

Eindopdracht Simulaties

Eddy van Beekveld

June 27, 2016

Contents

1	Introductie	2
2	Opdracht	2
2.1	Aannames	2
2.2	Modelbeschrijving	3
2.3	Simulaties	5
2.4	Interpretatie van de modellen	5
2.5	Conclusies	14
2.6	Reflectie	14

1 Introductie

Voor het vak simulaties hebben we een eindopdracht. Van deze eindopdracht moeten we onze bevindingen presenteren, dit door middel van een verslag. Dit verslag is mijn presentatie van mijn bevindingen. De opdracht omschrijving is letterlijk zoals gegeven.

2 Opdracht

Het basisinkomen is in concept een bedrag dat maandelijks wordt uitgekeerd aan iedere Nederlander van 18 jaar en ouder zonder dat daar voorwaarden aan verbonden zijn. In Nederland wordt sinds 2015 geëxperimenteerd met het basisinkomen als vervanger van de bijstand. Omdat gemeenten de reglementen rond het aanvragen en behouden van een bijstandsuitkering als te streng ervaren, onderzoeken zij de mogelijkheden om de bijstandsgerechtigden een maandelijks bedrag uit te keren zonder dat daar een tegenprestatie, zoals sollicitatieplicht, tegenover staat. Tilburg is een van de gemeenten waar een experiment zal worden uitgevoerd met het basisinkomen in samenwerking met Tilburg University onder de noemer ‘Vertrouwensexperiment’.

Voorstanders van het basisinkomen wijzen op de groeiende onzekerheid voor werknemers als gevolg van de flexibilisering van de arbeidsmarkt, de globalisering en de automatisering. Een groot deel van de beroepsbevolking zou geen reële kans hebben op werk met een fatsoenlijk inkomen.

Tegenstanders van het basisinkomen argumenteren dat het concept onbetaalbaar is, de arbeidsmoraal vernietigt en slecht is voor de economie. Sommigen suggereren een alternatief voor het basisinkomen, de basisbaan. Een basisbaan kan worden omschreven als een normale baan waarvoor tenminste het minimumloon wordt betaald. Ieder lid van de Nederlandse beroepsbevolking zou recht hebben op een basisbaan, gegarandeerd en bekostigd door de overheid.

De discussie tussen voorstanders van het basisinkomen en voorstanders van de basisbaan wordt vaak gevoerd op ethisch niveau en zonder feitelijke argumentatie aangaande financiële gevolgen van het basinkomen of de basisbaan. Jouw opdracht is om de voor- of tegenstanders te voorzien van numerieke argumentatie. Door modellen te beschrijven en implementeren voor zowel basisinkomen als basisbaan, kun je de discussie voeren met numerieke argumentatie.

Stapsgewijs worden in de opdracht aannamen benoemd en de modellen opgesteld. Middels Monte Carlo simulatie kun je het effect van de modellen, de financiële gevolgen van de implementatie van het basisinkomen of de basisbaan, benaderen en daarna uitkomsten met elkaar vergelijken.

2.1 Aannames

Voor deze opdrachten zijn er een aantal aannames gemaakt. Dit zijn de volgende:

- Het minimumloon is 1524 euro. Dit is zonder vakantiegeld welke 8 procent bedraagt. Het basisinkomen komt dan op $1524 * 1.08 = 1645.92$ euro.
- de totale beroepsbevolking is 8.9 miljoen
- Het totaal aantal 18+ers is $16.9 * 0.7971 = 13.47$ miljoen
- Het aantal arbeidsongeschikte is 815 duizend mensen
- De kosten uit het huidige sociale stelsel zijn 140 miljard

Deze aannames zijn gegeven.

2.2 Modelbeschrijving

Voor deze opdracht hadden we twee modellen. Het basisinkomen en de basisbaan model. In deze sectie vertel ik hoe deze modellen tot stand zijn gekomen.

2.2.1 Basisinkomen

Het kostenbatenmodel van het basisinkomen bestaat uit 4 onderdelen welke samen het gehele model maken.

Als eerste heb je de directe kosten:

$$\text{directekosten} = \text{aantal18 + ers * basisinkomen} \quad (1)$$

Deze directe kosten zijn rechttoe rechtaan. het aantal volwassen vermenigdvuldigd met het basisinkomen. Hieruit zijn de directe kosten ontstaan.

Als tweede heb je administratiekosten. De administratiekosten zijn onbekend maar we kunnen wel een schatting maken. De administratieve kosten volgen een normaal verdeling met als μ gelijk aan 250 en σ gelijk aan 75:

$$\text{administratiekosten} = \sim \mathcal{N}(\mu = 250, \sigma = 75) \quad (2)$$

Als derde hebben we een schatting nodig of het arbeid toe of afneemt door introductie van het basisinkomen nodig. Hier hebben we een aantal vergelijkingen voor nodig.

