de persoon in een magneetveld ligt en dit zorgt voor interne verhitting. Ook is het gebruik van het apparaat waarin de patiënt ligt duur. Om deze redenen is het zinvol om te bepalen wat de minimale scanduur moet zijn.

In een onlangs verschenen artikel is een formule afgeleid om het minimale aantal fMRI-scans te bepalen. Met deze formule kan bepaald worden wat het minimale aantal scans moet zijn om bij een gegeven kans een signaal te kunnen detecteren. Omdat er in de afleiding benaderingen en aannamen worden gemaakt, is het zinvol om te onderzoeken of deze formule in praktijk werkt. Om dit te controleren is een simulatieonderzoek verricht.

Het simulatieonderzoek is verricht met behulp van data van twee proefpersonen verkregen van de medische faculteit van de VU. Met deze data is een meervoudige lineair regressiemodel gebouwd. Vervolgens zijn een aantal parameters berekend om het signaal te simuleren. Hierbij hebben wij onderscheid gemaakt van twee typen ruis: Gaussische en fMRI. In de simulatie is het onderscheidend vermogen bepaald en vergeleken met het verwachte onderscheidend vermogen. Vervolgens is gekeken of er sprake is van een verschil.

Het gesimuleerde onderscheidend vermogen bleek lager te zijn dan verwacht voor Gaussische én ook voor fMRI-ruis. Echter werd in het artikel waarin de formule is afgeleid al opgemerkt dat de het aantal scans onderschat wordt. Na het uitvoeren van een correctie op het aantal scans, was het resultaat beter in beide gevallen (Gaussische en fMRI-ruis). In veel gevallen was het gesimuleerde onderscheidend vermogen groter of gelijk dan het verwachte onderscheidend vermogen. Verder merkten wij op dat het gesimuleerde onderscheidend vermogen in het geval van Gaussische ruis over het algemeen groter was dan in het geval van fMRI-ruis. Een mogelijke verklaring hiervoor zou kunnen zijn dat de normaliteitveronderstelling van de fMRI-ruis niet geheel correct is. Ook zijn er verschillen gevonden tussen de beide proefpersonen. De resultaten van het simulatieonderzoek was in veel gevallen beter voor proefpersoon 1 dan voor proefpersoon 2. Om dit verschil nader te onderzoeken is meer data gewenst. Eén van de belangrijkste bevindingen was dat er in het begin van een experiment meer ruis bleek te zijn dan op het eind. Als hier rekening mee wordt gehouden dan is het resultaat voor beiden typen ruis aanzienlijk beter dan voorheen.

Executive Summary

Functional MRI is a relatively new application of the existing MRI. It is widely used to do research in brain activity. Although there several advantages in using this technique compared to other existing ones, there are also disadvantages. A long scan duration can be a strain on the patient, because this person lies in a magnetic field which causes internal heating in the body. Furthermore the use of a scanner is expensive. This makes it useful to have a formula such that a minimal scan duration can be calculated.

In a recently published article a formula was derived to calculate the minimal scan duration. With this formula it is possible to calculate the minimal scan duration to have a certain chance of detecting a signal. Because this formula was derived by making some assumptions as well as approximations it is desirable to find out whether it works in practice. To achieve this a simulation study has been performed.

The simulation is performed using data from two subjects. These data were made available by the medical faculty of the Free University Amsterdam. Firstly we built a multiple linear regression model with the data. Then we used this model to calculate a few parameters to help us simulate the signal. We simulated two specific cases: using Gaussian noise and real fMRI noise. We calculated the power we achieved in the simulation and compared this with the expected power to see whether there were notable differences.

The simulated power was in both cases (Gaussian and fMRI noise) lower than what we expected. But in the article mentioned before it was noted that the minimal scan duration is undervalued. This could be the reason for a lower power. To correct this we performed a correction on the number of scans. The result was a higher power for both Gaussian and fMRI noise. In some cases it was equal or even greater than the expected power. We also noted that the simulated power with Gaussian noise was in most cases higher than with fMRI noise. A possible explanation would be that the assumption that fMRI noise is normally distributed is wrong. We also found differences in simulated power between subject 1 and 2. The results for subject 1 were better compared to subject 2. To study these differences between subjects more data is necessary. One of the main findings was the fact that there is more noise in the beginning of an experiment than in the end of one. If this is taken into account then the result substantially better than before.

Inhoudsopgave

Voorwoord.....	II
Samenvatting.....	III
Executive Summary.....	IV
1. Inleiding.....	1
2. Functionele MRI.....	2
2.1 Geschiedenis.....	2
2.2 fMRI.....	5
3. Meervoudig lineaire regressiemodel & power analyse.....	7
3.1 Regressieanalyse.....	7
3.2 Power analyse.....	10
3.3 Het minimale aantal scans in een fMRI experiment.....	11
4. Simulatieonderzoek.....	14
4.1 Onderzoeksopzet.....	14
4.2 Data.....	15
4.3 Dataverkenning.....	16
4.4 Simulatie.....	22
5. Conclusie & discussie.....	29
5.1 Conclusie.....	29
5.2 Discussie.....	30
Literatuurlijst.....	31
Bijlage 1: Code.....	33
Bijlage 2: Tabellen & figuren van simulatieonderzoek.....	37

1. Inleiding

Functionele Magnetic Resonance Imaging (fMRI) is een techniek die gebruikt wordt voor hersenactiviteitonderzoeken. Men wil bepalen of er sprake is van activiteit in bepaalde gebieden van de hersenen bij het verrichten van een specifieke taak (bijvoorbeeld het bewegen van de linkerwijsvinger gedurende een halve minuut).

De fMRI data kan gefit worden in een meervoudig lineaire regressiemodel. Met behulp van het model kan men bijvoorbeeld statistische toetsen en/of analyses uitvoeren.

Het bepalen van het aantal scans dat nodig is om een signaal te detecteren, kan worden bewerkstelligd door allerlei software. Daarin zijn meestal iteratieve methoden geïmplementeerd die het aantal scans bepalen. Dit kan erg tijdrovend zijn. Bovendien is het gebruik van dit soort speciale softwarepakketten soms niet wenselijk, omdat de gebruiker niet weet hoe het resultaat tot stand komt. In een onlangs verschenen artikel [27] is het minimale aantal scans bepaald wat nodig is om een signaal te detecteren. Maar deze is volgens Bijma et al. in [28] niet correct omdat het niet is afgeleid met behulp van een power analyse. Bovendien wordt door Murphy et. al. in [27] aangenomen dat het detecteren van een signaal deterministisch is en niet stochastisch. In [28] is het minimale aantal scans wel afgeleid met de juiste wiskundige onderbouwing.

Het gebruik van een apparaat als de MRI-scanner brengt kosten met zich mee én lange scanperioden zijn belastend voor de patiënt. Het is dus wenselijk om een methode voor het bepalen van het minimale aantal scans te beschikken die makkelijk toepasbaar is, maar ook wiskundig correct is.

Door middel van een simulatieonderzoek willen wij nagaan of deze formule correct is of niet. Bovendien zijn wij geïnteresseerd in het verschil dat eventueel ontstaat, doordat er een aantal aannamen en benaderingen worden gemaakt. De hoofdvraag van dit BWI-Werkstuk is als volgt geformuleerd:

Werkt de gestelde n voor het minimale aantal scans voor Gaussische en fMRI ruis?

Om de bovenstaande vraag te beantwoorden, zullen wij ook een aantal deelvragen trachten te beantwoorden, namelijk:

- Is het verschil in waargenomen onderscheidend vermogen en het onderscheidend vermogen van de toets groot of klein? Is het verschil systematisch of juist niet?
- Is het eventuele verschil te verklaren?

Dit verslag is opgedeeld in drie stukken: de resultaten van de literatuurstudie, een beschrijving van het simulatieonderzoek en ten slotte de resultaten en conclusies van het simulatieonderzoek.

Er is literatuurstudie gedaan in fMRI en de statistische toetsen en modellen die gebruikt zijn in het afleiden van het minimale aantal scans en in het simulatieonderzoek. In hoofdstuk 2 is een korte inleiding gegeven van de geschiedenis van de fMRI. Bovendien zijn ook de technieken en principes besproken die gebruikt worden in het maken van fMRI afbeelding. Hoofdstuk 3 bevat een beknopte beschrijving van de statistische toetsen en technieken die toegepast zijn in de analyse van het fMRI signaal. Ook is het bepalen van het minimale aantal scans geformuleerd. Hoofdstuk 4 bevat een uitgebreide beschrijving van de data, de analyse hiervan en de verschillende simulaties die uitgevoerd zijn. De laatste paragraaf van dit hoofdstuk bevat de belangrijkste bevindingen van dit onderzoek. Ten slotte bevat het vijfde en laatste hoofdstuk de conclusie en discussie van het simulatieonderzoek. Het simulatieonderzoek is uitgevoerd in het mathematische softwarepakket Matlab. De code van de simulaties zijn opgenomen in de bijlagen.

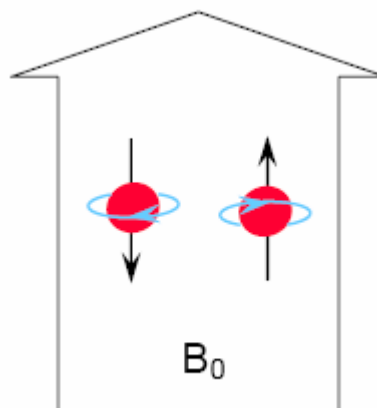
Wiskundige kennis is niet vereist om de hoofdstukken twee en vijf te kunnen begrijpen. Voor de overige hoofdstukken wordt algemene kennis in statistiek en lineaire algebra wel verondersteld.

2. Functionele MRI

2.1 Geschiedenis

De voorganger van de functionele MRI is de MRI. Deze afkorting staat voor: “Magnetic Resonance Imaging” oftewel kernspintomografie. MRI is een techniek die gebruikt maakt van kernspinresonantie. Dit principe is voor het eerst in 1946 beschreven door de wetenschappers Edward Mills Purcell en Felix Bloch ([1], [2]). Beiden hebben samen met hun eigen team in Harvard en Stanford onderzoek gedaan in het observeren van een magnetische resonantie in nucleï van gassen, vloeibare en vaste stoffen. Deze twee onderzoeken vonden in dezelfde periode plaats en waren onafhankelijk van elkaar. Het resultaat van deze onderzoeken leidde tot wat nu bekend staat als kernspinresonantie. Voor deze uitvinding ontvingen beide heren in 1952 de Nobelprijs voor Natuurkunde [3].

Kernspinresonantie, wat in het Engels “Nuclear Magnetic Resonance” (NMR) wordt genoemd, is een natuurkundig verschijnsel. Een atoomkern is een geladen deeltje en zal een magnetisch veld genereren als deze in beweging is. Echter gebeurt dit alleen als de kern de eigenschap spin heeft. Een sterker extern magnetisch veld kan ervoor zorgen dat de kern in een andere richting zal draaien (spinnen). Dit kan parallel of antiparallel zijn. Een voorbeeld van parallelle en een antiparallelle kern onder de invloed van een magneetveld B_0 is te zien in de onderstaande figuur 2.1:



Figuur 2.1: Antiparallel (links) en parallel (rechts) kernen in een magneetveld B_0 [5].

De kern keert terug naar zijn voorkeursrichting als het niet meer blootgesteld is aan het externe magnetisch veld. Hierdoor worden er radiogolven geproduceerd. Door middel van deze signalen is het mogelijk om een MRI afbeelding te maken. ([4], [5]).

De uitvinding van de MRI-scanner heeft plaatsgevonden in de jaren zeventig. Er is enige onduidelijkheid wie dit apparaat heeft uitgevonden. Echter in de literatuur worden drie namen (Raymond Vahan Damadian, Sir Peter Mansfield en Paul Christian Lauterbur) genoemd voor het leveren van een belangrijke bijdrage aan de uitvinding van de MRI-scanner.

In 1971 stelde Damadian [6] voor dat men door middel van kernspinresonantie tumoren kan onderscheiden van gewoon weefsel. Hoewel zijn methode niet praktisch bleek te zijn, heeft hij een aantal patenten in MRI-scan gerelateerde zaken [7]. Verder richtte Damadian zijn eigen bedrijf FONAR op. Dit bedrijf specialiseert zich in het produceren van MRI-scanners. Volgens Damadian [8] is hij ook de uitvinder van het apparaat en zou de scanner niet eens bestaan als hij niet geboren was. Toch zijn het Mansfield en Lauterbur die in 2003 de Nobelprijs voor Geneeskunde hebben gekregen voor het uitvinden van de MRI-scanner [9].

Lauterbur [10] beschreef in 1973 een techniek voor het maken van MRI-scans van objecten. In zijn wetenschappelijk artikel stelde hij ook voor dat dit mogelijk moest zijn voor mensen en dieren. In de daarop volgende jaren bedacht Mansfield [11] een aantal technieken ([12], [13]) waardoor het produceren van een MRI-scan aanzienlijk sneller en duidelijker werd. Dit leidde uiteindelijk tot het eerste prototype van de MRI-scanner in 1980. Het eerste apparaat kwam in 1982 uit. Sindsdien is het gebruik en het aantal apparaten sterk toegenomen [9]. De MRI-scanner bestaat uit een aantal delen, namelijk een supergeleidende magneet, de gradiëntspoelen, versterkers en ten slotte computers waarmee de data wordt gecreëerd [14]. In de onderstaande figuur 2.2 is een hedendaagse MRI-scanner afgebeeld:



Figuur 2.2: Een MRI-scanner bij een experiment.

Door middel van een MRI-scanner is het mogelijk om gedetailleerde beelden van de inwendige structuren van het lichaam te produceren. Een voorbeeld van een fMRI afbeelding is te zien in figuur 2.3. Het interpreteren van een dergelijke afbeelding wordt meestal gedaan door gespecialiseerde artsen, radiologen of röntgenologen in een MRI-onderzoek. Jaarlijks worden er ongeveer 60 miljoen van dit soort onderzoeken verricht [11]. Zoals de meeste hardware kent de MRI-scanner een proces van verbetering. Dit heeft geleid tot meerdere typen MRI-scanners voor meerdere gebruiksmogelijkheden.



Figuur 2.3: (f)MRI afbeelding van een proefpersoon.

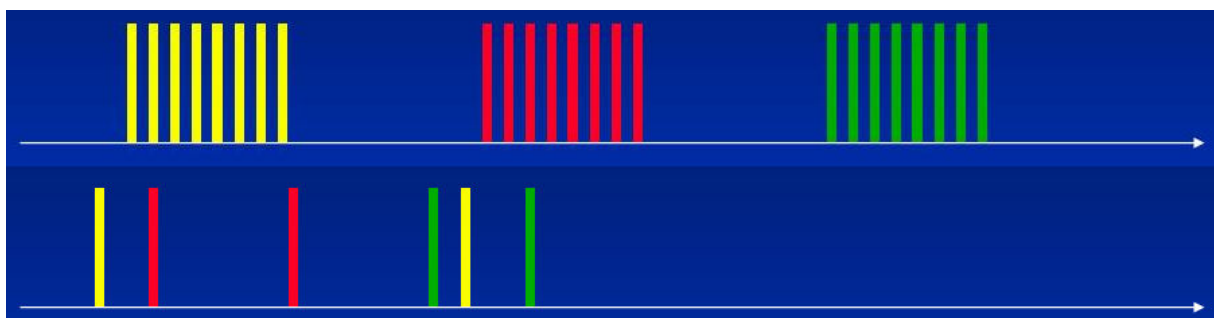
2.2 fMRI

Functionele MRI wordt tegenwoordig veelal gebruikt voor hersenactiviteitenonderzoek. Dit in tegenstelling tot de MRI, waarbij alleen wordt gekeken naar de menselijke anatomie. In een typisch onderzoek wordt bekeken of er sprake is van activiteit in een bepaald gebied van de hersenen bij het uitvoeren van een specifieke taak. Dit kan variëren van het bewegen van een vinger gedurende een halve minuut tot het bekijken van een aantal afbeeldingen [15].

De activiteit van de hersenen wordt indirect gemeten. Als er sprake is van activiteit dan is de bloedtoevoer naar dat gebied van de hersenen groter dan normaal. De zogeheten neuronen hebben bij activiteit zuurstof nodig. Dit wordt verzorgd door de rode bloedcellen van het lichaam. In de rode bloedcellen zit hemoglobine en deze bevat een belangrijke eigenschap. Zuurstofrijke hemoglobine is diamagnetisch en zuurstofarme hemoglobine is paramagnetisch. Deze verschillen in magnetische susceptibiliteit zorgen ervoor dat de MRI-scanner een signaal kan detecteren bij een experiment. Zo kan er vervolgens bepaald worden óf en belangrijker wáár er sprake is van activiteit. Het gemeten signaal wordt ook wel BOLD-signaal (Engels: *Blood Oxygen Level Dependant*) genoemd ([16], [17], [18]).

Voordat een persoon deel kan nemen aan een fMRI onderzoek moet die persoon eerst voorbereid worden. De proefpersoon vult meestal een vragenlijst in zodat er geen misverstanden kunnen ontstaan bij de onderzoekers en de proefpersoon zelf. Verder is het niet toegestaan om metalen objecten, zoals sieraden of piercings, mee te nemen in de onderzoekkamer vanwege het sterke magneetveld. Dit betekent ook dat mensen met pacemakers niet onderzocht kunnen worden, omdat het magneetveld de pacemaker kan verstoren.

In een experiment worden meestal twee condities met elkaar vergeleken. Vervolgens wordt het verschil in activiteit van de hersenen bij de twee condities afgebeeld ([19], [20]). Er zijn twee soorten experimenten die veelal gebruikt worden: 'block' en 'event-related' designs. De 'block' design methode heeft een aantal voordelen ten opzichte van de 'event related'. Eén daarvan is dat het statistisch gezien beter is in het detecteren van een signaal ([21], [22]). In een 'event-related' experiment voert de proefpersoon één of meerdere taken gedurende het gehele experiment. In een 'block' design voert de proefpersoon een bepaalde taak uit gedurende een aantal seconden. Vervolgens volgt er een rustperiode. Dit wordt dan een aantal malen herhaald. Er zijn verschillende taken die uitgevoerd kunnen worden in een experiment. Dit en ook de lengte van het experiment hangt af van het soort experiment dat uitgevoerd wordt. Een aantal van deze taken kunnen zijn: het drukken van knoppen bij bepaalde afbeeldingen, het zien van afbeeldingen of stukjes fragmenten, het horen van bepaalde geluiden, het bewegen van de vinger van de linkerhand of het onthouden van afbeeldingen.



Figuur 2.4: Het fMRI signaal bij een 'block' (boven) en 'event-related' (onder) design.

In de bovenstaande figuur 2.4 is het verschil te zien tussen een 'block' en 'event-related' design. In het experiment worden afbeeldingen getoond van een drietal personen. Elke persoon heeft een verschillende kleur.

In het bovenste plaatje van figuur 2.4 wordt in een periode met vaste lengte een aantal afbeeldingen van 1 persoon getoond en vervolgens volgt een rustperiode met eveneens een vaste lengte. Waarna dit patroon zich herhaalt tot aan het einde van het experiment. In het onderste plaatje wordt tijdens het

gehele experiment alle afbeeldingen getoond zonder dat deze een volgorde hebben. Bovendien is de tijd tot de eerst volgende afbeelding willekeurig.

Tijdens en na het uitvoeren van een experiment wordt de data¹ samengesteld. Het gemeten signaal van elke volumetrische pixel (voxel) van de hersenen wordt vastgelegd. Alvorens de data gebruikt kan worden, voor bijvoorbeeld statistische analyses, moet het bewerkt worden. Hiervoor bestaan er een aantal preprocessing stappen. Het corrigeren van de data is belangrijk, omdat de ruwe data ruis bevat. Dit kan veroorzaakt worden door de apparatuur, het experiment of de proefpersoon.