Als eerste de non participatie factor. Het is waarschijnlijker dat het basisinkomen arbeidsparticipatie meer aanmoedigt dan ontmoedigt. Daar door hebben we deze factor uniform verdeeld met als ondergrens -0.05 en als boven grens 0.10.

$$\text{nonparticipatie} = \sim U(a = -0.05, b = 0.10) \quad (3)$$

Vervolgens hebben we ook de arbeidsproductie nodig. Dit is gebaseerd op mensen met een marginale productiviteit en een laag inkomen. Dit omdat het meerendeel van de mensen die niet meer werken eerder in deze doelgroep zullen vallen dan een andere doelgroep. De productiviteit is als volgt verdeeld:

$$\text{arbeidsproductiviteit} = \sim \mathcal{N}(\mu = 10, \sigma = 1) \quad (4)$$

Bij deze arbeidsproductiviteit gaan we uit van een 40 urige werkweek, 50 weken per jaar.

Met deze twee formules kunnen we de derde formule voor ons model op gaan stellen:

$$\text{veranderingarbeid} = -1 * (\text{aantalvolwassen} - \text{beroepsbevolking} - \text{arbeidsongeschikte}) * \text{nonparticipate} * 40 * 50 * \text{arbeidsproductiviteit} \quad (5)$$

Als laatste voor ons model van het basisinkomen hebben we het zogenoemde j.k Rowling effect. Dit is iemand die geen baan heeft maar creatief bezig is en kans heeft om door te breken en dan ook een grote waarde genereert:

$$\text{rowlingwaarde} = 10^9 * \sim \text{bin}N = \text{aantalwerklozen}, p = \frac{1}{10^7} \quad (6)$$

Het totaal model voor het basisinkomen is:

$$\text{model} = \text{directekosten} + \text{administratiekosten} + \text{veranderingarbeid} + \text{rowlingwaarde} \quad (7)$$

2.2.2 Basisbaan

Het tweede model gaat over de basisbaan. Het model lijkt op die van het basisinkomen. Volwassen die bij de beroepsbevolking horen, komen in aanmerking voor een basisbaan. Het salaris is vergelijkbaar met het minimulloon.

Als eerste hebben we nodig de toegevoegde waarde van een basisbaan per uur. Deze is uniform verdeeld en gegeven door de volgende functie:

$$toegevoegdewaarde = \sim U(a = 0, b = 9) \quad (8)$$

Als tweede hebben we hier ook administratiekosten zowel per basisbaan als per arbeidsongeschikte. Deze kosten per basisbaan zullen vele malen hoger uit vallen dan die van de arbeidsongeschikte omdat er ook leidinggevende betaald moeten worden. Verder zijn ze beide normaal verdeeld:

$$administratiekostenarbeidsongeschikte = \sim \mathcal{N}(\mu = 500, \sigma = 150) \quad (9)$$

$$administratiekostenbasisbaan = \sim \mathcal{N}(\mu = 5000, \sigma = 1500) \quad (10)$$

Nu moeten we ook schatten welk aandeel van de volwassen niet beroepsbevolking dit model zal kiezen voor een basisbaan. We nemen een uniforme verdeling hiervoor met een afname van niet arbeiders van 20 % en een toename van niet arbeiders met 5 %

$$veranderingnietarbeiders = \sim U(a = -0.20, b = 0.05) \quad (11)$$

Als laatste moeten we nog definieren hoe het aantal basisbanen wordt gedefinieerd voordat we het model kunnen opstellen. Het aantal basisbanen is als volgt gedefinieerd:

$$aantalbasisbanen = aantalvolwassenen - arbeidsongeschikte - beroepsbevolking \quad (12)$$

we kunnen nu de kosten baten per vervulde basisbaan en per arbeidsongeschikte opstellen. Deze bij elkaar opgeteld representeert het model voor de basisbaan.

$$basisbaankostenbaten = basisinkomen + administratiekostenbasisbaan + 40 * 50 * toegevoegdewaarde \quad (13)$$

$$arbeidsongeschiktekostenbaten = basisinkomen + administratiekostenarbeidsongeschikte \quad (14)$$

Het model voor de basisbaan is dan:

$$basisbaanmodel = aantalbasisbanen * basisbaankostenbaten + aantalarbeidsongeschikte * arbeidsongeschiktekostenbaten \quad (15)$$

Deze twee modellen heb ik samen met de aannames in het bestand model.R geïmplementeerd. Deze file is bijgesloten en heeft zelf allemaal comments die uitlegt wat er staat.

2.3 Simulaties

Als volgende heb ik deze twee modellen gesimuleerd met behulp van een monte carlo simulatie. Dit heb ik 100,000 keer gedaan voor beide modellen om zo accuraat mogelijke data te krijgen.

Deze data heb ik opgeslagen naar het bestand data.csv. Dit heeft meerdere redenen, maar de voornaamste reden is dat de gesimuleerde data dan gelijk blijft en je over de data conclusies kan trekken. Ook omdat het 200 honderduizend keer model berekenen even duurt en dat is lastig voor mijn verslag, dit omdat deze de grafieken zelf genereert.

2.4 Interpretatie van de modellen

Als eerste kijken we naar wat KPI's, voor zowel het basisinkomen als de basisbaan.

Basisbaan:

	Minimum	Maximum	Median	Mean	Total
kosten	7,262,047,658	113,518,396,854	60,500,100,817	60,532,988,494	6.053299e+15

Basisinkomen:

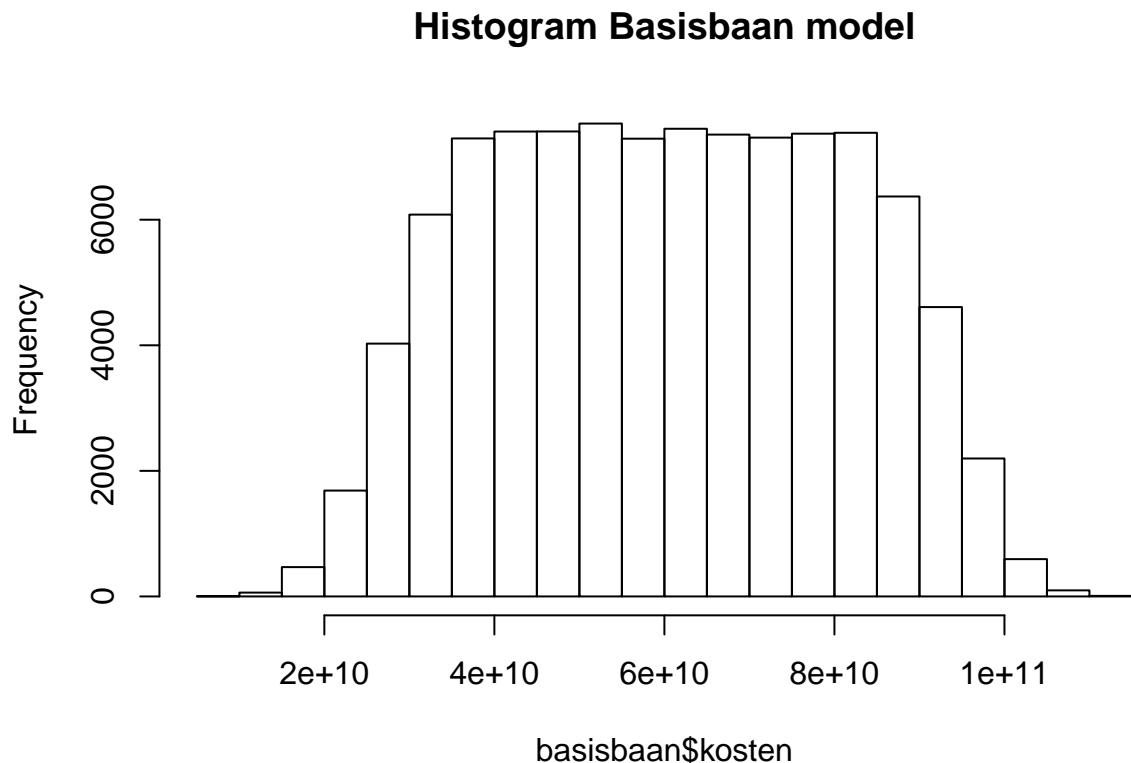
	Minimum	Maximum	Median	Mean	Total
kosten	12,525,154,967	28,583,908,707	20,384,402,936	20,365,286,861	2.036529e+15

Dit geeft ons al snel wat inzichten in de data. Zo zie je dat het gemiddelde van basisinkomen lager is dan die van de basisbaan. Ook het maximum is groter bij basisbaan. Maar het minimum is daarentegen weer lager bij de basisbaan.