Het nadeel van de fMRI is dat de apparatuur vrij duur is [14]. Dit betekent dat de fMRI niet geschikt is voor commerciële toepassingen en het aantal onderzoeken voor medische doeleinden relatief gering is. Verder ligt de proefpersoon in een magneetveld en dit zorgt ervoor dat het lichaam van binnen verhit. Hierdoor is het niet mogelijk om de lengte van een experiment te lang te maken. De proefpersoon die in een MRI-scanner ligt, ligt in feite in een nauwe buis. Omdat dit vergelijkbaar is met een kleine ruimte, is dit dus niet geschikt voor claustrofobische personen. Ook maakt het apparaat veel geluid en dit is niet bevorderlijk voor het gehoor van de proefpersoon. Ten slotte kan het signaal verstoord worden door lichte bewegingen van proefpersoon.

De fMRI techniek heeft echter ook een aantal voordelen ten opzichte van andere bestaande technieken zoals de PET of EEG [15]. Ten eerste duurt een experiment relatief vrij kort en wordt de data vrij snel verkregen. In een tijdbestek van 15 minuten kan er een experiment uitgevoerd worden. Verder is een fMRI-scan een niet-invasieve behandeling. Het doet geen pijn en heeft geen ongewenste effecten. Ten slotte wordt de proefpersoon niet bestraald.

¹ De data wordt besproken in hoofdstuk 4.

3. Meervoudig Lineaire Regressiemodel & Power Analyse

3.1 Regressieanalyse

In een regressieanalyse wordt onderzocht of er een verband is tussen twee soorten variabelen: de response en de verklarende variabelen. De respons variabele wordt ook wel de afhankelijke variabele genoemd en deze hangt af van de onafhankelijke variabelen. Als deze relatie lineair is, dan spreken wij van een lineair verband. In dat geval kan de data gefit worden in een enkel- of meervoudige lineair regressiemodel. Het doel van het opstellen van een lineair regressiemodel is om de responsvariabele te verklaren door middel van een (lineaire) functie van de verklarende variabelen. In deze functie zijn er een aantal onbekende parameters die geschat moeten worden. Verder wordt er vanuit gegaan dat de responsvariabele waargenomen wordt met een meetfout [23].

Het lineaire regressiemodel met p verklarende variabelen en n observaties ziet er als volgt uit:

$$Y_i = \beta_0 + x_{i1}\beta_1 + \dots + x_{ip}\beta_p + e_i \quad i=1, \dots, n \quad (3.1)$$

Waarbij Y_i de i^{de} respons is, e_i de i^{de} stochastische meetfout, x_{ij} de i^{de} waarde van de j^{de} verklarende variabele en β_j de bijbehorende (onbekende) parameter van de j^{de} verklarende variabele.

Er wordt aangenomen dat de meetfouten onafhankelijk en identiek normaal verdeeld zijn met verwachting 0 en variantie σ^2 :

$$\begin{aligned} e_i &\sim N(0, \sigma^2) \\ Ee_i &= 0 \\ Ee_i e_j &= \begin{cases} \sigma^2 & i = j \\ 0 & i \neq j \end{cases} \end{aligned} \quad (3.2)$$

De vergelijkingen (3.1) en (3.2) kunnen ook in matrix vorm worden geschreven. Dit ziet er dan als volgt uit:

$$\begin{aligned} Y &= X\beta + e \\ e &\sim N_{n \times n}(0, \sigma^2 I_{n \times n}) \\ Ee &= 0 \\ \text{cov}(e) &= \sigma^2 I_{n \times n} \end{aligned} \quad (3.3)$$

De matrix X heeft de dimensie $n \times (p + 1)$. Er zijn immers n observaties en p verklarende variabelen. De eerste kolom, ook wel de 0^{de} kolom genoemd, is de kolomvector met n enen. Y en e , de n waargenomen responsvariabelen en de stochastische meetfouten, zijn vectoren met lengte n . De vector β bevat de onbekende parameters en heeft de lengte p .

De onbekende parameters van vergelijking (3.3) worden in het geval van normaal verdeelde meetfouten met de kleinste kwadratenmethode geschat [24]. Dit wordt gedaan door de kwadraatsom te minimaliseren:

$$S(\beta) = \sum_{i=1}^n (Y_i - EY_i)^2 \quad (3.4)$$

Omdat $EY = X\beta + Ee = X\beta$ geldt het volgende:

$$S(\beta) = (Y - X\beta)^T (Y - X\beta) \quad (3.5)$$

De vergelijking (3.5) wordt geminimaliseerd door deze te differentiëren naar β en gelijk te stellen aan 0. Wij krijgen dan het volgende:

$$X^T (Y - X\beta) = 0 \rightarrow X^T X \beta = X^T Y \quad (3.6)$$

De oplossing die hieraan voldoet is:

$$\hat{\beta} = (X^T X)^{-1} X^T Y \quad (3.7)$$

Echter bestaat de inverse van $X^T X$ alleen als deze van maximale rang is. Dit betekent dus dat de matrix X maximale rang moet hebben. Anders gezegd: de rang van matrix X van vergelijking (3.3) moet gelijk zijn aan $p + 1$. De vergelijking (3.5) in het punt $\hat{\beta}$ wordt ook wel de residuele kwadraatsom genoemd.

De covariantiematrix van $\hat{\beta}$ is het volgende:

$$\text{cov}(\hat{\beta}) = \sigma^2 (X^T X)^{-1} \quad (3.8)$$

Waarbij σ^2 geschat wordt met het volgende:

$$\hat{\sigma}^2 = \frac{S(\hat{\beta})}{(n - p - 1)} \quad (3.9)$$

De (geschatte) waarden van de diagonaal van de covariantiematrix zeggen iets over de nauwkeurigheid van de schattingen van $\hat{\beta}$. Hoe groter deze waarden, hoe onnauwkeurig de schattingen van $\hat{\beta}$ zijn.

Een maat voor de kwaliteit van het model is de zogeheten determinatiecoëfficiënt. Dit is de fractie verklaarde variantie van het model. De determinatiecoëfficiënt is als volgt gedefinieerd:

$$\mathfrak{R}^2 = \frac{SYY - RSS}{SYY} \quad (3.10)$$

Waarbij SYY de residuele kwadraatsom van het model zonder verklarende variabelen is en RSS de residuele kwadraatsom van het model met verklarende variabelen. De waarde van de determinatiecoëfficiënt ligt tussen 0 en 1. Hoe dichter de waarde bij 1 ligt, hoe beter het model is.

Een belangrijk onderdeel van het opstellen van een regressiemodel is het bepalen van welke verklarende variabelen in het model worden toegevoegd. Dit wordt gedaan door te toetsen of de waarde van de parameter β significant is of niet. Dit kan gedaan worden de t-toets. De volgende hypothesen worden getoetst:

$$\begin{aligned} H_0 : \beta_k = 0; \beta_j \text{ willekeurig voor } 0 \leq j \leq p, j \neq k. \\ H_1 : \beta_k \neq 0; \beta_j \text{ willekeurig voor } 0 \leq j \leq p, j \neq k. \end{aligned} \quad (3.11)$$

De toetsingsgrootheid is als volgt gedefinieerd:

$$|t_k| = \frac{|\hat{\beta}_k|}{\sqrt{\widehat{\text{cov}}(\hat{\beta})_{kk}}} \quad (3.12)$$

Deze heeft onder de nulhypothese een t -verdeling met $(n - p - 1)$ vrijheidsgraden. De nulhypothese wordt verworpen als de overschrijdingskans kleiner is dan de gekozen onbetrouwbaarheidsdrempel α .

Het lineaire regressiemodel is een vrij simpel model en kan bovendien toegepast worden in allerlei situaties en voor verschillende doeleinden. Maar de kwaliteit van het model hangt af van een aantal zaken. Ten eerste moet er sprake zijn van een lineair verband. In het geval van tweedimensionale data kan dit door middel van scatterplots onderzocht worden. In het geval van meerdere verklarende variabelen is dit niet mogelijk met scatterplots, maar zijn er wel andere methoden om te bepalen welke variabelen toe te voegen aan het lineair regressiemodel. Ook is het wenselijk om extreme waarden van de data te verwijderen, omdat dit ervoor kan zorgen dat de schattingen van de parameters minder nauwkeurig zijn. Het vinden van extreme of invloedspunten is echter niet zo gemakkelijk. Ten slotte moet er opgepast worden voor collineariteit. Als hier sprake van is dan kan het zijn dat de matrix in vergelijking (3.5) geen maximale rang heeft en dus ook geen inverse. Verder kan de variantie van de geschatte parameter β groot zijn, wat betekent dat deze schatting erg onnauwkeurig is.

De data die verkregen wordt bij een fMRI experiment kan gefit worden in een meervoudige lineaire regressiemodel. Worsley [25] stelt een aanpak voor om een fMRI signaal te modelleren. Dit wordt gedaan in drie stappen: het corrigeren van de fMRI respons, het opstellen van het lineaire regressiemodel met een aantal verklarende variabelen en ten slotte het modelleren van de meetfouten van het model. Het corrigeren van de fMRI respons is noodzakelijk omdat er een vertraging optreedt in het signaal als er sprake is van activiteit. Vervolgens kan een lineair model worden samengesteld als in de vergelijking (3.1). Echter merkt Worsley op dat in sommige voxels in fMRI datapunten drift wordt getoond. Als deze niet verwijderd worden dan kan de schatting van de parameters van het lineaire regressiemodel onnauwkeurig zijn. Het verwijderen van deze drift kan gerealiseerd worden door het toevoegen van verklarende variabelen. Bijvoorbeeld het verwijderen van een polynomische drift van orde q kan verwijderd worden door het introduceren van q verklarende variabelen in het lineaire regressie model. Ten slotte moet er rekening gehouden worden met onderling gecorreleerde meetfouten. Het lineaire regressiemodel ziet er als volgt uit:

$$Y_i = \beta_0 + \underbrace{x_{i1}\beta_1 + \dots + x_{ip}\beta_p}_{\text{fMRI respons}} + \underbrace{x_{ip+1}\beta_{p+1} + \dots + x_{im}\beta_m}_{\text{drift}} + e_i \quad (3.13)$$

3.2 Power Analyse

Bij de meeste analyses is het niet mogelijk om de gehele populatie te toetsen. Dit kan door een aantal factoren komen, bijvoorbeeld de populatie is te groot om gemeten te worden of er zijn niet genoeg middelen om dit te realiseren. Om toch iets te kunnen zeggen over de gehele populatie wordt meestal een steekproefpopulatie gekozen.

In de meeste gevallen wordt er een bepaalde statistiek berekend om bepaalde karakteristieken van een populatie te beschrijven. Om iets zinnigs te zeggen over de gehele populatie moet de steekproefgrootte niet te klein zijn. Ook is het niet verstandig om dit aantal te groot te laten zijn, omdat hier waarschijnlijk ook tijd en kosten aan verbonden zijn. Het is voor een statische analyse van belang om de juiste steekproefgrootte te kiezen.

Bij een toets wordt een nulhypothese en een alternatieve hypothese opgesteld. Vervolgens wordt de toets uitgevoerd en wordt de nulhypothese verworpen of niet. Hierbij kunnen er twee soorten fouten gemaakt worden: de nulhypothese wordt ten onrechte verworpen of de nulhypothese wordt ten onrechte niet verworpen. De eerste fout wordt ook wel een fout van de eerste soort genoemd. De kans hierop is bij een “goede” toets kleiner dan de betrouwbaarheidsdrempel α . De tweede fout wordt ook wel een fout van de tweede soort of γ genoemd. Het onderscheidend vermogen (of in het Engels: *power*) van een toets is gelijk aan $1 - \gamma$. Met andere woorden een toets met een groot onderscheidend vermogen is dus gewenst. Dit betekent dat α en γ het liefst zo klein mogelijk moeten zijn. Echter is het zo klein mogelijk maken van de onbetrouwbaarheidsdrempel niet verstandig, omdat het onderscheidend vermogen van de toets ook afneemt, wat betekent dat de kans op een fout van de tweede soort erg groot wordt [26].

Het onderscheidend vermogen van een toets hangt mede af van de steekproefgrootte. Hoe groter de steekproefgrootte, hoe groter het onderscheidend vermogen van een toets. Het bepalen van de minimale steekproefgrootte zodat een toets een onderscheidend vermogen heeft dat groter is dan een vooraf bepaald getal, is dus wenselijk. In een power analyse bepaalt men hoe groot de steekproefpopulatie moet zijn wil de kans op een fout van de tweede soort kleiner of gelijk zijn dan γ , bij een gegeven onbetrouwbaarheidsdrempel.

3.3 Het minimale aantal scans in een fMRI experiment

In een onlangs verschenen artikel [27] is de relatie tussen het aantal scans in een fMRI experiment en de temporele signaal-ruisverhouding² (TSNR) onderzocht. Het resultaat is een formule voor het aantal scans bij het detecteren van een signaal.

$$N = 8 \left(\frac{\operatorname{erfc}^{-1}(P)}{(TSNR)(\operatorname{eff})} \right)^2 \quad (3.14)$$

Waarbij P de onbetrouwbaarheidsdrempel is, eff de relatieve sterkte van het signaal en erfc de complementaire error functie. Deze formule is vervolgens gevalideerd door deze toe te passen op vier datasets. Dit resulteerde in een verbeterde formule, waarbij detectie van een signaal wordt gegarandeerd.

Volgens Bijma et. al. in [28] is deze afleiding onjuist om een aantal redenen. Ten eerste wordt er vanuit gegaan dat het detecteren van een signaal deterministisch is. Dit is echter niet het geval en er is geen garantie dat de kans om een signaal te detecteren gelijk is aan 1. Dit zou overeenkomen met een kans op een fout van de tweede soort die gelijk is aan 0. Bovendien leiden ze de formule van het aantal scans af door deze uit te drukken in de temporele signaal-ruisverhouding in plaats van deze af te leiden door middel van een power analyse.

In [28] wordt allereerst een lineair model opgesteld als in vergelijking (3.13). Voor één voxel ziet het model met $(p+q)$ verklarende variabelen (p originele variabelen en q geïntroduceerde variabelen voor de drift (inclusief het intercept), ook wel confounders genoemd) er als volgt uit:

$$y = X_1 \beta_1 + X_2 \beta_2 + e \quad (3.15)$$

Waarbij y het gemeten signaal is. X_1 en X_2 de matrices waarbij elke kolom een verklarende variabele (in dit geval bevat elke kolom van matrix X_2 confounders) bevat en β_1 en β_2 de bijbehorende parameters die geschat moeten worden. Tot slot bevat e de stochastische meetfouten.

Ook in dit geval kunnen de onbekende parameters geschat worden met de kleinste kwadratenmethode. Omdat de parameters in β_2 niet interessant zijn, kunnen wij ook de parameters in β_1 in één keer uitrekenen³. Deze ziet er als volgt uit:

$$\hat{\beta}_1 = (X_1^T P_2^\perp X_1)^{-1} X_1^T P_2^\perp y \quad (3.16)$$

Waarbij P_2^\perp een projectiematrix is en als volgt is gedefinieerd:

$$P_2^\perp = I_{n \times n} - X_2 (X_2^T X_2)^{-1} X_2^T \quad (3.17)$$

² De TNSR van een tijdreeks x_i is gedefinieerd als het gemiddelde van de tijdreeks gedeeld door de standaardafwijking van de tijdreeks.

³ Ook in dit geval nemen wij aan dat de matrix $X = (X_1 \ X_2)$ maximale rang heeft en de meetfouten onderling onafhankelijk normaal verdeeld zijn.

Vermenigvuldigen wij de bovenstaande projectiematrix van vergelijking (3.17) met de vergelijking (3.15) dan krijgen wij het deconfounded signaal:

$$\begin{aligned} P_2^\perp y &= P_2^\perp X_1 \beta_1 + \underbrace{P_2^\perp X_2 \beta}_{=0} + P_2^\perp e \\ P_2^\perp y &= P_2^\perp X_1 \beta_1 + P_2^\perp e \end{aligned} \quad (3.18)$$

Dit kunnen wij herschrijven als het volgende:

$$\begin{aligned} P_2^\perp y &= P_2^\perp X_1 \beta_1 + P_2^\perp e \\ \tilde{y} &= \tilde{X}_1 \beta_1 + \tilde{e} \end{aligned} \quad (3.19)$$

De corresponderende uitdrukkingen voor vergelijkingen (3.8) en (3.9) zijn in dit geval:

$$\begin{aligned} \text{cov}(\hat{\beta}_1) &= \hat{\sigma}^2 (\tilde{X}_1^T \tilde{X}_1)^{-1} \\ \hat{\sigma}^2 &= \frac{S(\hat{\beta}_1)}{(n-p-q)} \end{aligned} \quad (3.20)$$

De toetsingsgrootheid van vergelijking (3.12) vereenvoudigt in het geval van één verklarende variabele ($p = 1$):

$$T = \frac{\hat{\beta}_1}{\hat{\sigma}(\tilde{x}_1^T \tilde{x}_1)^{\frac{1}{2}}} \quad (3.21)$$

Deze heeft onder de nulhypothese een t -verdeling met $(n - q)$ vrijheidsgraden.

Om het minimale aantal scans te bepalen definiëren wij allereerst b als de ware waarde van β_1 . Bij een hypothese van $H_0 : \beta_1 \leq 0$, is het onderscheidend vermogen voor een zekere waarde van $\beta_1 = b > 0$ gelijk aan de volgende kans:

$$P_b(T > t_{n-q-1; \alpha}) \quad (3.22)$$

Kortom, voor een kans op een fout van de tweede soort kleiner dan γ , moet de kans in vergelijking (3.22) groter of gelijk zijn aan $1 - \gamma$. Deze kans is in [28] benaderd door het volgende te gebruiken:

$$t_{n-q-1; \alpha} \approx \xi_\alpha \quad (3.23)$$

Hieruit volgt dat vergelijking (3.22) uitgewerkt kan worden:

$$P_b(T > t_{n-q-1; \alpha}) \approx P_b\left(T - \sqrt{n} \frac{b\sigma_x}{\sigma} > \xi_\alpha - \sqrt{n} \frac{b\sigma_x}{\sigma}\right) \quad (3.24)$$

De kans in vergelijking (3.24) kan benaderd worden door de normale verdeling. Wij krijgen dan het volgende:

$$P_b\left(T - \sqrt{n} \frac{b\sigma_x}{\sigma} > \xi_\alpha - \sqrt{n} \frac{b\sigma_x}{\sigma}\right) \approx 1 - \Phi\left(\xi_\alpha - \sqrt{n} \frac{b\sigma_x}{\sigma}\right) \quad (3.25)$$

Hieruit volgt dat het minimale aantal scans n voor zekere waarden van b en een onderscheidend vermogen van minimaal $1-\gamma$ gelijk is aan het volgende:

$$n \geq (\xi_\alpha + \xi_\gamma)^2 \frac{\sigma^2}{b^2 \sigma_x^2} \quad (3.26)$$

Waarbij ξ_α en ξ_γ de α - en γ -bovenkwantiel van de standaard normale verdeling. $\left| \frac{b\sigma_x}{\sigma} \right|$ kan geïnterpreteerd worden als de signaal-ruisverhouding en σ_x^2 is als volgt gedefinieerd:

$$\sigma_x^2 = \frac{1}{n_{\tilde{x}_1}} \|\tilde{x}_1\|^2 \quad (3.27)$$

Het minimale aantal scans voor de tweezijdige nulhypothese worden op dezelfde manier afgeleid in [28].

4. Simulatieonderzoek

4.1 Onderzoeksopzet

De hoofdvraag van dit onderzoek is als volgt geformuleerd:

Werkt de gestelde n voor het minimale aantal scans voor Gaussische en fMRI ruis?

De formule hiervoor is afgeleid in [28] voor de een- en tweezijdige nulhypothese. In dit onderzoek zullen wij ons alleen richten op de eenzijdige nulhypothese. Wij gebruiken dus de formule in vergelijking (3.26). Deze formule is afgeleid door een aantal benaderingen en aannamen te maken.