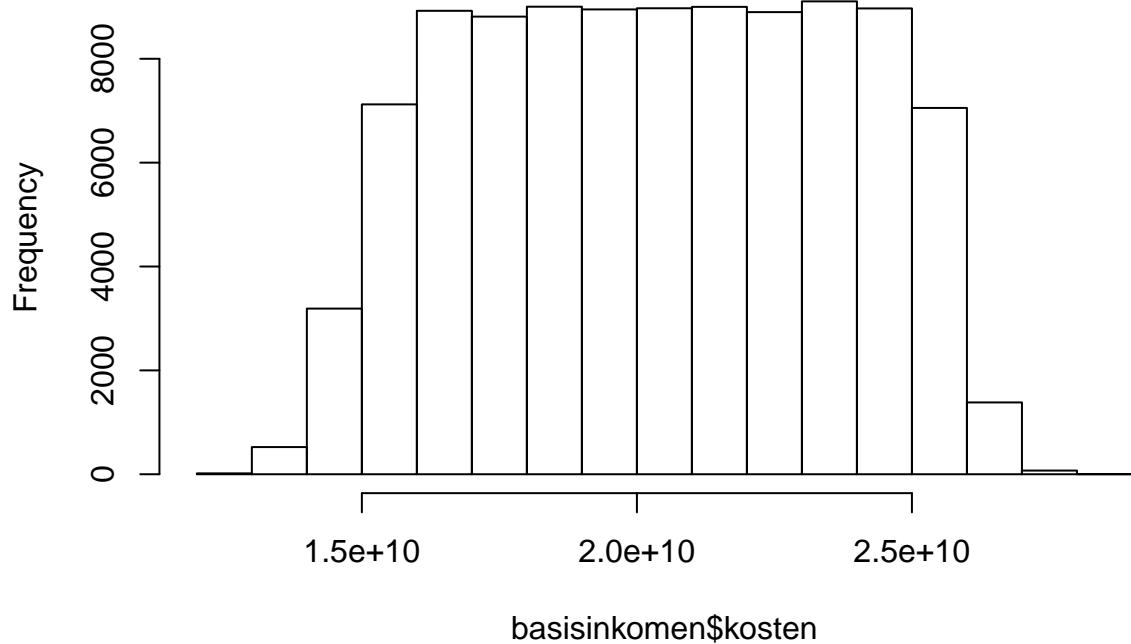
2.4.1 Histogram

Tijd om de data wat meer in kaart te brengen.

Als eerste heb ik van beiden modellen een histogram gemaakt. Dit geeft al een mooie indicatie voor de data.

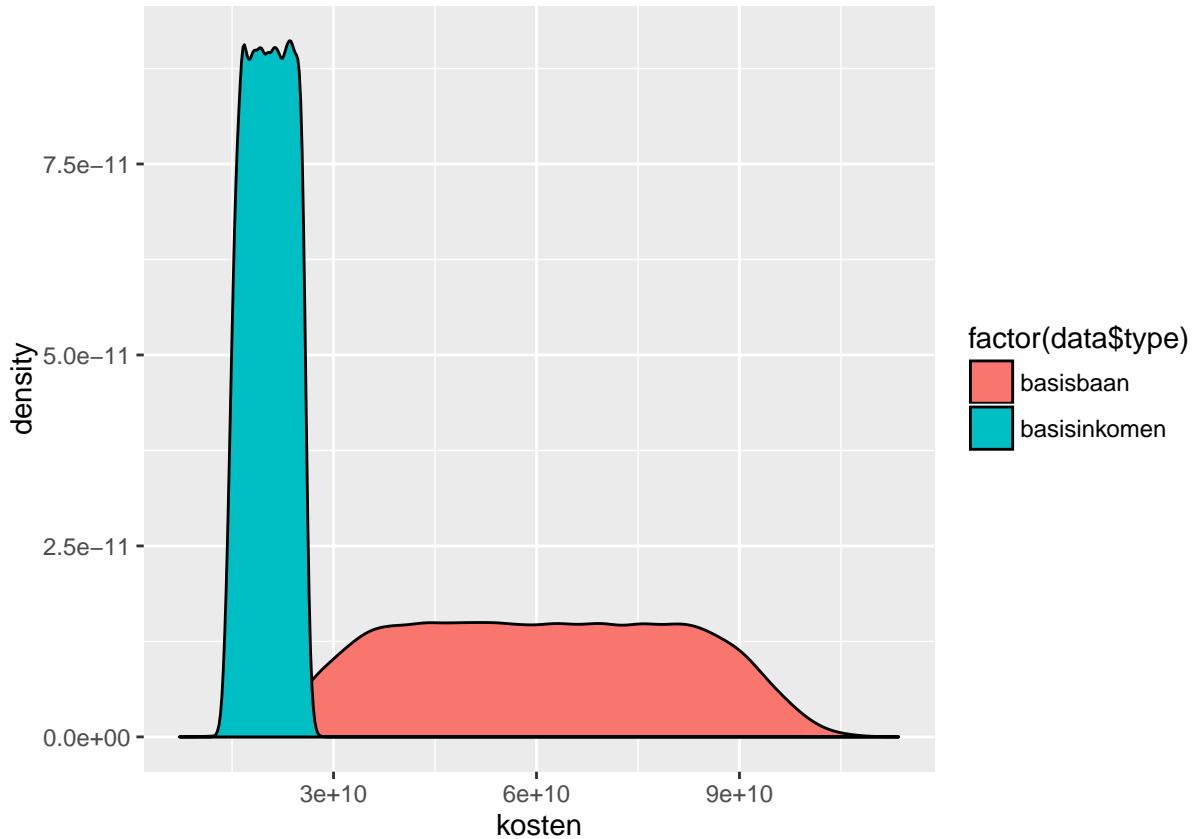


Histogram basisinkomen model



2.4.2 Density plot

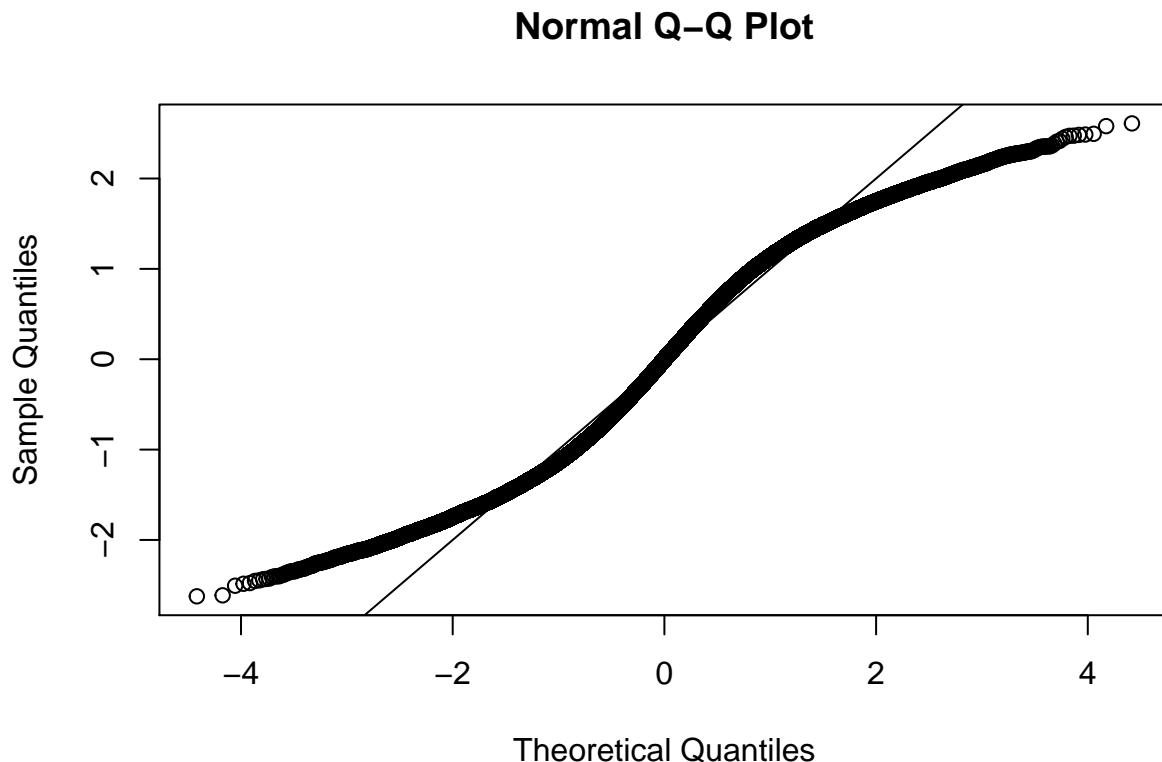
Daarna heb ik een densityplot gemaakt van beide modellen in een grafiek:



Dit geeft al veel meer indruk hoe de data verdeeld zal zijn. Bij het basisinkomen lijkt het uniform verdeeld te zijn, en bij de basisbaan kan het zowel uniform als normaal verdeeld zijn. Het vermoeden wat ik heb is dat deze normaal verdeeld is.