Wij maken verschil tussen twee typen ruis: Gaussische en fMRI. In het voorgaande hoofdstuk hebben wij aangenomen dat de meetfouten normaal verdeeld zijn. Hoewel deze aanname niet ongewoon is, moeten wij dit controleren. Om het signaal te simuleren, zullen wij de meetfouten op twee manieren simuleren. Namelijk uit de normale verdeling (Gaussische ruis) en het gebruik van geschatte meetfouten afkomstig uit fMRI data (fMRI ruis). Wij bekijken dus of de formule voor het bepalen van het minimale aantal scans voor beide gevallen werkt.

In [28] is de kans in vergelijking (3.22) exact berekend en benaderd. Het verschil tussen het minimale aantal scans dat berekend is, varieert van 1% tot maximaal 10%. Het benaderde aantal is altijd lager dan het exacte aantal. Het verschil wordt groter naarmate de signaal-ruisverhouding toeneemt. Wij vermoeden dus dat het gesimuleerde onderscheidend vermogen kleiner zal zijn dan het werkelijke onderscheidend vermogen. In de simulatie corrigeren wij het aantal minimale scans door er 10% bij op te tellen. In sommige gevallen is het minimale aantal scans kleiner dan 100. Echter komt in de praktijk een scanduur van minder dan vijf minuten⁴ niet voor. Ook hier voeren wij een correctie uit, namelijk dat het aantal minimale scans groter of gelijk is aan 100.

Wij hebben de volgende stappen in ons simulatieonderzoek uitgevoerd:

0. Gebruik data om β_I te schatten en σ , σ_x en ε te bepalen.
1. Bepaal het aantal scans n voor een zekere b , onbetrouwbaarheidsdrempel α en onderscheidend vermogen $1-\gamma$. Voer correctie uit op n :

$$n = \max(100, 1.1 \cdot n) \quad (4.1)$$

2. Simuleer het deconfounded signaal:

$$\tilde{y}_{1,sim}, \dots, \tilde{y}_{n,sim} \quad \tilde{y}_{i,sim} = b \cdot \tilde{x}_i + P_2^\perp \cdot \text{ruis} \quad (4.2)$$

Gebruik voor ruis de normale verdeling of fMRI ruis:

$$\begin{aligned} \text{Gaussische ruis: } \varepsilon &\sim N_{n \times n}(0, \sigma^2 I_{n \times n}) \\ \text{fMRI ruis: } \varepsilon &= \hat{\varepsilon}_1, \dots, \hat{\varepsilon}_n \end{aligned} \quad (4.3)$$

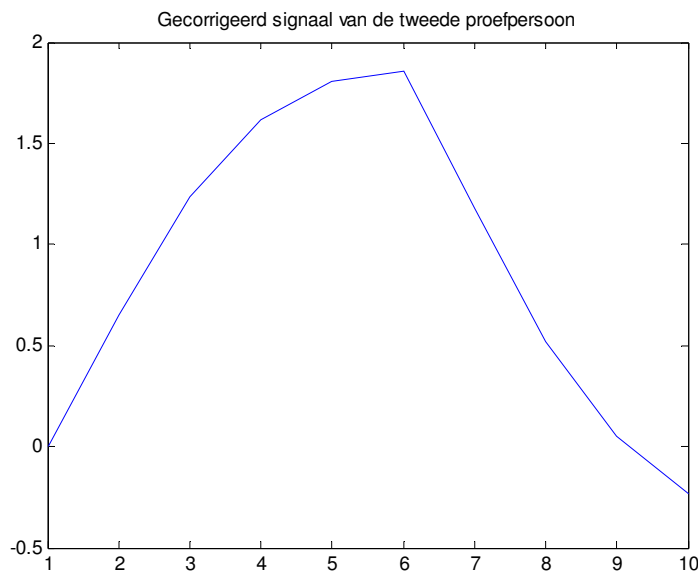
3. Schat de waarde voor β_I uit de gesimuleerde data en toets deze op significantie met de t-toets.
4. Herhaal stappen 1, 2 en 3 B keer en bepaal het onderscheidend vermogen:

$$\text{onderscheidend vermogen} = \frac{\text{aantal significante } \beta_I \text{'s}}{B} \quad (4.4)$$

⁴ Elke scan heeft een duur van 3 seconden. Dit betekent dat 100 scans een duur heeft 300 seconden, dus vijf minuten.

4.2 Data

De data die gebruikt is in het simulatieonderzoek is verkregen door de medische faculteit van de VU. De data bestaat uit twee delen: het fMRI-signaal van een aantal voxels en de designmatrix. De designmatrix bevat een aantal confounders (X_2) en één onafhankelijke variabele (x_1). Het fMRI-signaal is een matrix met dimensie $n_t \times n_v$. Waarbij n_t gelijk is aan het aantal scans en n_v gelijk is aan het aantal voxels. In [25] wordt opgemerkt dat het fMRI-signaal gecorrigeerd moet worden, omdat er een kleine vertraging optreedt in de bloedtoevoer als er sprake is van activiteit in het neuron. Dit vertaalt zich in een klein dipje in het signaal. Om deze correctie uit te voeren, zijn een aantal technieken voorgesteld [25]. De data die wij gebruikt hebben, is al gecorrigeerd. In de onderstaande figuur 4.1 is een voorbeeld te zien van het gecorrigeerde signaal:



Figuur 4.1: Voorbeeld van een gecorrigeerd signaal van de tweede proefpersoon.

De proefpersonen zijn via een advertentie gerekruteerd. Proefpersoon 1 was een vrouw, geboren in 1985 en is gescand op 14 februari 2006. De tweede proefpersoon was een man, geboren in 1974 en is gescand op 24 januari 1974.

De motortask was een toevoeging aan het standaard EEG/fMRI resting state protocol. De proefpersoon moest gedurende 30 seconden niets doen, 30 seconden knijpbewegingen maken met rechter hand, etc. De task werd aangegeven via een geprojecteerde dia. Het experiment duurde ongeveer 15 minuten.

In de onderstaande tabel 4.1 zijn de kenmerken van de data van beide proefpersonen gegeven:

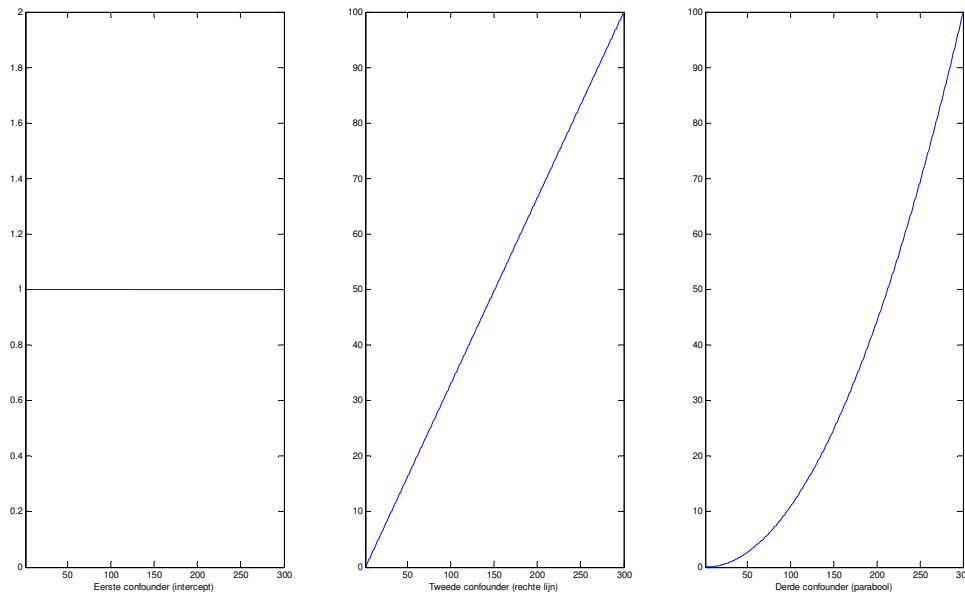
<i>Proefpersoon</i>	<i>Aantal voxels:</i>	<i>Aantal confounders:</i>	<i>Aantal scans:</i>
1	154	10	300
2	166	9	300

Tabel 4.1: kenmerken van de data die gebruikt is in het simulatieonderzoek

Wij hebben de data geanalyseerd door deze grafisch en numeriek weer te geven. Bovendien hebben wij een aantal aannamen getoetst. Ook hebben wij extreme waarden en collineariteit onderzocht. De resultaten zijn in de volgende paragraaf weergegeven.

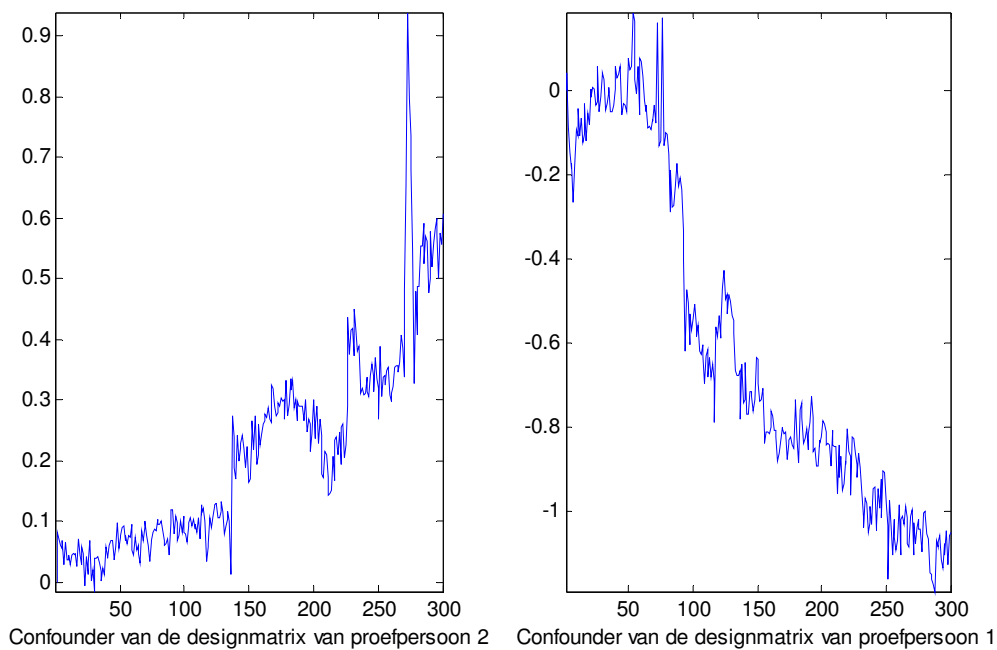
4.3 Dataverkenning

Zoals te zien is in tabel 4.1 is het aantal confounders voor beide proefpersonen verschillend. In beide gevallen is de eerste confounder het intercept, de tweede een rechte lijn en de derde is een parabolische lijn. In figuur 4.2 zijn deze confounders weergegeven:



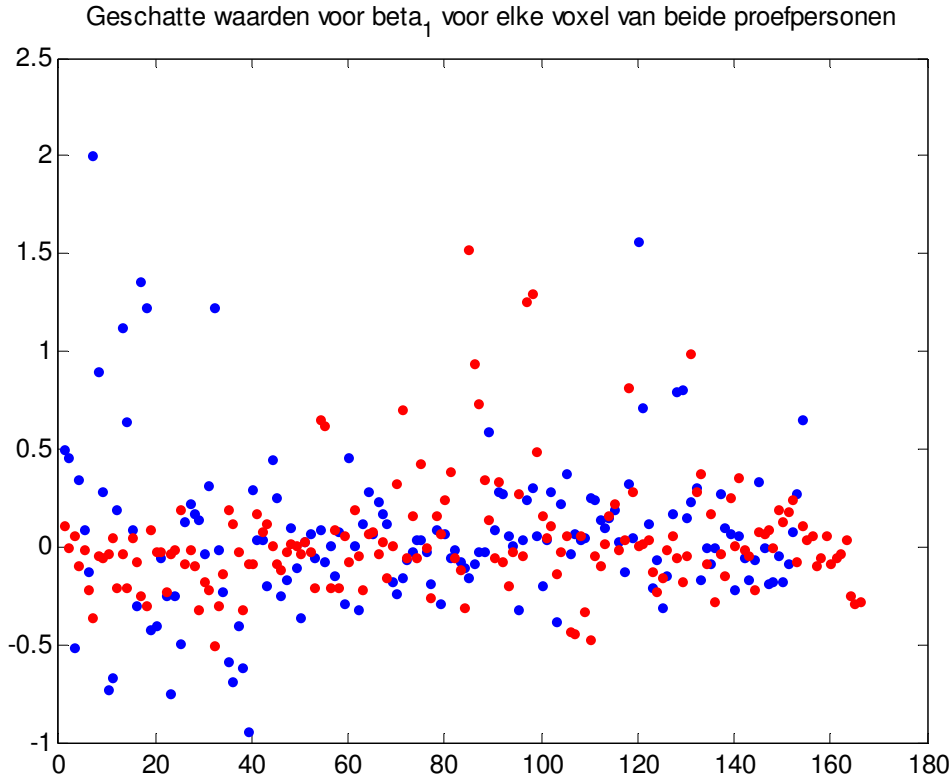
Figuur 4.2: De eerste drie confounders van beide proefpersonen.

De overige confounders zijn verschillend en in figuur 4.3 zijn een aantal van deze weergegeven:



Figuur 4.3: Voorbeeld van confounders van de designmatrix van beide proefpersonen.

Ook het aantal voxels in de data zijn verschillend voor beide proefpersonen. De data van proefpersoon 1 bevat metingen van 154 voxels en die van proefpersoon 2 166 (tabel 4.1). Voor elke voxel van beide proefpersonen hebben wij de onbekende parameter β_l geschat. De geschatte waarden liggen in het interval $[-1, 2]$. In het geval van proefpersoon 1 liggen meer dan 85% van alle geschatte waarden in het interval $[-0.5; 0.5]$. In het geval van proefpersoon 2 is dit meer dan 90%. In figuur 4.4 is het resultaat te zien:



Figuur 4.4: Geschatte waarden voor β_l voor elke voxel van proefpersoon 1 (blauw) en 2 (rood).

In figuur 4.2 is te zien dat de tweede en derde confounders vergelijkbaar zijn. Dit zou kunnen betekenen dat zij collineair zijn. Dit kan onderzocht worden door conditie indices en variantie decomposities uit te rekenen van $\hat{\beta}$.

Deze zijn als volgt gedefinieerd: de k^{de} conditie-index is gelijk aan:

$$K_k(X) = \frac{\max_j \mu_j}{\mu_k} \quad (4.5)$$

Waarbij μ_j de singuliere waarden van de matrix X is.

De variantiedecompositie:

$$\pi_{kj} = \frac{v_{jk}^2 / \mu_j^2}{\sum_{k=1}^{p+1} v_{jk}^2 / \mu_j^2}, \quad k = 1, \dots, q+1; \quad j = 0, 1, \dots, q. \quad (4.6)$$

Waarbij v_{jk} de waarden van de eigenvector V_k van $X^T X$ behorend bij μ_k is.

In figuur 4.5 zijn een aantal conditie indices en variantie decomposities van $\hat{\beta}$ voor beide proefpersonen weergegeven⁵:

$$\begin{pmatrix} \text{cond.ind.} & \text{prop}(\hat{\beta}_2) & \text{prop}(\hat{\beta}_3) \\ 87.898 & 0.97863 & 0.79096 \end{pmatrix}$$

$$\begin{pmatrix} \text{cond.ind.} & \text{prop}(\hat{\beta}_0) & \text{prop}(\hat{\beta}_2) & \text{prop}(\hat{\beta}_3) \\ 58.946 & 0.85754 & 0.94907 & 0.9639 \end{pmatrix}$$

Figuur 4.5: Conditie indices en variantiedecompositie van de geschatte parameters β_2 en β_3 voor proefpersoon 1 (boven) en β_1 , β_2 en β_3 voor proefpersoon 2 (onder).

Zoals te zien is in figuur 4.5 zijn de conditie-indices corresponderend met μ_{10} bij proefpersoon 1 en μ_{11} bij proefpersoon 2 relatief groot. Bovendien corresponderen deze conditie indices met twee of meer variantieproporties dicht bij 1. In beide gevallen zijn de tweede en derde kolom van de designmatrix collineair. Dit zijn de tweede en derde confounder (rechte en parabolische lijn). Verder merken wij ook op dat in het geval van de tweede proefpersoon het intercept ook collineair is met de tweede en derde confounder. Hoewel deze confounders toegevoegd zijn in de data om drift te modelleren, ontstaat er een ongewenste situatie van collineariteit. Dit kan ervoor zorgen dat de schattingen van de onbekende parameters niet nauwkeurig zijn.

Wij hebben aangenomen dat de meetfouten normaal verdeeld zijn. Om te onderzoeken of dit daadwerkelijk het geval is, maken wij een QQplot van de residuen van een aantal voxels uitgezet tegen de normale verdeling. Het is niet zinvol om alle residuen te onderzoeken, omdat dit aantal te groot is. Wij hebben dus besloten om een willekeurig aantal residuen te bekijken⁶ en hieruit onze conclusies te trekken. In figuur 4.6 en 4.7 is het resultaat te zien van de voxels 4 en 42 in het geval van proefpersoon 2. Zoals te zien is in figuur 4.6 en 4.7 liggen de punten vrijwel op de (rode) rechte lijn. Echter merken wij op dat dit lang niet geldt voor alle punten in de figuren. De punten van deze figuren liggen soms boven en soms beneden de rechte lijn. Vooral de punten in de staart wijken af van de rechte lijn. Een nader onderzoek is dus gewenst. Alvorens wij de normaliteit toetsen, concluderen wij aan de hand van de bovenstaande figuren dat ons vermoeden (aanname) dat de meetfouten afkomstig kunnen zijn uit de normale verdeling bevestigd wordt.

Wij gaan de normaliteitveronderstelling van de meetfouten toetsen. De hypothese luidt:

$$\begin{aligned} H_0 &: F \in F_0 \\ H_1 &: F \notin F_0 \end{aligned} \tag{4.7}$$

Waarbij F_0 de familie van normale kansverdelingen met verwachting $\mu = 0$ en σ^2 onbekend.

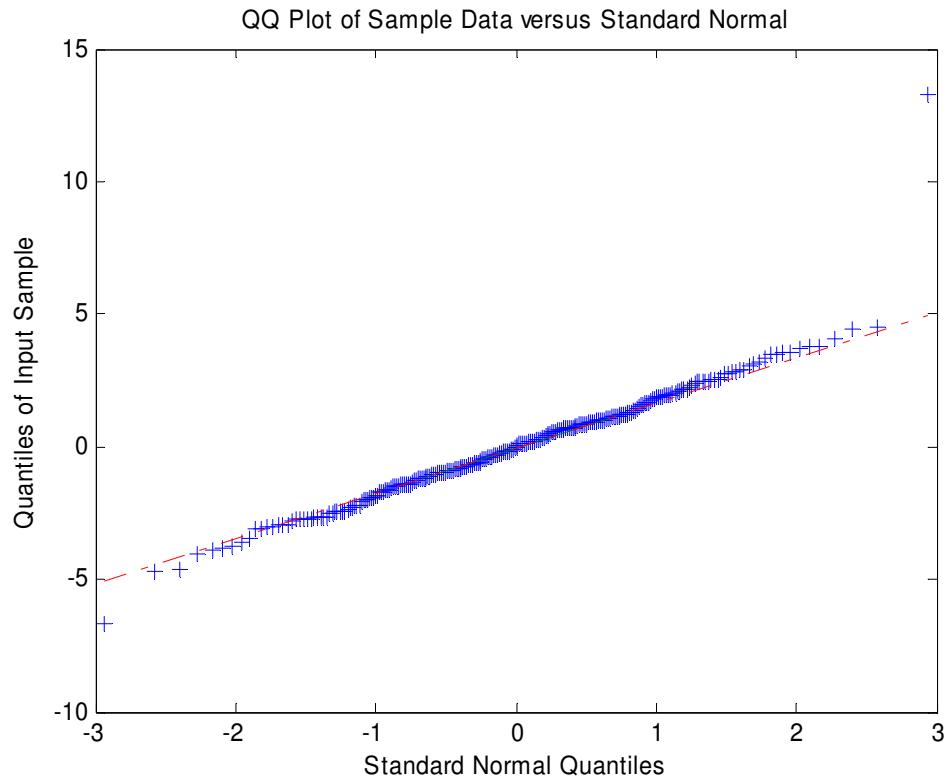
De toetsingsgrootheid is:

$$D_n = \sup_{-\infty < x < \infty} \left| \hat{F}_n(x) - \Phi(x/S) \right| \tag{4.8}$$

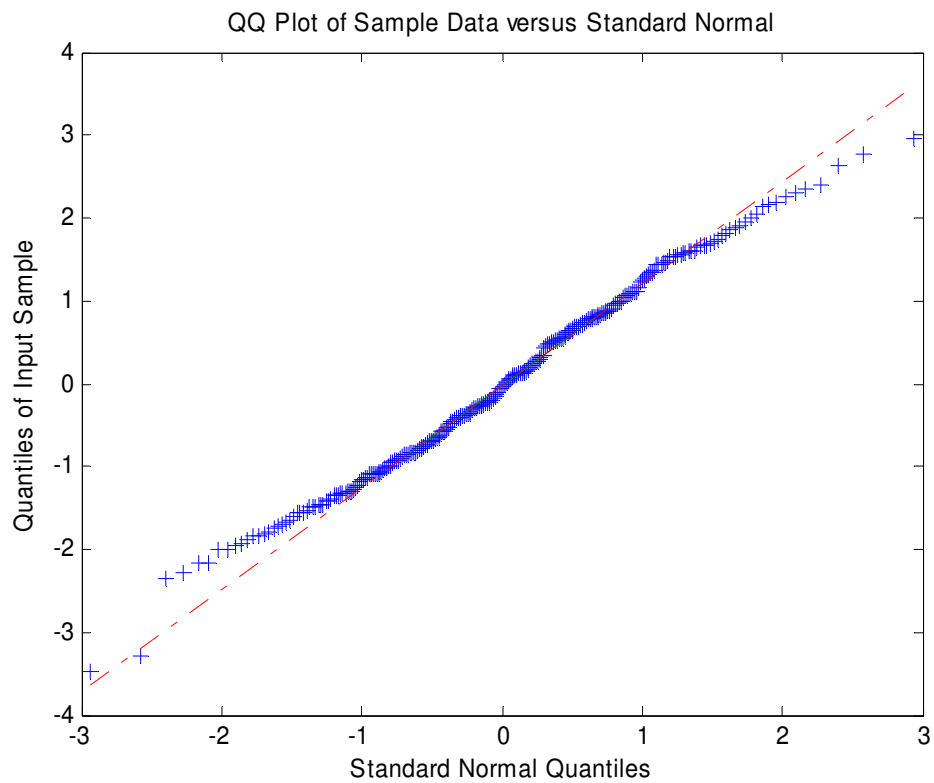
Waarbij $\hat{F}_n(x)$ de empirische verdelingsfunctie van de residuen is. De toetsingsgrootheid is onder de nulhypothese verdelingsvrij.

⁵ Beide designmatrices zijn geschaald.

⁶ In de bijlagen zijn de QQplots van een zestal voxels opgenomen.



Figuur 4.6: QQplot van de residuen van voxel 4 uitgezet tegen de normale verdeling.



Figuur 4.7: QQplot van de residuen van voxel 42 uitgezet tegen de normale verdeling.

De waargenomen waarden voor S^2 en D_n voor de residuen van het tweetal voxels is te zien in tabel 4.2. Om de bijbehorende overschrijdingskansen te bepalen, hebben wij 1000 bootstrapwaarden van D_n gegenereerd en vervolgens berekend wat de fractie is dat deze groter is dan de waargenomen waarde van D_n .

Voxel	S^2	D_n	Overschrijdingskans
4	3.96	0.056	0.026
42	1.35	0.040	0.346

Tabel 4.2: Waargenomen waarden voor S^2 en D_n en de bijbehorende overschrijdingskans.

Beide overschrijdingskansen zijn groter dan het onbetrouwbaarheidsdrempel van 0.01. Wij verwerpen de nulhypothese niet en concluderen dat de meetfouten waarschijnlijk uit de normale verdeling afkomstig zouden kunnen zijn. Hoewel dit voor de meeste voxels geldt, zijn er ook voxels waarbij de overschrijdingskans ongeveer gelijk is aan de onbetrouwbaarheidsdrempel.

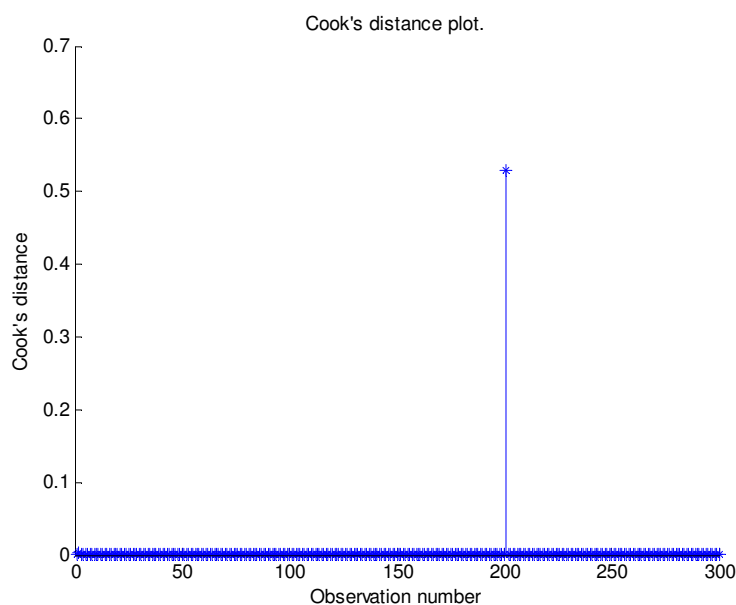
Om te bekijken of de dataset invloedspunten bevat, berekenen wij de Cook's afstand van elk punt. Punten met een relatief grote Cook's afstand of een Cook's afstand van groter of gelijk aan 1 hebben een grote invloed. Dit kunnen dus extreme waarden zijn. Het is wenselijk om deze punten van de dataset te verwijderen.

De Cook's afstand D_i voor het i^{de} punt is gedefinieerd als:

$$D_i = \frac{(\hat{Y}_{(i)} - \hat{Y})^T (\hat{Y}_{(i)} - \hat{Y})}{q\hat{\sigma}^2} \quad (4.9)$$

Waarbij $\hat{Y} = X\hat{\beta}$ en $\hat{Y}_{(i)} = X\hat{\beta}_{(i)}$. De index (i) betekent met het i^{de} waarnemingspunt weggelaten.

Ook in dit geval kunnen wij een aantal voxels bekijken. Wij besluiten om een willekeurige voxel van proefpersoon 1 te bekijken en de Cook's afstand van de waarnemingen te bepalen. In figuur 4.8 is het resultaat te zien:



Figuur 4.8: Cook's afstanden van de punten in het geval van voxel 16 van proefpersoon 1.

Zoals te zien is in figuur 4.8 is er punt met een relatieve grote Cook's afstand. Dit is mogelijk een invloedspunt en het is verstandig om dit punt te verwijderen van de dataset. Hoewel wij in dit geval maar één invloedspunt vinden, kan het zo zijn dat er nog meer punten zijn in de dataset die een relatieve grote Cook's afstand hebben.

Een andere manier om extreme punten te identificeren, is het verschil van elk punt met zijn voorganger te berekenen. En vervolgens deze punten te onderzoeken met een grote absoluut verschil. Het voordeel van deze manier is, is de simpliciteit. Echter kan dit bij een grote dataset veel tijd in beslag nemen. Zoals wij eerder hebben aangegeven, is de dataset verkregen door de medische faculteit van de VU. Deze hebben geprobeerd om extreme waarden in de data op te sporen. De volgende punten konden verwijderd worden van de dataset van beide proefpersonen:

<i>Proefpersoon 1</i>	<i>Proefpersoon 2</i>
201	136
202	137
	227
	271
	272
	273
	274

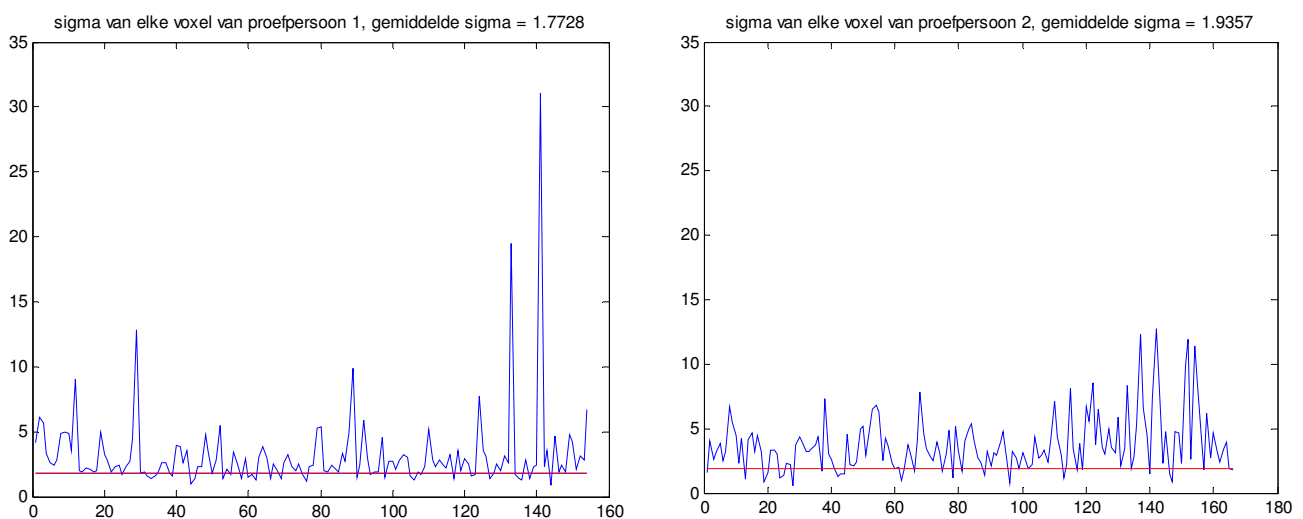
Tabel 4.3: Punten die verwijderd moeten worden van de data van beide proefpersonen.

4.4 Simulatie

In het simulatieonderzoek hebben wij een aantal verschillende situaties onderzocht. Alvorens wij dit beschrijven, hebben wij de parameters in stap 0 bepaald. In de vorige paragraaf hebben wij voor elke voxel van beide proefpersonen de geschatte waarde voor β_1 bepaald. Dit bleek te liggen in het interval $[-1, 2]$. In onze simulaties hebben wij voor b ook waarden gekozen uit dit interval. De waarden voor σ en σ_x zijn bepaald met de vergelijkingen (3.20) en (3.27). In het geval van fMRI-ruis hebben wij ε van elke voxel bepaald met de volgende vergelijking:

$$\varepsilon = y - (x_1 \hat{\beta}_1 + X_2 \hat{\beta}_2) \quad (4.10)$$

In het geval Gaussische ruis hebben wij ruis gegenereerd uit de normale verdeling zoals beschreven in de bovenste regel van vergelijking (4.3). In de onderstaande figuur 4.9 en tabel 4.4 is het resultaat te zien van het bepalen van de waarden van σ en σ_x beide proefpersonen:



Figuur 4.9: Variantie van de ruis van elke voxel van beide proefpersonen.

In figuur 4.9 is te zien dat beide ruisvectoren verschillend zijn. De spreiding van de ruis van de tweede proefpersoon lijkt groter, maar de ruisvector van proefpersoon 1 bevat een aantal grote (extreme) waarden. De waarde voor het gemiddelde σ is echter niet echt verschillend.

	<i>Proefpersoon 1</i>	<i>Proefpersoon 2</i>
σ	1.7728	1.9357
σ_x	0.6637	0.6894

Tabel 4.4: Waarden voor de parameters van stap 0.

Aantal simulaties (B in stap 4)

Alvorens wij begonnen zijn met het simulatieonderzoek, hebben wij eerst onderzocht wat een goede waarde voor het aantal simulaties is. Wij hebben voor een aantal situaties het onderscheidend vermogen van vergelijking (4.4) uitgerekend. Hieronder volgt een voorbeeld.

De volgende parameters⁷ zijn gekozen:

<i>Proefpersoon</i>		1
<i>b</i>		0.35
<i>σ</i>		1.7728
<i>σ_x</i>		0.6637
<i>α</i>		0.05
<i>γ</i>		0.4
<i>n</i>		232
<i>B</i>		[50, 100, 150, 500, 1000, 5000, 10000, 50000, 100000, 1000000]

Tabel 4.5: Gekozen parameters voor het bepalen van de juiste waarden van B.

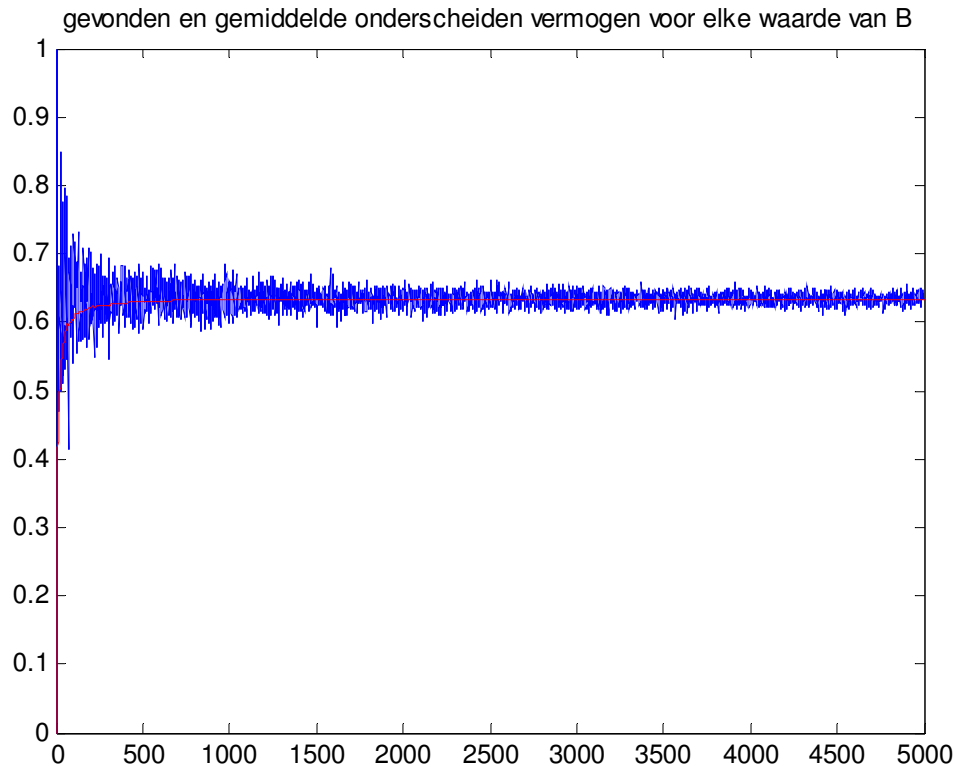
In de onderstaande figuur 4.10 is het resultaat van de simulaties te zien:



Figuur 4.10: Het gevonden onderscheidend vermogen in de verschillende simulaties, waarbij x = 1 ≡ B = 50, x = 2 ≡ B = 100, x = 3 ≡ B = 150, etc.

⁷ Het aantal scans is gecorrigeerd met 10%. Dit aantal was gelijk aan 210. Deze correctie wordt later besproken.

In figuur 4.10 zien wij dat het onderscheidend vermogen na punt 5-6 ($B = 1000-5000$) lijkt te convergeren. Wij vermoeden dat het uitvoeren van het aantal simulaties in dit interval genoeg is. Vervolgens hebben wij het aantal simulaties gevarieerd van 1 tot en met 5000. In de onderstaande figuur 4.11 is het gevonden en gemiddelde onderscheidend vermogen afgebeeld:



Figuur 4.11: Gevonden (blauw) en gemiddelde (rood) onderscheidend vermogen voor B gelijk aan 1 tot en met 5000.

Zoals wij kunnen zien in figuur 4.11 convergeert het gemiddelde al snel naar 1 specifieke waarde. Wij concluderen dat elke waarde voor B die tussen 1000 en 5000 (of hoger) ligt genoeg is voor betrouwbare resultaten.

Simulatie met Gaussische ruis

In de simulaties met de Gaussische ruis, hebben wij ruis gegeneerd door waarden uit de normale verdeling te trekken zoals beschreven in vergelijking (4.3). Vervolgens hebben wij dit voor een aantal parameters, zoals b , a en $1-\gamma$, gevarieerd. Dit hebben wij voor beide proefpersonen gedaan. In de onderstaande tabel 4.6 zijn de belangrijkste resultaten weergegeven:

proefpersoon	1	1	2	2
$(1-\gamma)\alpha$	0.05	0.0001	0.05	0.0001
0.6	0.4538	0.4504	0.4549	0.2878
0.7	0.6007	0.5648	0.5527	0.4157
0.8	0.7261	0.7292	0.6501	0.5859
0.9	0.8648	0.8643	0.8477	0.7825
0.95	0.9367	0.9356	0.914	0.8903

Tabel 4.6: Gesimuleerde waarden voor het onderscheidend vermogen met Gaussische ruis.

Er zijn een aantal zaken die ons zijn opgevallen in de simulaties met Gaussische ruis. Het minimale aantal scans verschilt per proefpersoon. Dit resulteert blijkbaar ook in een verschil in gesimuleerde onderscheidend vermogen. Het verschil in onderscheidend vermogen is te verwaarlozen voor proefpersoon 1 bij een kleinere betrouwbaarheidsdrempel, terwijl dit verschil toeneemt bij proefpersoon 2. Wij merken ook op dat het verschil in gesimuleerde onderscheidend vermogen en het ingestelde onderscheidend vermogen afneemt bij een grotere ingestelde onderscheidend vermogen. Tot slot concluderen wij dat het onderscheidend vermogen in onze simulaties kleiner zijn dan het ingestelde onderscheidend vermogen.

Simulatie met fMRI-ruis

In de simulaties met fMRI-ruis, hebben wij willekeurig residuvectoren geselecteerd als ruis. Zoals te zien in tabel 4.1 zijn er 154 voxels in de dataset van proefpersoon 1 en 166 voxels in de dataset van proefpersoon 2. Dit is dan ook het aantal residuvectoren van beide proefpersonen. Kortom het is niet zinvol om het aantal simulaties groter te laten zijn dan het aantal residuvectoren. Dit voorkomt dat dezelfde residuvector meerdere malen geselecteerd wordt in de simulatie. Wij hebben B gelijkgenomen aan 150. Het resultaat is te zien in de onderstaande tabel 4.7:

<i>proefpersoon</i>	<i>1</i>	<i>1</i>	<i>2</i>	<i>2</i>
$(1-\gamma)\alpha$	0.05	0.0001	0.05	0.0001
0.6	0.36	0.64	0.4467	0.2533
0.7	0.7333	0.4533	0.4533	0.4733
0.8	0.56	0.7	0.4933	0.5
0.9	0.9267	0.8	0.6467	0.7533
0.95	0.9667	0.9133	0.8933	0.86

Tabel 4.7: Gesimuleerde waarden voor het onderscheidend vermogen met fMRI ruis.