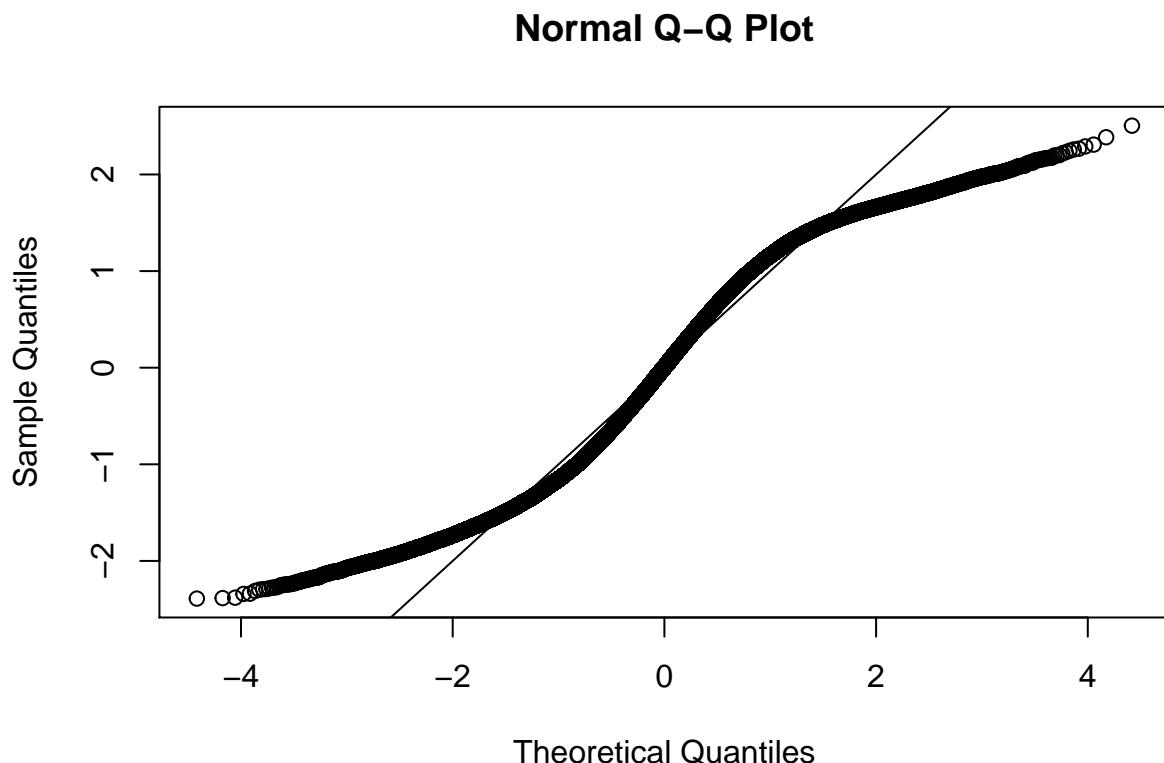
2.4.3 QQ plots

Ik heb nu voor het beide modellen een QQ plot gemaakt, bij het model van de basisbaan om te kijken of de data normaal verdeeld is en bij het basisinkomen om te kijken of het ons wellicht meer inzicht geeft of dat het naar verwachting niet normaal verdeeld is.



In deze grafiek zie je dat ook basisbaan waarschijnlijk niet normaal verdeeld is. De punten van de qqplot liggen niet goed in lijn met de getrokken lijn. Welke wanneer de data normaal verdeeld was beter zou matchen.

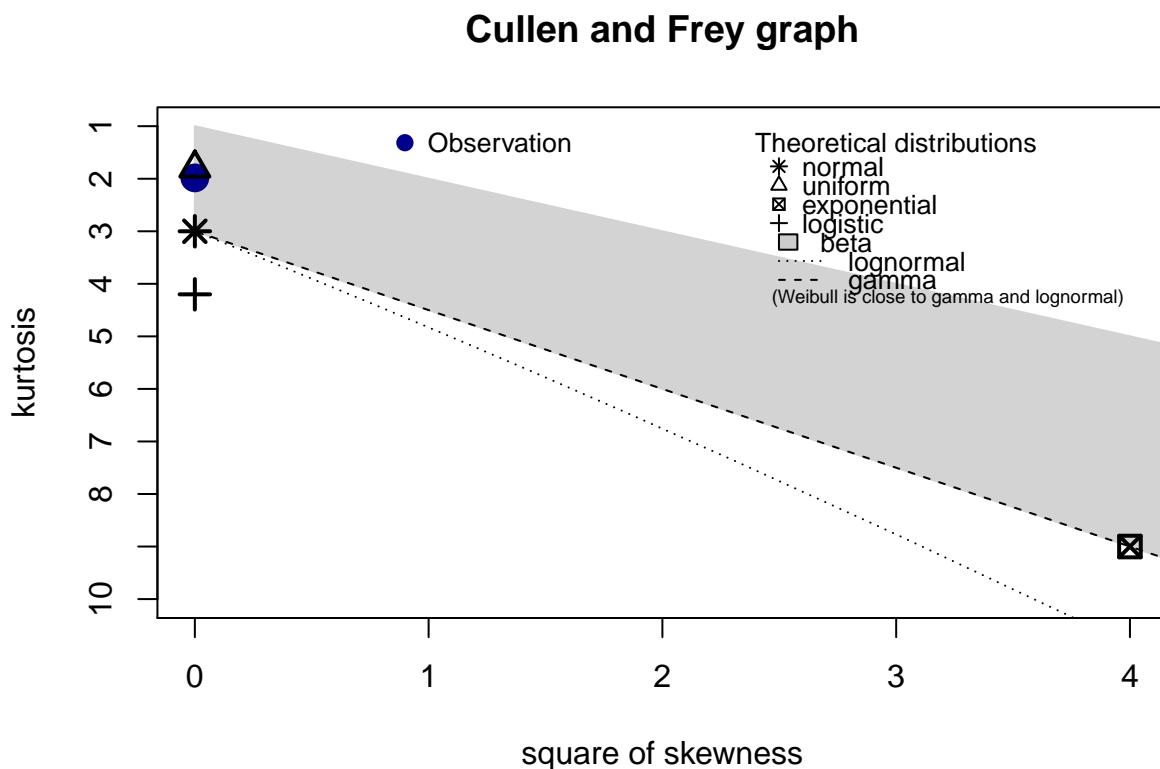
Als volgt de QQplot van het basisinkomen echter verwachten we dat deze ook niet normaal verdeeld is.