Ook in dit geval kunnen wij dezelfde conclusies trekken als bij de simulaties met Gaussische ruis. Het onderscheidend vermogen is over het algemeen beter bij proefpersoon 1 dan bij 2. Het gesimuleerde onderscheidend vermogen is bij een kleinere γ beter. Echter zijn er wel nu wel verschillen in het gesimuleerde onderscheidend vermogen in het geval van andere onbetrouwbaarheidsdrempels. Ook zien wij een vreemde situatie in de middelste rij van tabel 4.7: het gesimuleerde onderscheidend vermogen is vrij laag. Dit kan liggen aan het feit dat aantal simulaties te laag is ($B = 150$) voor betrouwbare resultaten.

Correctie op het minimale aantal scans

In de voorgaande simulaties is ons opgevallen dat het gesimuleerde onderscheidend vermogen lager is dan wat wij zouden verwachten. Zoals wij eerder hebben aangegeven, is een correctie op het aantal scans verstandig, omdat in [28] het verschil in benaderde en exacte aantal scans oploopt van 1 tot 10%. Wij hebben besloten om een correctie uit te voeren op het aantal scans door er 10% bij op te tellen. In de onderstaande tabellen is het resultaat te zien van de simulaties met het gecorrigeerde aantal:

Gauss					fMRI				
<i>proefpersoon</i>	<i>1</i>	<i>1</i>	<i>2</i>	<i>2</i>	<i>proefpersoon</i>	<i>1</i>	<i>1</i>	<i>2</i>	<i>2</i>
$(1-\gamma)\alpha$	0.05	0.0001	0.05	0.0001	$(1-\gamma)\alpha$	0.05	0.0001	0.05	0.0001
0.6	0.8848	0.5774	0.79	0.5202	0.6	0.94	0.6067	0.7333	0.3867
0.7	0.893	0.7204	0.7832	0.6332	0.7	0.9667	0.7867	0.7933	0.5
0.8	0.8748	0.8264	0.7982	0.7626	0.8	0.94	0.7333	0.7267	0.7933
0.9	0.8822	0.9236	0.799	0.9236	0.9	0.9267	0.9333	0.7133	0.9467
0.95	0.9408	0.9558	0.8964	0.9184	0.95	0.92	0.9667	0.72	0.9333

Tabel 4.8: Gesimuleerde waarden van het onderscheidend vermogen met gecorrigeerd aantal.

In de bovenstaande tabellen kunnen wij zien dat er een toename is in onderscheidend vermogen. Vooral voor kleinere γ is de toename erg groot. Echter merken wij op dat voor de onbetrouwbaarheidsdrempel van 0.0001 en γ groter dan 0.2 het gesimuleerde onderscheidend vermogen nog steeds tegen valt. Over het algemeen zien wij een veel betere gesimuleerde onderscheidend vermogen.

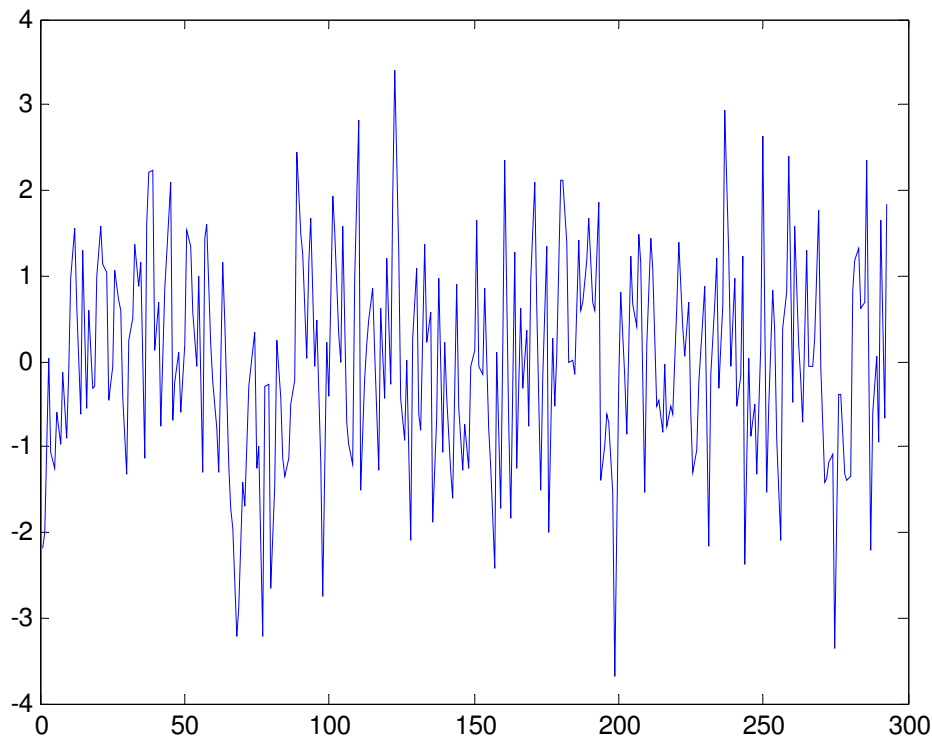
Willekeurig startpunt in de verklarende variabele

In de simulaties hebben wij opgemerkt dat het gesimuleerde onderscheidend vermogen hoger blijkt te zijn hoe meer punten van de designmatrix wij van het eind gebruiken dan punten van het begin. Stap 2 van de simulatie veranderen wij als volgt:

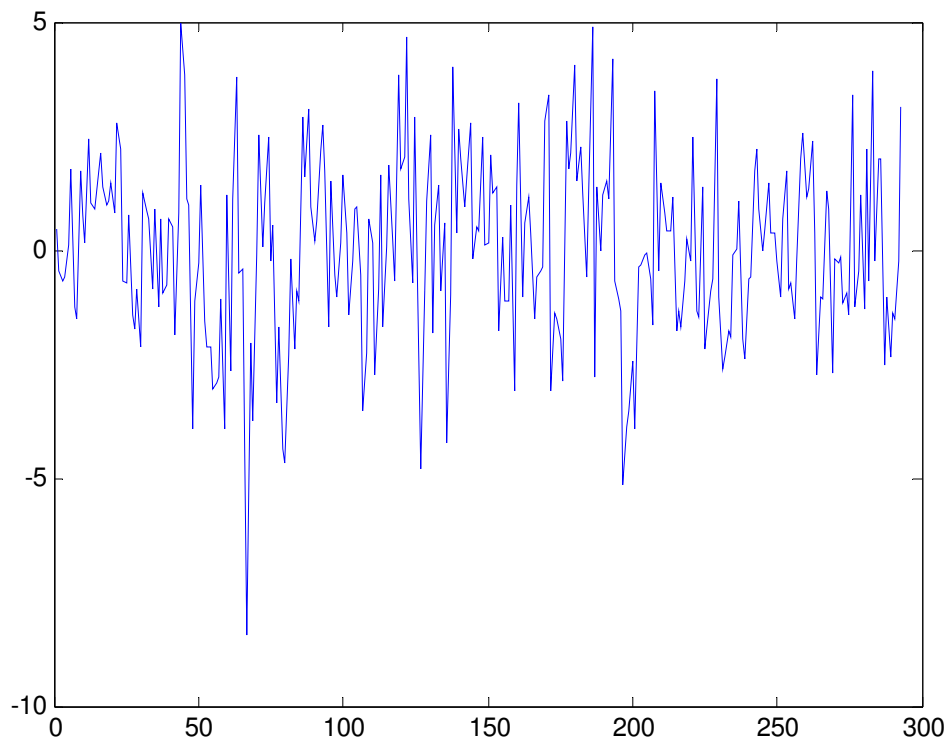
$$\tilde{y}_{a,sim}, \dots, \tilde{y}_{a+n,sim} \quad \tilde{y}_{i,sim} = b \cdot \tilde{x}_i + P_2^\perp \cdot \text{ruis} \quad 1 \leq i \leq (n_x + 1 - n) \quad (4.1)$$

Het gesimuleerde onderscheidend vermogen van de rechte tabel van tabel 4.8 met onbetrouwbaarheidsdrempel gelijk aan 0.0001 en $1-\gamma$ gelijk aan 0.7 neemt toe van 0.5 naar groter dan 0.8.

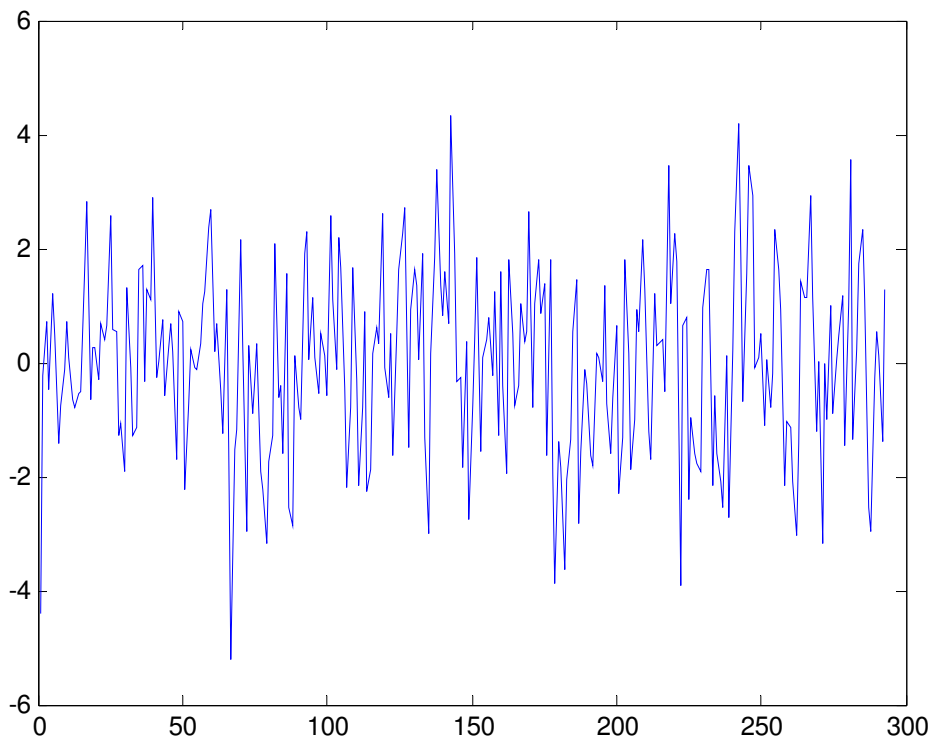
Ons vermoeden is dan ook dat de residuen in absolute waarden groter zijn in het begin dan op het eind. Allereerst plotten wij een aantal residu vectoren om te zien of wij dit beeld te zien krijgen. Wij beschikken over 293 residu vectoren in het geval van proefpersoon 2. Het is niet zinvol om alle vectoren te bekijken, dus kiezen wij een aantal vectoren willekeurig. Vervolgens bekijken wij ook het gemiddelde van alle vectoren. In onderstaande figuur is het resultaat te zien:



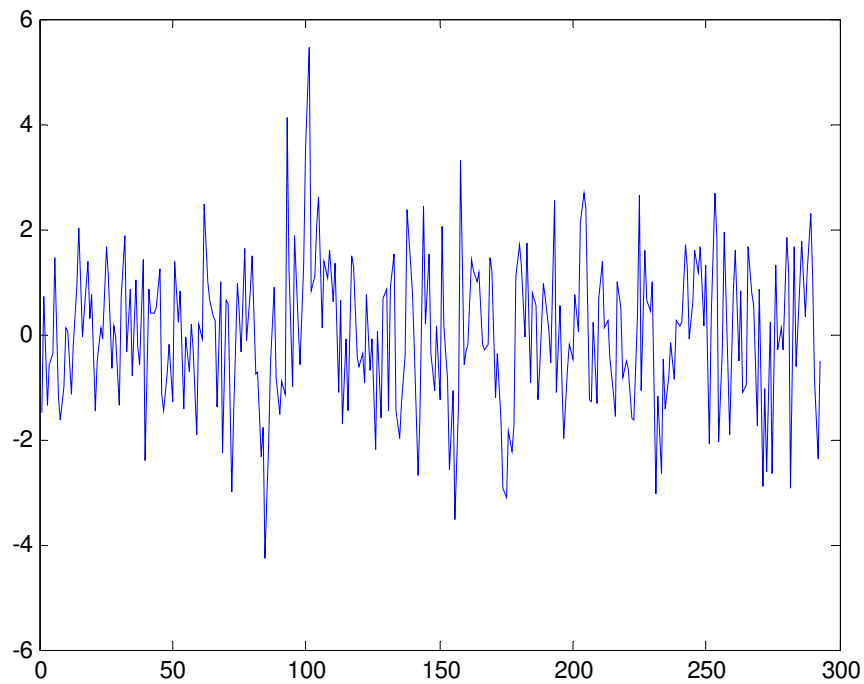
Figuur 4.12: Residuen van het meervoudig regressiemodel in voxel 1 van proefpersoon 2.



Figuur 4.13: Residuen van het meervoudig regressiemodel in voxel 2 van proefpersoon 2.

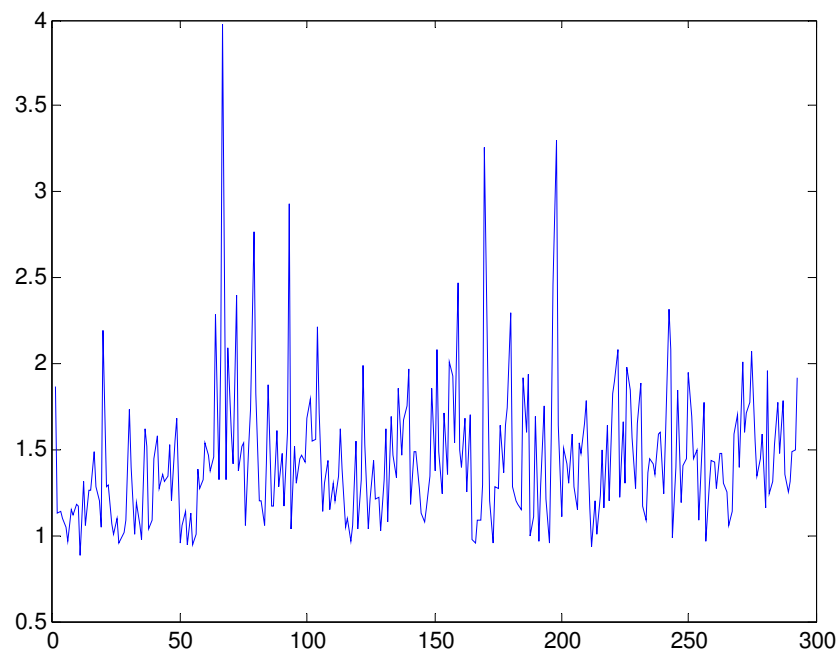


Figuur 4.14: Residuen van het meervoudig regressiemodel in voxel 40 van proefpersoon 2.



Figuur 4.15: Residuen van het meervoudig regressiemodel in voxel 90 van proefpersoon 2.

Als wij de bovenstaande figuren in ogenschouw nemen, dan kunnen wij niet concluderen dat de residuen groter zijn in absolute waarden in het begin vergeleken met de residuen op het eind. Echter valt ons op dat aan een aantal punten met de grootste absolute waarden in het interval 0 en 100 liggen. Dit beeld zien wij ook in de onderstaande figuur waarin de gemiddelde absolute residuen zijn geplott:



Figuur 4.16: Gemiddelde absolute residuen van het meervoudig regressiemodel.

5. Conclusie

5.1 Conclusie

Het doel van dit onderzoek was het beantwoorden van de vraag of de formule van vergelijking (3.26) correct is. Met deze formule is het mogelijk om het minimale aantal fMRI-scans te bepalen om een signaal te detecteren bij een bepaalde kans. Deze formule is afgeleid in [28] door middel van een power analyse. In de afleiding worden een aantal benaderingen en aannamen gemaakt. Hierdoor ontstaat de vraag of deze formule werkt in de praktijk.

Met behulp van echte fMRI-data van twee proefpersonen hebben wij een lineair regressiemodel gebouwd. Vervolgens hebben wij een aantal parameters bepaald om fMRI data te kunnen simuleren in het mathematische softwarepakket Matlab. Met behulp van de simulaties hebben wij het onderscheidend vermogen uitgerekend en deze vergeleken met het gebruikte onderscheidend vermogen.

Allereerst hebben wij onderzocht hoe lang wij moesten simuleren om betrouwbare resultaten te krijgen. Het aantal simulaties hebben wij gevarieerd tussen 10 en 1 miljoen keer. Het is gebleken dat het kiezen van een aantal tussen 1000 en 5000 betrouwbare resultaten oplevert (figuur 4.11). Dit correspondeert in een rekentijd van ongeveer 5-6 seconden.

In [28] werd al opgemerkt dat het minimale aantal scans onderschat wordt met ongeveer 1 tot 10%. Dit zagen wij ook terug in de simulaties. Het gesimuleerde onderscheidend vermogen was is veel gevallen lager dan verwacht. In sommige gevallen was het verschil aanzienlijk. Wij hebben daarom besloten om een correctie uit te voeren op het aantal scans door hier 10% bij op te tellen. Dit resulteerde in veel betere gesimuleerde onderscheidend vermogen.

Verder hebben wij onderzocht of er een verschil is tussen Gaussische en fMRI-ruis. In de simulaties is gebleken dat het onderscheidend vermogen over het algemeen beter was in het geval van Gaussische ruis. Uit een klein onderzoek in de verdeling van de fMRI-ruis is gebleken dat niet alle fMRI-ruisvectoren normaal verdeeld zijn. Dit is een aanname die wij gemaakt hebben bij het modelleren van het lineaire regressiemodel.

Ook vonden wij een verschil in het gesimuleerde onderscheidend vermogen tussen de proefpersonen. Omdat niet elke proefpersoon dezelfde ruis heeft, is het aantal scans meestal verschillend. Zo kan het zijn dat er meer scans nodig zijn om een signaal te kunnen detecteren. Echter blijkt er ook een verschil te zijn in gesimuleerde onderscheidend vermogen. Over het algemeen was dit hoger voor proefpersoon 1 dan bij proefpersoon 2.

Wij hebben ook een andere aanpak geprobeerd in de simulatie. Namelijk het kiezen van een willekeurig startpunt in de designmatrix bij het simuleren van een signaal. Als wij in plaats van de eerste n punten uit de designmatrix, de laatste n punten selecteren, dan stijgt het gesimuleerde onderscheidend vermogen aanzienlijk en is deze groter dan verwacht.

Wij concluderen dat de formule om het minimale aantal scans te bepalen goed werkt mits er een aantal correcties worden uitgevoerd. Het aantal wordt onderschat en geadviseerd wordt om dit te verhogen met ongeveer 10%. Ook is er sprake van meer ruis in het begin van het experiment en hier zou dus rekening mee gehouden moeten worden.

5.2 Discussie

Uit het simulatieonderzoek is gebleken dat er verschillen zijn tussen proefpersonen in ruis en onderscheidend vermogen. Om deze verschillen te onderzoeken om deze bijvoorbeeld te verklaren, zou meer data nodig zijn. In dit BWI-Werkstuk is gebruik gemaakt van data van twee proefpersonen. Dit maakte het onmogelijk om deze verschillen te onderzoeken en de formule voor het bepalen van het minimale fMRI-scans in het algemeen. Echter constateren wij ook dat data duur is, dus het verzamelen van de juiste hoeveelheid data kan problematisch zijn.

Een ander aspect wat onderzocht kan worden, is de collineariteit van de confounders en het effect hiervan op het lineaire regressiemodel bij het weglaten van de collineaire confounders. Het is op dit moment niet zeker welke confounders weggelaten kunnen worden en welke juist niet. Bovendien heeft het weglaten van confounders effect op een aantal andere parameters zoals de geschatte coëfficiënten en de ruisvectoren van het lineaire regressiemodel. Dit beïnvloedt de simulaties ook.