2.4.4 Cullen and frey grafiek

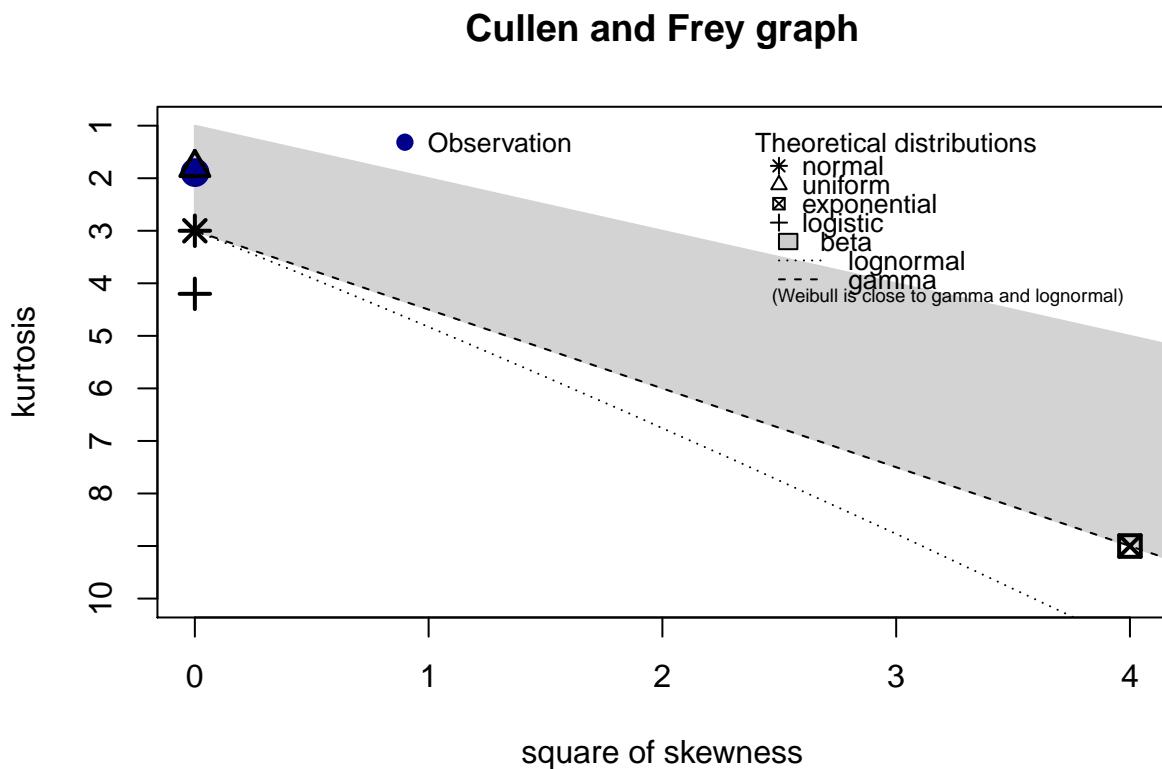
Dus gaan we kijken of de basisbaan dan uniform verdeeld is. Dit doen we door middel van De Cullen en Frey grafiek. Deze grafiek geeft ons inzicht in welke verdeling de data ongeveer volgen.

Allereerst de basisbaan:



In de grafiek kan je zien dat de data uniform verdeeld is. De observatie(de data) ligt zelfs voor een gedeelte op de plek van de distributie. Dit geeft dus een goede indicatie dat de data voor de basisbaan uniform verdeeld is.

Als tweede de Cullen and Frey grafiek voor basisinkomen



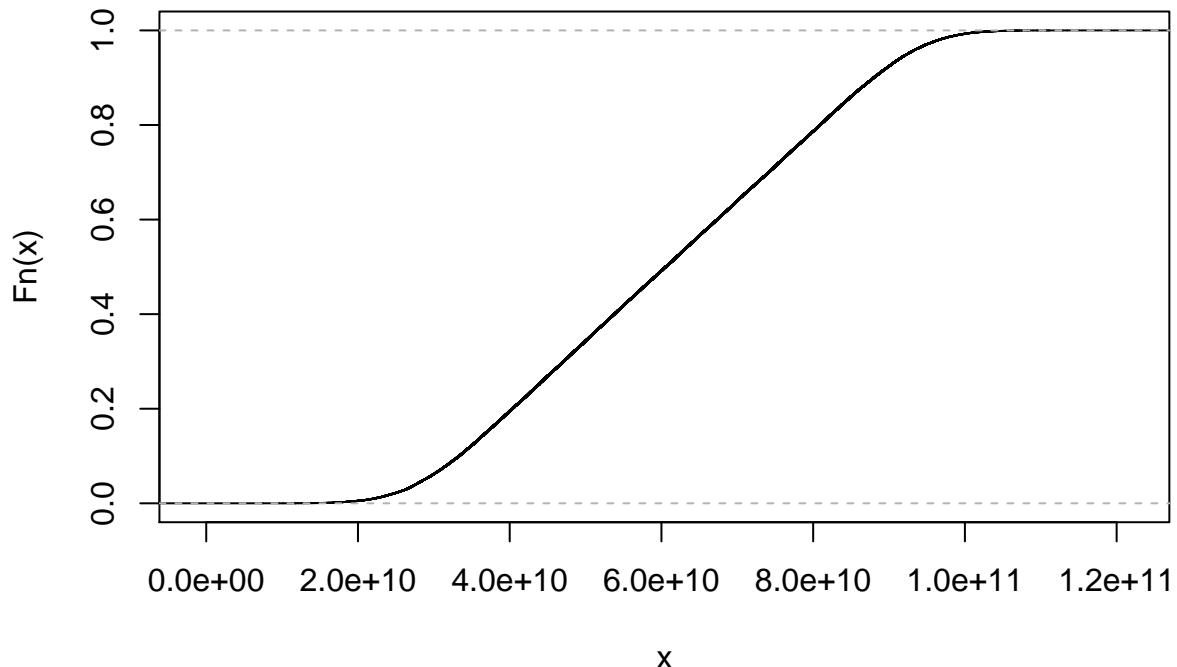
Deze is naar verwachting ook te zien als uniform verdeeld.

Beide modellen zijn dus uniform verdeeld.

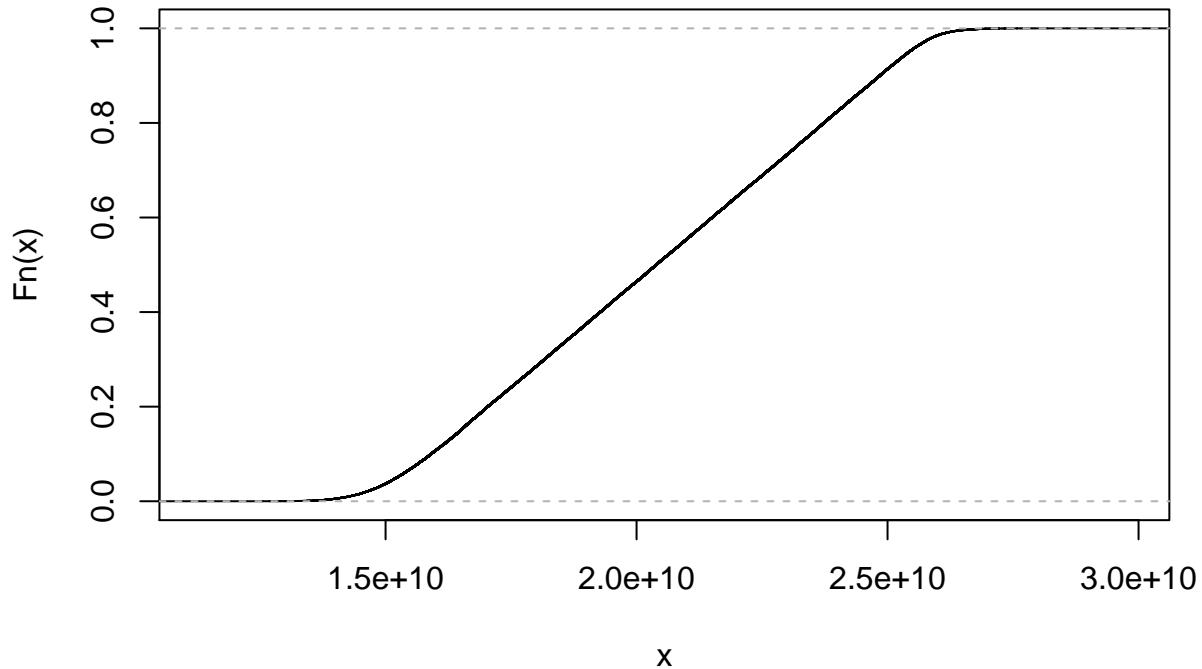
2.4.5 Emperische cumulatieve distributie

Als laatste test heb ik nog een ECDF gemaakt van beide modellen. Dit geeft ons ook iets meer inzicht in hoe het verdeeld is. Naar verwachting, omdat hij uniform verdeeld is zal het een aardig rechte lijn zijn met op het begin en op het einde een kleine bolling.

Emperische cumulatieve distributie functie Basisbaan



Emperische cumulatieve distributie functie basisinkomen



2.5 Conclusies

Beide data waren uniform verdeeld, maar het kosten baten plaatje van het basisinkomen is toch een stuk beter. De kosten zijn gemiddelde gezien lager en ook het maximum ligt significant lager. Wel lijken de kosten van een van de huidige situatie significant lager. De vergelijking daarmee is dus onmogelijk.

Maar van beiden modellen is te zeggen dat het basisinkomen beter is qua kosten baten plaatje dan de basisbaan.

2.6 Reflectie

De modellen waren niet te vergelijken met de huidige situatie. Wellicht zit er een foutje in mijn modellen. Maar ik heb ze meermaals gecontroleerd en heb het niet kunnen vinden. Wellicht dat ik de volgende keer eerder hierna kan kijken want, ik had de data al eerder gemaakt echter het rapport pas vandaag. Verder denk ik dat het een leuke leerzame opdracht is. Ook heb ik een betrouwbaarheidsinterval gemaakt echter was deze relatief gezien zo klein om het gemiddelde dat ik niet het gevoel had dat deze correct was. De formule voor het betrouwbaarheidsinterval die ik gebruikte heb was:

$$\text{mean}(\text{basisinkomen}) + c(-1, 1) * 1.96 * \text{sd}(\text{basisinkomen}) / \sqrt{100000} \quad (16)$$

Met als uitkomst: min:20,344,950,520 max:20,385,623,201 voor een betrouwbaarheidsinterval van 95%.