Tot slot is in dit onderzoek gekozen voor een simpele simulatie aanpak. Het nadeel van een dergelijke aanpak is dat sommige aspecten vereenvoudigd worden. Dit gaat ten koste van het doel van de simulatie, namelijk het creëren van een zo echt mogelijke situatie. Ook het simuleren van de fMRI-ruis kan op een andere manier worden gerealiseerd. In dit onderzoek is gekozen om willekeurige ruisvectoren te selecteren om de afhankelijkheid van de meetfouten in de tijd te behouden. Een ander manier om dit te simuleren zou kunnen onderzocht worden de verdeling van de fMRI-ruis grondig te onderzoeken.

Literatuurlijst

- [1] Purcell, E. M., Torrey, H. C., Pound, R. V., 1946, 'Resonance absorption by nuclear magnetic moments in a solid'. *Physical Review* 69, 37–38.
- [2] Bloch, F., Hansen, W. W., Packard, 1946, 'M. Nuclear induction'. *Physical Review* 69, 127.
- [3] http://nobelprize.org/nobel_prizes/physics/laureates/1952/index.html
- [4] Hornak, J.P., 1997, *The Basics of NMR*. (<http://www.cis.rit.edu/htbooks/nmr>)
- [5] Bink, E.J., 2004, *MRI Basis Principles; Voor iedereen die geen graad in fysica heeft*. (<http://www.mri-physics.com/>)
- [6] Damadian, R.V., 1971, 'Tumor Detection by Nuclear Magnetic Resonance'. *Science*, 171, 1151-1153.
- [7] <http://www.fonar.com/>
- [8] Montgomery, D., 2003, 'In a Funk Over the No-Nobel Prize; Overlooked MRI Pioneer Lobbies Against Decision'. *The Washington Post*, Washington, C.01. (<http://www.washingtonpost.com/ac2/wp-dyn/A6083-2003Oct9?language=printer>)
- [9] http://nobelprize.org/nobel_prizes/medicine/laureates/2003/
- [10] Lauterbur, P.C., 1973, 'Image Formation by Induced Local Interaction: Examples Employing Nuclear Magnetic Resonance'. *Nature* 242, 190–191.
- [11] O Leach, M., 2004, 'Nobel Prize in Physiology or Medicine 2003 awarded to Paul Lauterbur and Peter Mansfield for discoveries concerning magnetic resonance imaging'. *Physics in Medicine and Biology* 49.
- [12] Mansfield, P., Grannell, P.K., 1973, 'NMR 'diffraction' in solids?'. *Journal of Physics C: Solid State Physics* 6, L422-L426.
- [13] Mansfield, P., Garroway, A.N., Grannell P.K., 1974, 'Image Formation using Nuclear Magnetic Resonance'. *Journal of Physics C: Solid State Physics* 7, L457.
- [14] Smidts, A., Rossiter, J.R., 2002, 'Wat onthoudt een consument van een tv-commercial? Een kijkje in het brein met neuro-imaging technieken'. *Jaarboek MarktOnderzoekAssociatie*, 145-155.
- [15] <http://www.fmrib.ox.ac.uk/education/fmri/introduction-to-fmri/>
- [16] Savoy, R.L., 2001, 'History and future directions of human brain mapping and functional neuroimaging'. *Acta Psychologica* 107, 9-42.
- [17] Ogawa S., Lee T.M., Kay A.R., Tank D.W., 1990, 'Brain Magnetic Resonance Imaging with Contrast Dependent on Blood Oxygenation'. *Proceedings of the National Academy of Sciences* 87, 9868-9872.
- [18] Logothetis, N.K., 2002, 'The neural basis of the blood-oxygen-level-dependent functional magnetic resonance imaging signal'. *Philosophical Transactions of The Royal Society* 357, 1003-1037.

- [19] Noll, D.C., 2001, 'A Primer on MRI and Functional MRI'. (<http://www.eecs.umich.edu/~dnoll/>)
- [20] Kok, A., 2004, *Het hiërarchisch brein: inleiding tot de cognitieve neurowetenschap*. Van Gorcum, 28-31.
- [21] Chee, M.W.L., Venkatraman, V., Westphal, C., Siong, S.C., 2003, 'Comparison of Block and Event-Related fMRI Designs in Evaluating the Word-Frequency Effect'. *Human Brain Mapping* 18, 186-193.
- [22] de Haan, B., Rorden, C., 'fMRI Guide'. (http://www.sph.sc.edu/comd/rorden/fmri_guide/index.html)
- [23] De Gunst, M.C.M, Van der Vaart, A.W., 2006, *Statistische Data Analyse*, Vrije Universiteit, Amsterdam.
- [24] De Gunst, M.C.M, 2006, *Statistical Models*, Vrije Universiteit, Amsterdam.
- [25] Worsley, K.J., Liao, C.H., Aston, J., Petre, V., Duncan, G.H., Morales, F., Evans, A.C., 2002, A general statistical analysis for fMRI data. *NeuroImage* 15, 1-15.
- [26] Kleijn, B., Van der Vaart, A.W., 2005, *Algemene Statistiek*, Vrije Universiteit, Amsterdam.
- [27] Murphy, K., Bodurka, J., Bandetti, P.A., 2007, How long to scan? The relationship between fMRI temporal signal to noise ratio and necessary scan duration. *NeuroImage* 34, 565-574.
- [28] Bijma, F., Gonçalves, S.I., Jonker, M.A., De Munck, J.C., 2008, Study on power and false negatives in fMRI. *IEEE Signal Processing Special Issue on fMRI Analysis for Human Brain Mapping*.

Bijlage 1: Code

Het inlezen van de data van beide proefpersonen:

```
function [Y, x1, X2, RX] = inlezenDataSubject1(stringVar1, stringVar2)

% inlezen van data
[a, b, c, d, e, f, g, h, i, j, k, l, m, n, o, p]=textread(stringVar1, '%n
%n %n %n %n %n %n %n %n %n %n %n %n %n %n %n');
X = [a, b, c, d, e, f, g, h, i, j, k, l, m, n, o, p];

% maken van design matrix
x1 = X(:,5);
X2 = [X(:,3:4) X(:,6:12)];
X2 = [ones(size(X2,1),1) X2];
RX = [x1 X2];

% signaal Y inlezen
Y = textread(stringVar2);
Y = Y(:,2:155);

% (outliers) case 201 en 202 verwijderen
Y = [Y(1:200,:);Y(203:300,:)];
x1 = [x1(1:200,:);x1(203:300,:)];
X2 = [X2(1:200,:);X2(203:300,:)];
RX = [RX(1:200,:);RX(203:300,:)];
return

function [Y, x1, X2, RX] = inlezenDataSubject2(stringVar1, stringVar2)

% inlezen van data
[a, b, c, d, e, f, g, h, i, j, k]=textread(stringVar1, '%n %n %n %n %n %n
%n %n %n %n %n');
X = [a, b, c, d, e, f, g, h, i, j, k];

% maken van design matrix
x1 = X(:,5);
X2 = [X(:,3:4) X(:,6:11)];
X2 = [ones(size(X2,1),1) X2];
RX = [x1 X2];

% signaal Y inlezen
Y=textread(stringVar2);
Y=Y(:,2:167);

% (outliers) case 136, 137, 227, 271, 272, 273 en 274 verwijderen
Y = [Y(1:135,:);Y(138:226,:);Y(228:270,:);Y(275:300,:)];
x1 = [x1(1:135,:);x1(138:226,:);x1(228:270,:);x1(275:300,:)];
X2 = [X2(1:135,:);X2(138:226,:);X2(228:270,:);X2(275:300,:)];
RX = [RX(1:135,:);RX(138:226,:);RX(228:270,:);RX(275:300,:)];
return
```

Parameters bepalen voor simulatie:

```
function [residuen, sigmas] = regressie(Y, RX)

% memory allocatie matrices
residuen = [];
sigmas = [];

% parameters
nY = size(Y,2);           % aantal voxels
nT = size(RX,1);         % aantal tijd samples per voxel
q = size(RX,2);          % aantal confounders + beta_1

% residuen, variantie van residuen en beta_1 bepalen
for i=1:nY
    beta = regress(Y(:,i),RX);
    residuen = [residuen (Y(:,i) - RX*beta)];
    sigma2 = sum((Y(:,i) - RX*beta).^2)/(nT-q);
    sigmas = [sigmas; sigma2];
end
return
```

Het aantal scans bepalen:

```
function [n, nOud] = aantalScans(b, x1, X2, RX, sigmas, alpha, gamma)

nT = size(RX,1);

%bepalen van sigma_x, sigma
x_slang = (eye(length(x1))-X2*inv(X2'*X2)*X2')*x1;
sigma_x = sqrt((x_slang'*x_slang)/nT);

sigma = sqrt(mean(sigmas));

z_alpha = norminv(1-alpha,0,1);
z_gamma = norminv(1-gamma,0,1);
n = (z_alpha+z_gamma)^2*(sigma/(b*sigma_x))^2;
n = ceil(n);
nOud = n;

% in het geval dat het aantal minimale scans groter dan het aantal
% timestamps is
if n > nT
    error(['Het aantal minimale scans (' num2str(ceil(n)) ') is groter dan '
    ' num2str(nT) '!']);
else
    n = max([1.1*n, 100]);
    n = ceil(n);
end

n = min([n, nT]);
return
```


Het onderscheidend vermogen bepalen in het geval van Gaussische ruis:

```
function power = functie_Gauss(b, n, X2, RX, sigma, B, alpha)

q = size(RX,2);
betas = b;
aantalSignificant = 0;

% Projectie matrix uitrekenen
P2 = eye(n)-X2(1:n,:)*inv(X2(1:n,:)'*X2(1:n,:))*X2(1:n,:);

% y_1,...,y_n simuleren
for i = 1:B
    Y_sim = P2*RX(1:n,1)*betas + P2*(sigma.*randn(n,1));

    beta = regress(Y_sim,P2*RX(1:n,1));
    sigma2 = sum((Y_sim - P2*RX(1:n,1)*beta(1)).^2)/(n-q);
    covBeta = sigma2*inv((P2*RX(1:n,1))'*(P2*RX(1:n,1)));

    % significantie van geschatte beta_1 toetsen
    t = beta(1)/sqrt(covBeta(1,1));
    p = 1-tcdf(t, (n-q));

    if p < alpha
        aantalSignificant = aantalSignificant + 1;
    end
end

power = aantalSignificant/B;
return
```

Het onderscheidend vermogen bepalen in het geval van fMRI-ruis:

```
function power = functie_fmri(b, n, X2, RX, residuen, B, alpha)

q = size(RX,2);
betas = b;
aantalSignificant = 0;

% Projectie matrix uitrekenen
P2 = eye(n)-X2(1:n,:) * inv(X2(1:n,:) ' * X2(1:n,:)) * X2(1:n,:)' ;

% y_1, ..., y_n simuleren
for i = 1:B

    % willekeurige residuvector selecteren
    r = 1 + (size(residuen,2)-1) .* rand(1,1);
    r = round(r);

    Y_sim = P2 * RX(1:n,1) * betas + P2 * (residuen(1:n,r));

    beta = regress(Y_sim, P2 * RX(1:n,1));
    sigma2 = sum((Y_sim - P2 * RX(1:n,1) * beta(1)).^2) / (n-q);
    covBeta = sigma2 * inv((P2 * RX(1:n,1))' * (P2 * RX(1:n,1)));

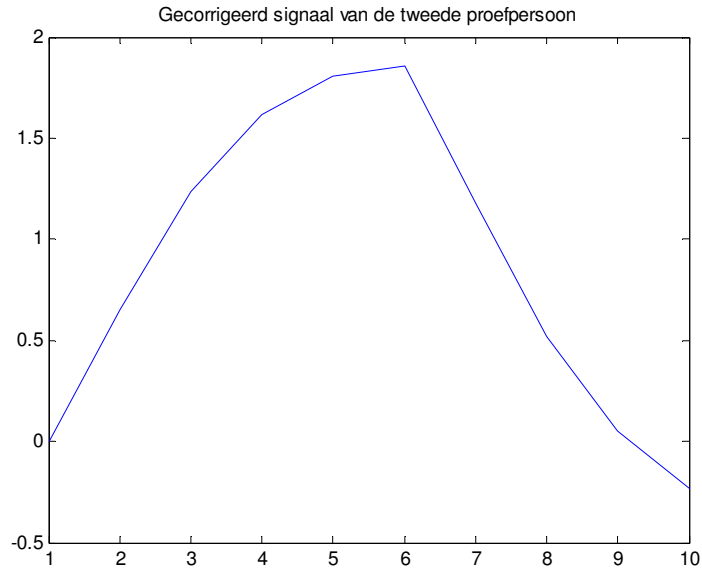
    % significantie van geschatte beta_1 toetsen
    t = beta(1) / sqrt(covBeta(1,1));
    p = 1 - tcdf(t, (n-q));

    if p < alpha
        aantalSignificant = aantalSignificant + 1;
    end
end

power = aantalSignificant / B;
return
```

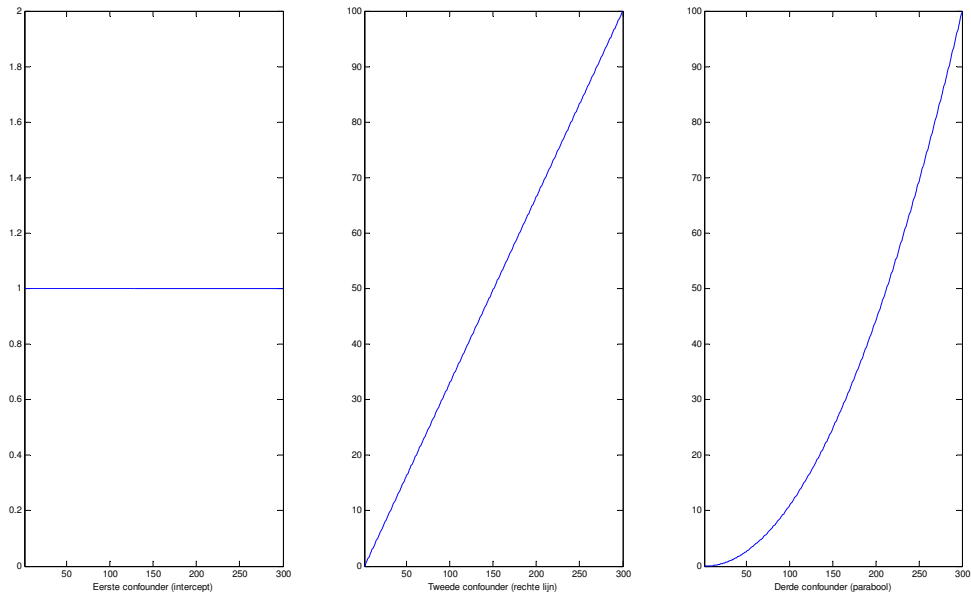
Bijlage 2: Tabellen & figuren van simulatieonderzoek

Het gecorrigeerde signaal van proefpersoon 2



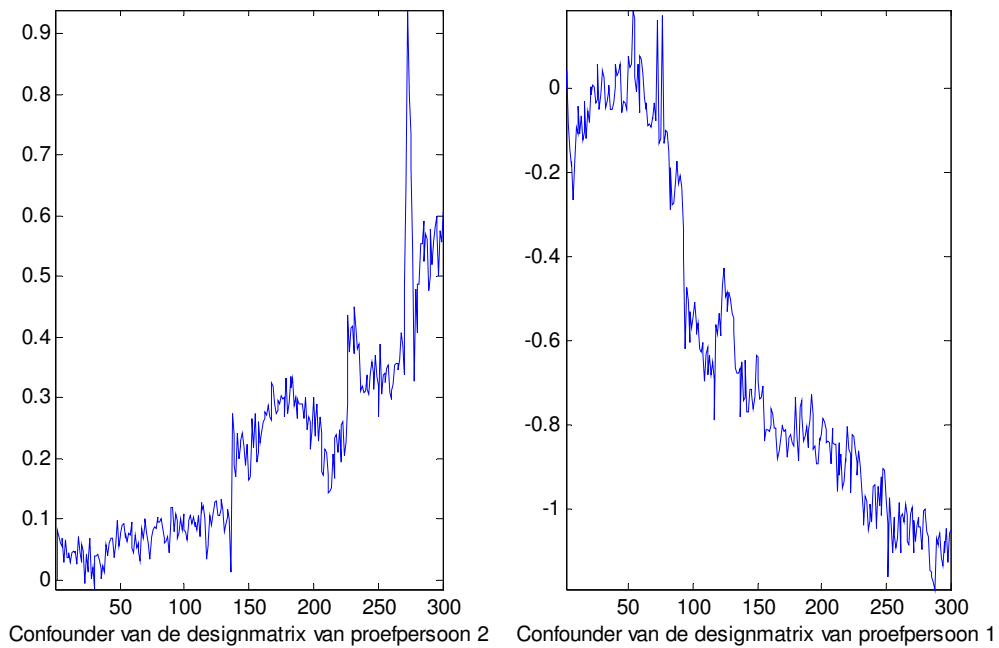
Figuur II.1: Voorbeeld van een gecorrigeerd signaal van de tweede proefpersoon.

De eerste drie confounders van beide proefpersonen



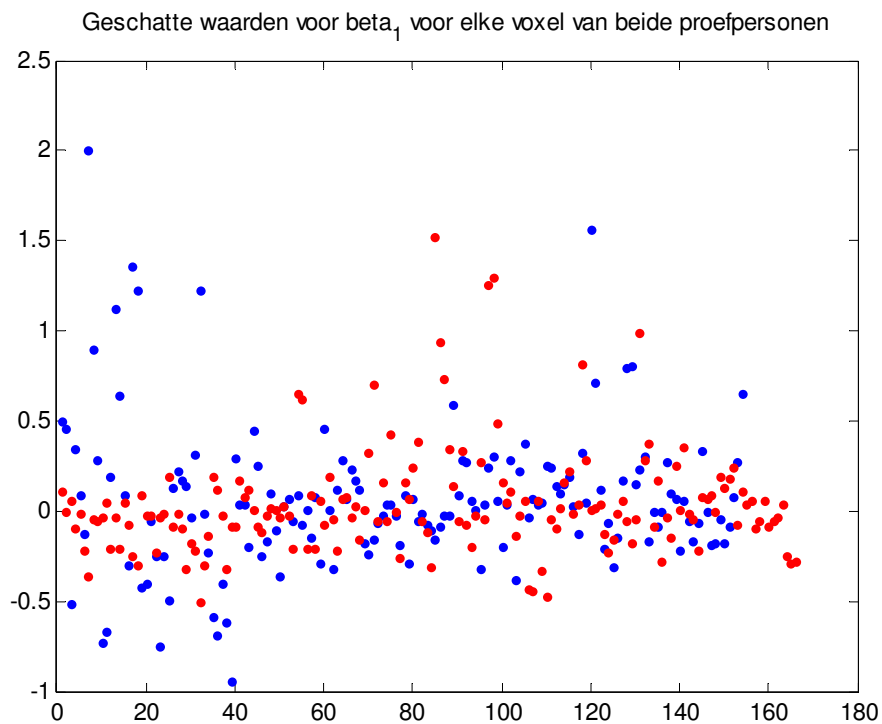
Figuur II.2: De eerste drie confounders van beide proefpersonen.

Voorbeeld van de overige confounder van beide proefpersonen



Figuur II.3: Voorbeeld van confounders van de designmatrix van beide proefpersonen.

Geschatte waarden voor de onbekende parameters van het lineair regressiemodel



Figuur II.4: Geschatte waarden voor β_1 voor elke voxel van proefpersoon 1 (blauw) en 2

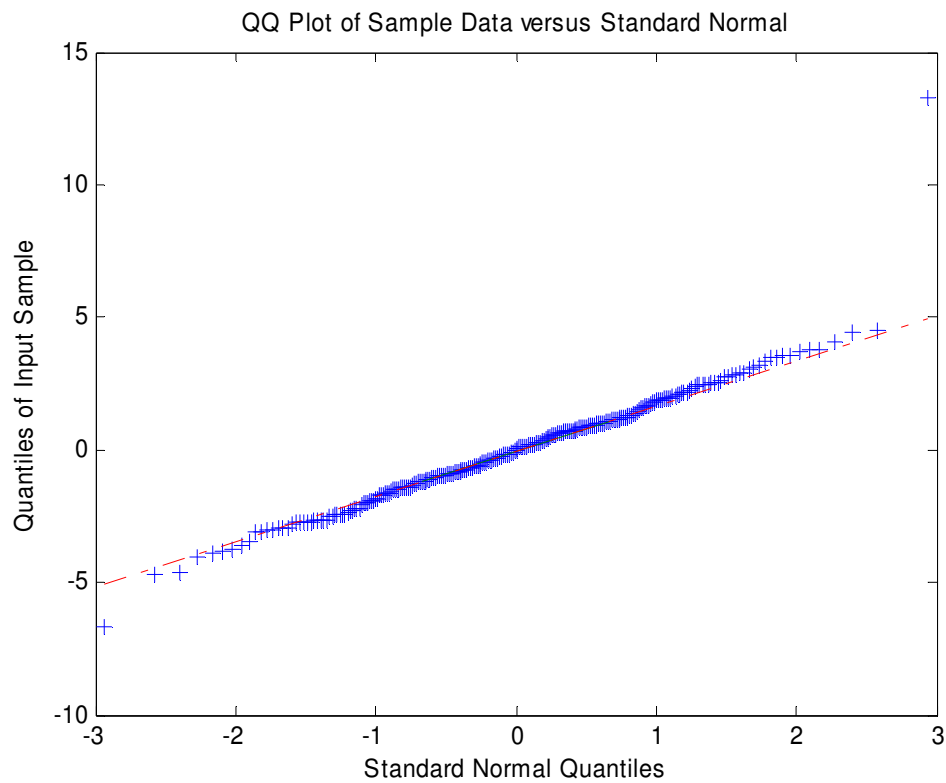
Conditie indices en variantiedecomposities van de geschatte parameters

$$\begin{pmatrix} \text{cond.ind.} & \text{prop}(\hat{\beta}_2) & \text{prop}(\hat{\beta}_3) \\ 87.898 & 0.97863 & 0.79096 \end{pmatrix}$$

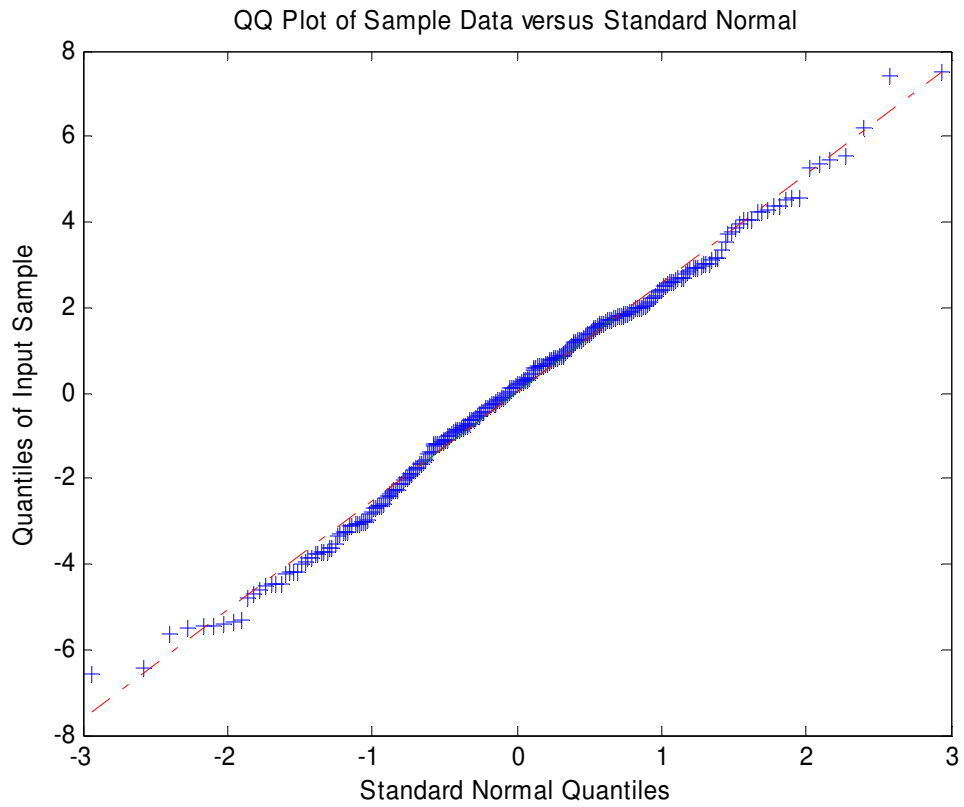
$$\begin{pmatrix} \text{cond.ind.} & \text{prop}(\hat{\beta}_0) & \text{prop}(\hat{\beta}_2) & \text{prop}(\hat{\beta}_3) \\ 58.946 & 0.85754 & 0.94907 & 0.9639 \end{pmatrix}$$

Figuur II.5: Conditie indices en variantiedecompositie van de geschatte parameters β_2 en β_3 voor proefpersoon 1 (boven) en β_1 , β_2 en β_3 voor proefpersoon 2 (onder).

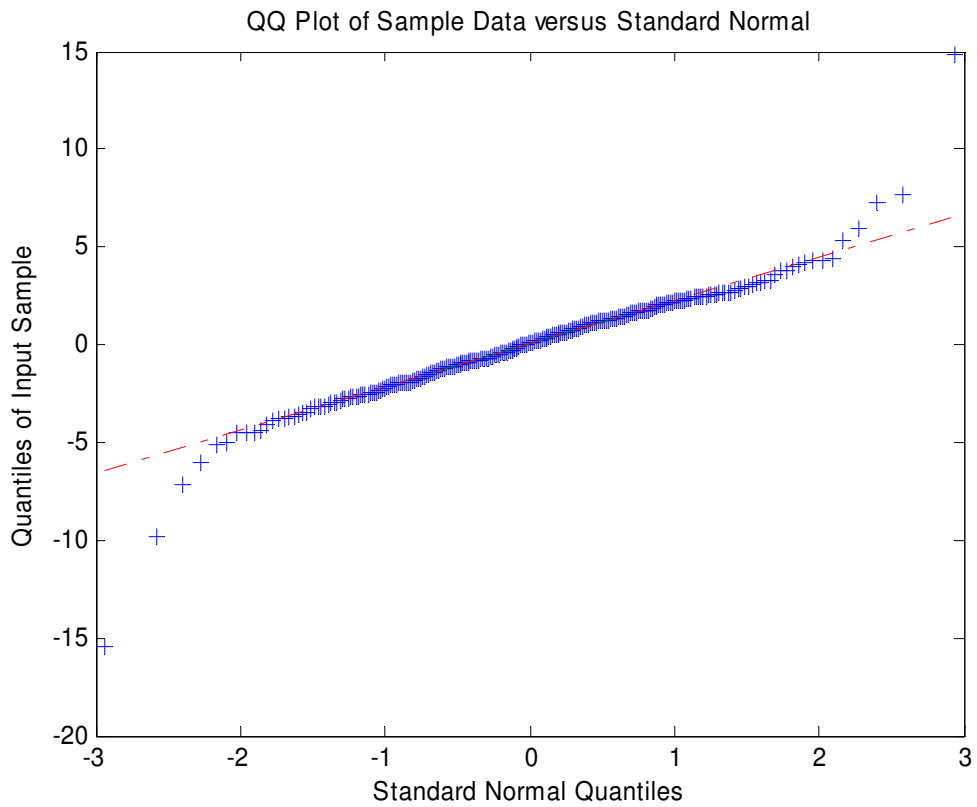
QQplots van de residuen van het lineaire regressiemodel



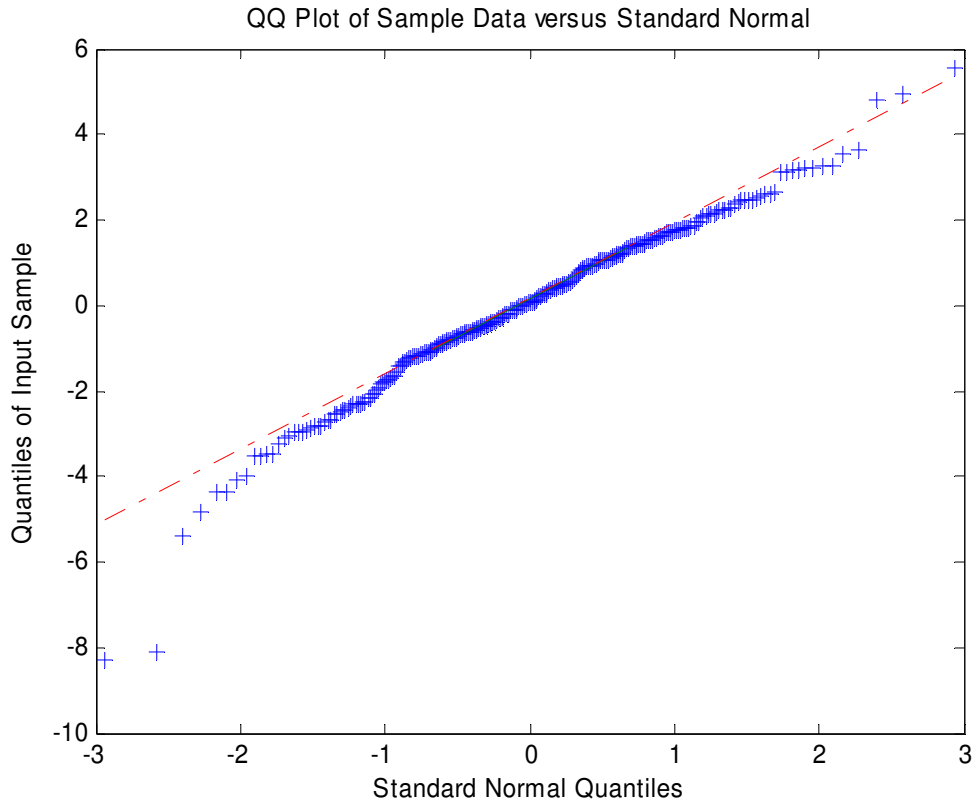
Figuur II.6: QQplot van de residuen van voxel 4 uitgezet tegen de normale verdeling.



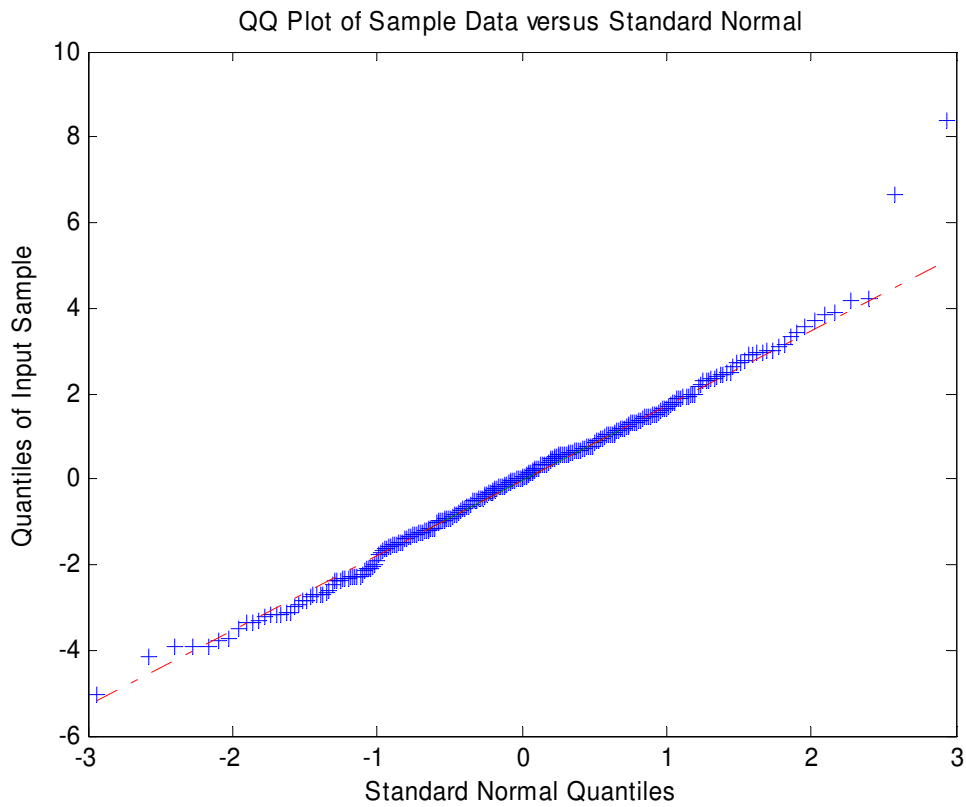
Figuur II.7: QQplot van de residuen van voxel 8 uitgezet tegen de normale verdeling.



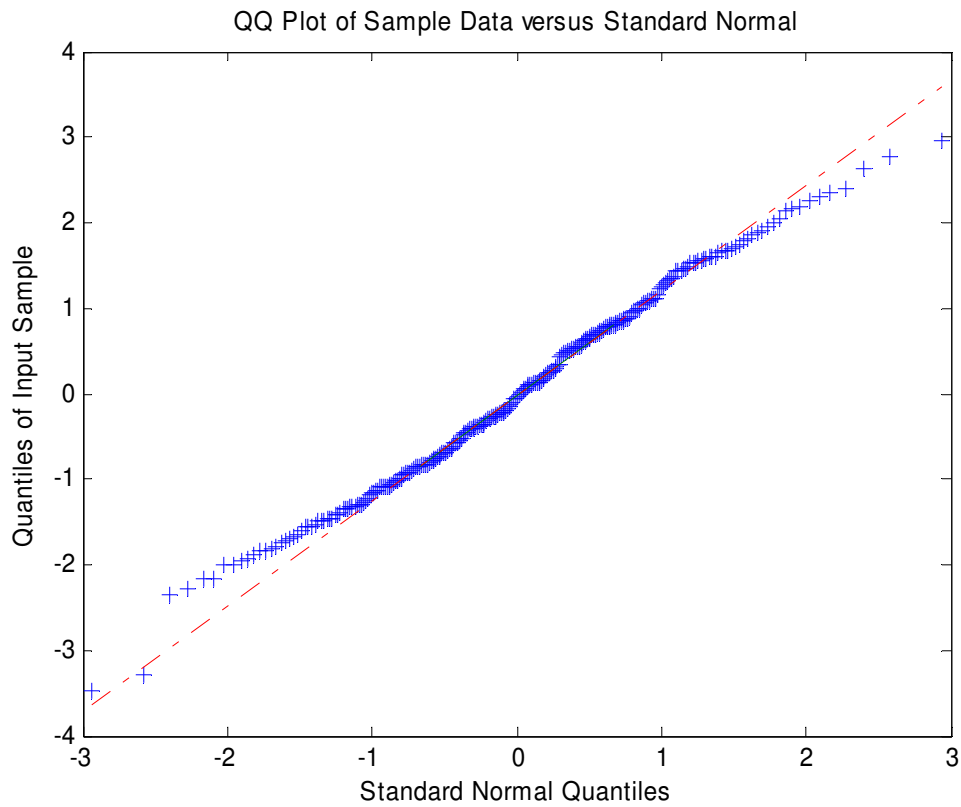
Figuur II.8: QQplot van de residuen van voxel 15 uitgezet tegen de normale verdeling.



Figuur II.9: QQplot van de residuen van voxel 16 uitgezet tegen de normale verdeling.

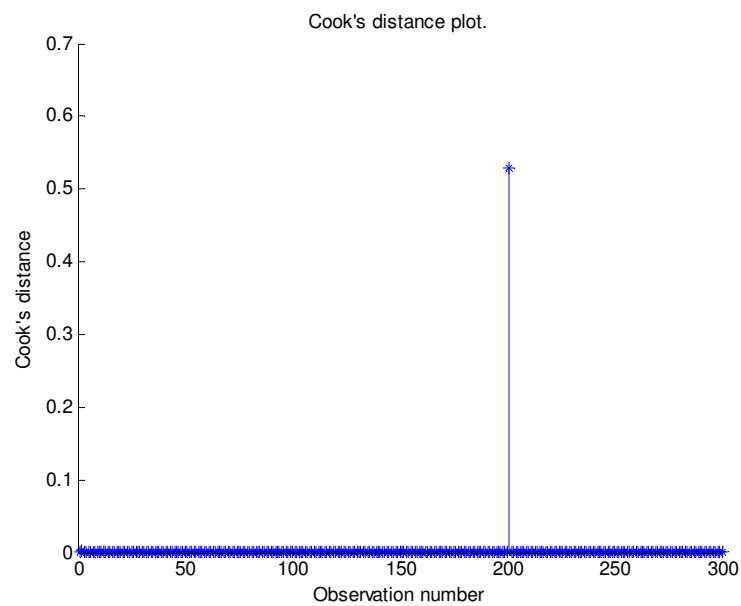


Figuur II.10: QQplot van de residuen van voxel 23 uitgezet tegen de normale verdeling.



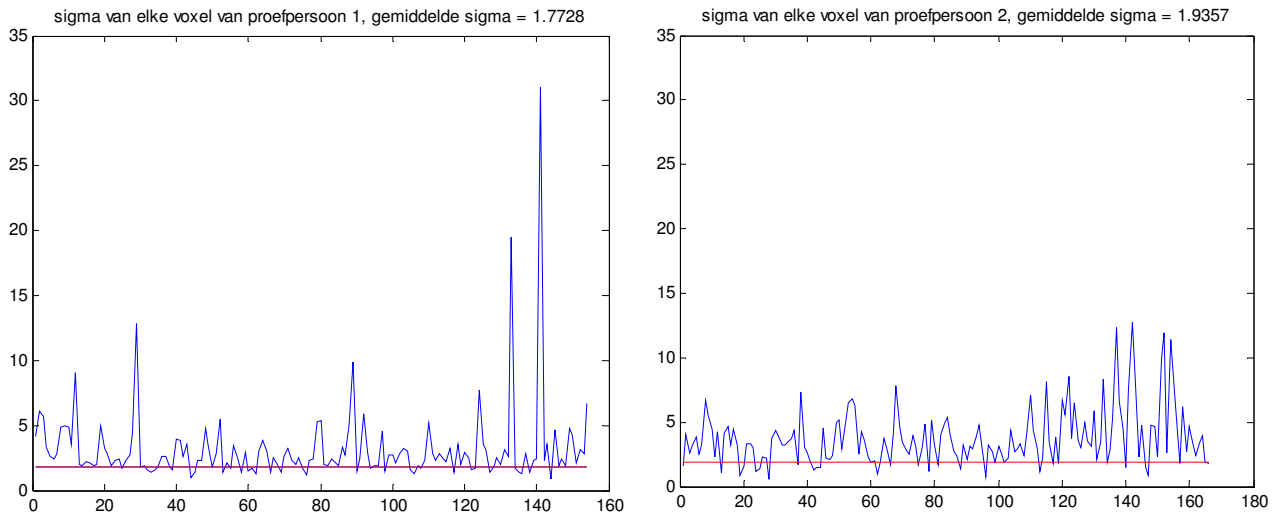
Figuur II.11: QQplot van de residuen van voxel 42 uitgezet tegen de normale verdeling.

Cook's afstanden van de punten in voxel 16 van proefpersoon 1



Figuur II.12: Cook's afstanden van de punten in het geval van voxel 16 van proefpersoon 1.

Variantie van de ruis van elke voxel van beide proefpersonen



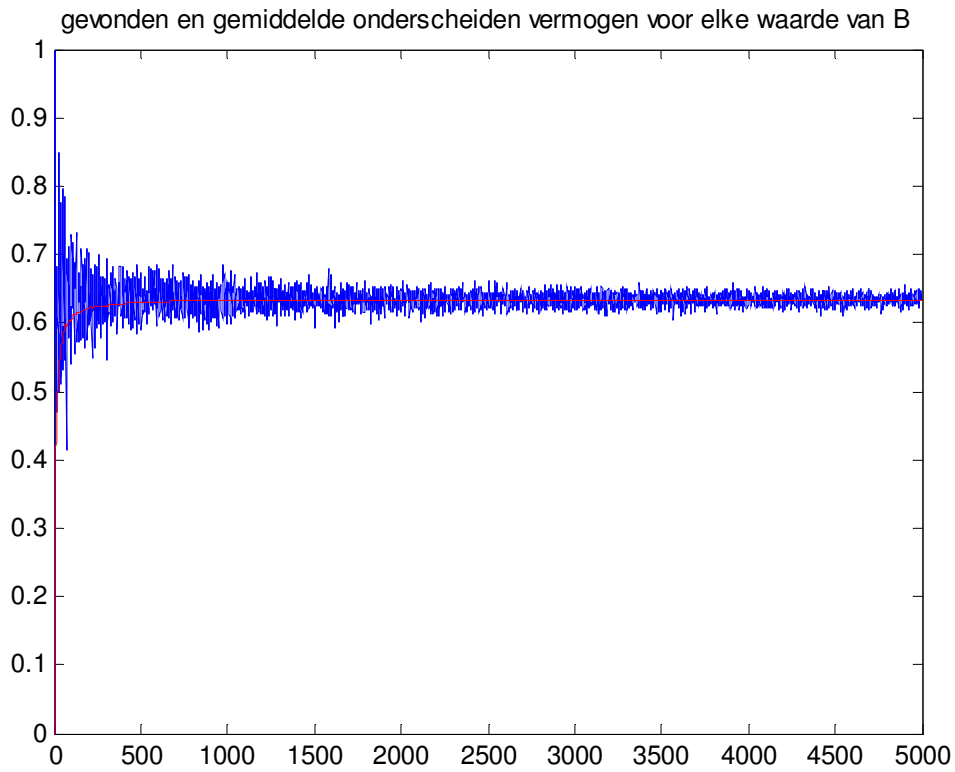
Figuur II.13: Variantie van de ruis van elke voxel van beide proefpersonen.

Onderscheidend vermogen voor verschillende waarden van B



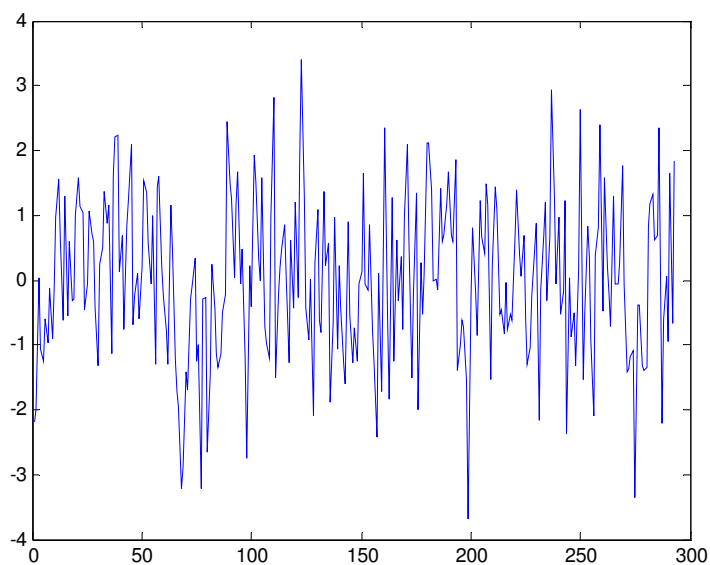
Figuur II.14: Het gevonden onderscheidend vermogen in de verschillende simulaties, waarbij $x = 1 \equiv B = 50$, $x = 2 \equiv B = 100$, $x = 3 \equiv B = 150$, etc.

Onderscheidend vermogen voor verschillende waarden van B

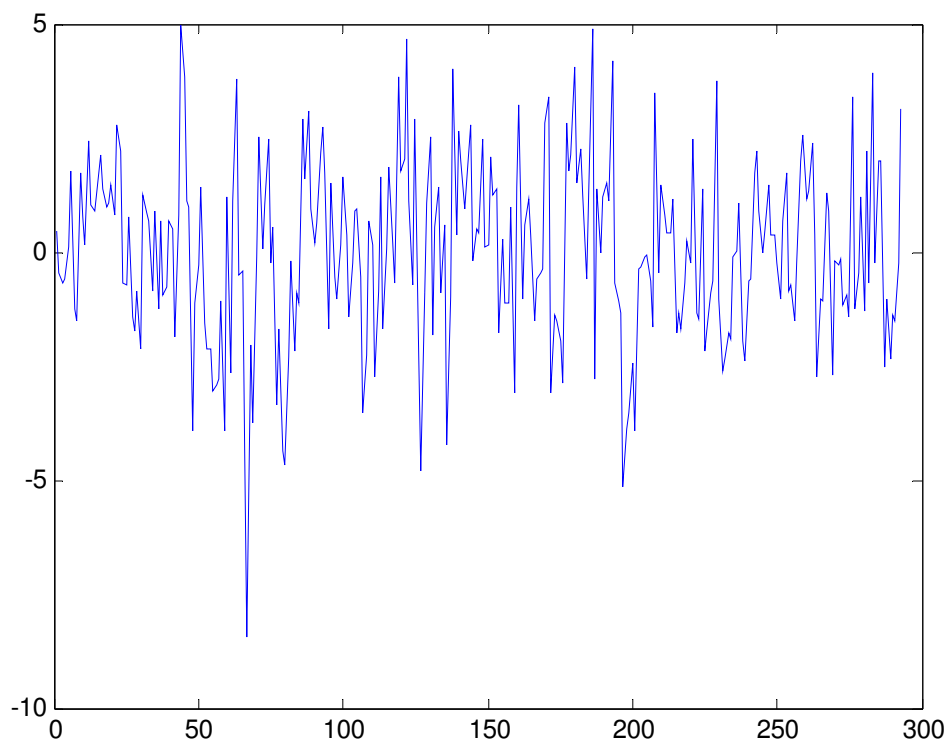


Figuur II.15: Gevonden (blauw) en gemiddelde (rood) onderscheidend vermogen voor B gelijk aan 1 tot en met 5000.

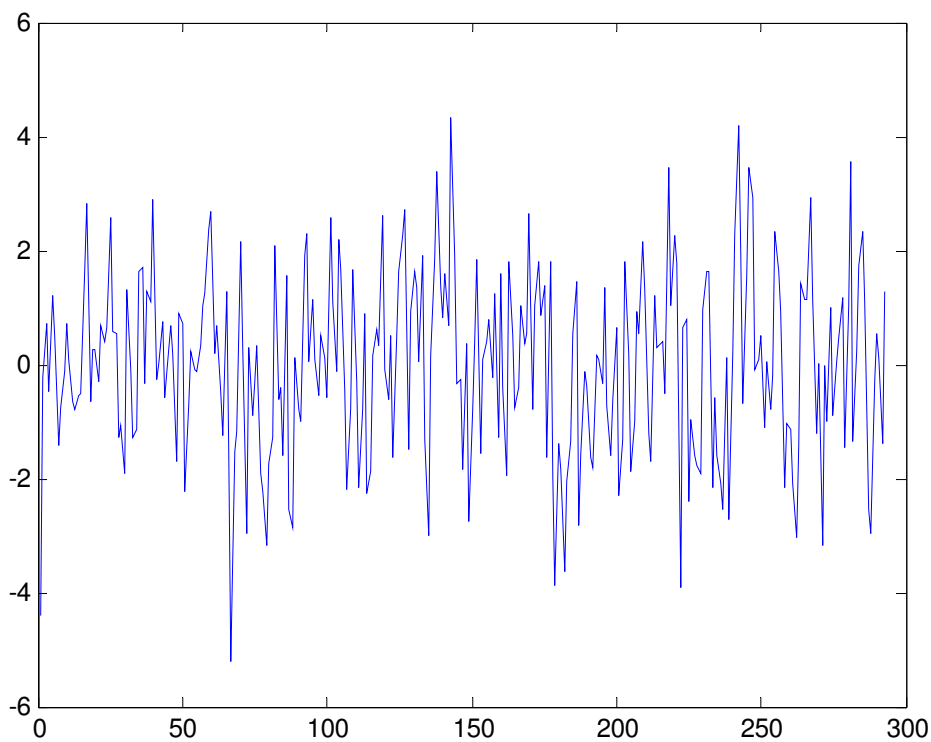
Residuen van het meervoudig regressiemodel in verschillende voxels van proefpersoon 2



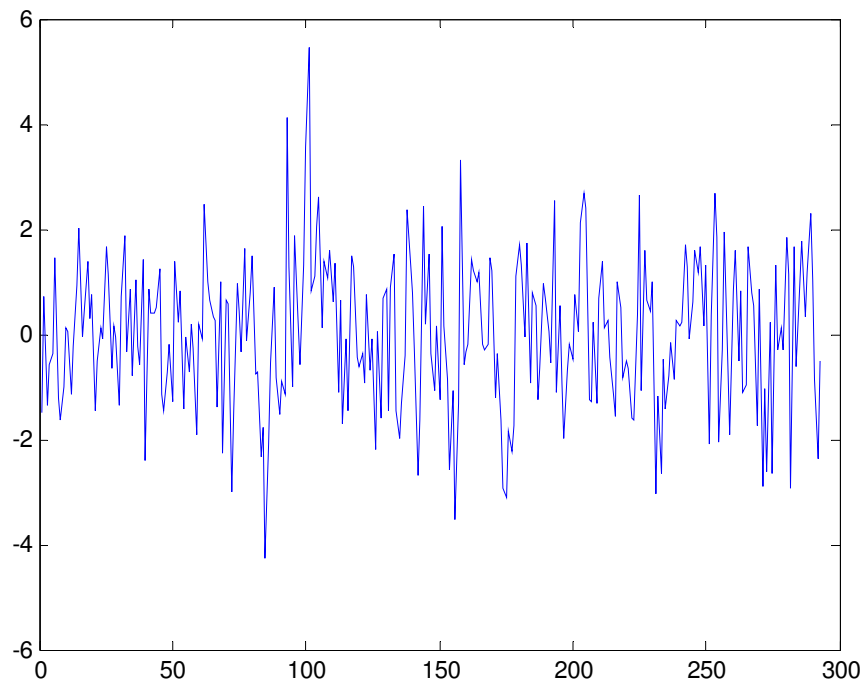
Figuur II.16: Residuen van het meervoudig regressiemodel in voxel 1 van proefpersoon 2.



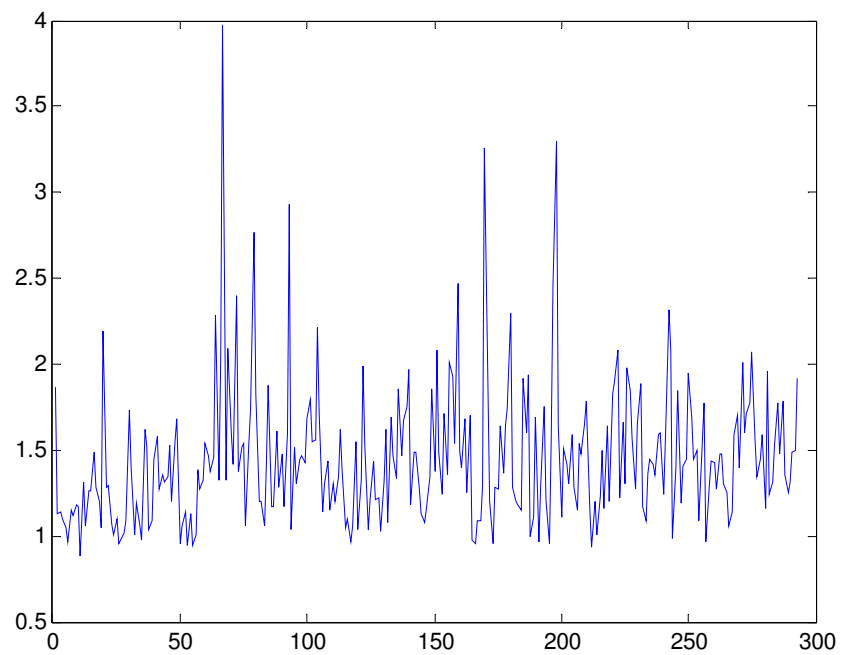
Figuur II.17: Residuen van het meervoudig regressiemodel in voxel 2 van proefpersoon 2.



Figuur II.18: Residuen van het meervoudig regressiemodel in voxel 40 van proefpersoon 2.



Figuur II.19: Residuen van het meervoudig regressiemodel in voxel 90 van proefpersoon 2.



Figuur II.20: Gemiddelde absolute residuen van het meervoudig regressiemodel.

Kenmerken van de data van beide proefpersonen

<i>Proefpersoon</i>	<i>Aantal voxels:</i>	<i>Aantal confounders:</i>	<i>Aantal scans:</i>
1	154	10	300
2	166	9	300

Tabel II.9: kenmerken van de data die gebruikt is in het simulatieonderzoek

Waargenomen waarden voor S^2 en D_n en de bijbehorende overschrijdingskans

<i>Voxel</i>	S^2	D_n	<i>Overschrijdingskans</i>
4	3.96	0.056	0.026
42	1.35	0.040	0.346

Tabel II.10: Waargenomen waarden voor S^2 en D_n en de bijbehorende overschrijdingskans.

Extreme waarden van data van beide proefpersonen

<i>Proefpersoon 1</i>	<i>Proefpersoon 2</i>
201	136
202	137
	227
	271
	272
	273
	274

Tabel II.11: Punten die verwijderd moeten worden van de data van beide proefpersonen.

Geschatte waarden voor σ en σ_x voor beide proefpersonen

	<i>Proefpersoon 1</i>	<i>Proefpersoon 2</i>
σ	1.7728	1.9357
σ_x	0.6637	0.6894

Tabel II.12: Waarden voor de parameters van stap 0.

Parameters voor het bepalen van het juiste aantal simulaties

Proefpersoon	1
b	0.35
σ	1.7728
σ_x	0.6637
α	0.05
γ	0.4
n	232
B	[50, 100, 150, 500, 1000, 5000, 10000, 50000, 100000, 1000000]

Tabel II.13: Gekozen parameters voor het bepalen van de juiste waarden van B .

Resultaten van de simulatie met $b=0.85$ en fMRI ruis voor beide proefpersonen

Subject1 fMRI					Subject2 fMRI				
B	150				B	150			
power\alpha	0.05	0.02	0.01	0.001	power\alpha	0.05	0.02	0.01	0.001
0.6	0.36	0.5933	0.6933	0.64	0.6	0.4467	0.4533	0.34	0.2533
0.7	0.7333	0.6	0.7733	0.4533	0.7	0.4533	0.3667	0.3133	0.4733
0.8	0.56	0.62	0.7	0.7	0.8	0.4933	0.5667	0.4867	0.5
0.9	0.9267	0.92	0.8867	0.8	0.9	0.6467	0.6133	0.62	0.7533
0.95	0.9667	0.9467	0.9267	0.9133	0.95	0.8933	0.84	0.8533	0.86

Tabel II.6: Resultaten van de simulatie met $b=0.85$ en fMRI ruis voor beide proefpersonen.

Resultaten van de simulatie met $b=0.85$ en Gaussische ruis voor beide proefpersonen

Subject1 Gauss						Subject2 Gauss					
B	5000					B	5000				
power\alpha	0.05	0.02	0.01	0.001	0.0001	power\alpha	0.05	0.02	0.01	0.001	0.0001
0.6	0.4538	0.4474	0.4627	0.4333	0.4504	0.6	0.4549	0.435	0.4199	0.352	0.2878
0.7	0.6007	0.5787	0.5763	0.5675	0.5648	0.7	0.5527	0.5008	0.4708	0.504	0.4157
0.8	0.7261	0.7081	0.7227	0.6757	0.7292	0.8	0.6501	0.6527	0.6438	0.6602	0.5859
0.9	0.8648	0.857	0.8624	0.8745	0.8643	0.9	0.8477	0.8316	0.8086	0.8117	0.7825
0.95	0.9367	0.9343	0.9316	0.9311	0.9356	0.95	0.914	0.916	0.9044	0.9029	0.8903

Tabel II.7: Resultaten van de simulatie met $b=0.85$ en Gaussische ruis voor beide proefpersonen.

Resultaten van de simulatie met $b=0.85$, fMRI ruis en gecorrigeerde n voor beide proefpersonen

Subject 1 fMRI						Subject 2 fMRI					
B	150					B	150				
power\alpha	0.05	0.01	0.001	0.0001	0.00001	power\alpha	0.05	0.01	0.001	0.0001	0.00001
0.6	0.94	0.76	0.5133	0.6067	0.7533	0.6	0.7333	0.5133	0.5	0.3867	0.5133
0.7	0.9667	0.7533	0.58	0.7867	0.8467	0.7	0.7933	0.5067	0.46	0.5	0.7667
0.8	0.94	0.8933	0.82	0.7333	0.8733	0.8	0.7267	0.6	0.52	0.7933	0.8667
0.9	0.9267	0.7867	0.8933	0.9333	0.9333	0.9	0.7133	0.72	0.7933	0.9467	0.7667
0.95	0.92	0.9667	0.9333	0.9667	0.94	0.95	0.72	0.8533	0.8533	0.9333	0.82

Tabel II.8: Resultaten van de simulatie met $b=0.85$, fMRI ruis en gecorrigeerde aantal voor beide proefpersonen.

Resultaten van de simulatie met $b=0.85$, Gaussische ruis en gecorrigeerde n voor beide proefpersonen

Subject 1 Gauss						Subject 2 Gauss					
B	5000					B	5000				
power\alpha	0.05	0.01	0.001	0.0001	0.00001	power\alpha	0.05	0.01	0.001	0.0001	0.00001
0.6	0.8848	0.6794	0.5082	0.5774	0.6076	0.6	0.79	0.537	0.4692	0.5202	0.5536
0.7	0.893	0.6914	0.6316	0.7204	0.7358	0.7	0.7832	0.5374	0.5762	0.6332	0.7428
0.8	0.8748	0.7346	0.7924	0.8264	0.8324	0.8	0.7982	0.6362	0.7154	0.7626	0.8032
0.9	0.8822	0.8644	0.9114	0.9236	0.8776	0.9	0.799	0.8268	0.8822	0.9236	0.7894
0.95	0.9408	0.953	0.9676	0.9558	0.8762	0.95	0.8964	0.9308	0.9346	0.9184	0.805

Tabel II.9: Resultaten van de simulatie met $b=0.85$, Gaussische ruis en gecorrigeerde aantal voor beide proefpersonen.

Onderzoek in willekeurig startpunt bij simulatie

Subject 2 fMRI					
B	150				
power\alpha	0.05	0.01	0.001	0.0001	0.00001
0.6	0.7333	0.5133	0.5	0.3867	0.5133
0.7	0.7933	0.5067	0.46	0.5	0.7667
0.8	0.7267	0.6	0.52	0.7933	0.8667
0.9	0.7133	0.72	0.7933	0.9467	0.7667
0.95	0.72	0.8533	0.8533	0.9333	0.82

Tabel II.10: Resultaten van de simulatie met $b=0.85$, fMRI ruis en gecorrigeerde aantal voor proefpersoon 2.

Gesimuleerde power:	Werkelijke power:	Nieuwe simulatie:	Met startpunt:	Nieuwe simulatie:	Met startpunt:
0.5	0.6	>0.6	110	>0.7	150
0.46	0.7	>0.8	120		
0.52	0.8	>0.8	90		
0.7933	0.9	>0.9	50		
0.8533	0.95	>0.95	25		
0.3867	0.6	>0.6	100		
0.5	0.7	>0.8	75		
0.5133	0.6	>0.65	45		

Tabel II.11: Resultaten van de simulatie met $b=0.85$, fMRI ruis, gecorrigeerde aantal en willekeurige startpunt voor proefpersoon 2.