



Universiteit
Leiden
The Netherlands

Playing dice with the universe: Bayesian statistical analyses of cosmological models and new observables

Cañas Herrera, G.

Citation

Cañas Herrera, G. (2022, October 19). *Playing dice with the universe: Bayesian statistical analyses of cosmological models and new observables*. *Casimir PhD Series*. Retrieved from <https://hdl.handle.net/1887/3483658>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/3483658>

Note: To cite this publication please use the final published version (if applicable).

Samenvatting

De fascinatie voor de nachtelijke hemel. Een eenvoudige blik op het prachtige landschap boven ons opent een venster naar de verbeelding. Uit deze fascinatie samen met onze aangeboren menselijke nieuwsgierigheid, is de passie voor het begrijpen van de wereld waarin wij leven geboren. Van oude culturen tot nu, dezelfde vragen hielden ons bezig: "Waar komen we vandaan?" en "Wat is onze plaats in het heelal?".

Het begrijpen en bestuderen van de kosmos is een onderwerp dat niet alleen wetenschappers, maar ook de bevolking in het algemeen. Een snelle zoektocht op de populairste fora op het internet toont aan dat ze allemaal actieve berichten hebben die gewijd zijn aan gesprekken over het heelal¹³. Bovendien wordt elke spannende astronomische ontdekking binnen een paar uur verspreid, zoals het geval was bij de allereerste foto van het superzware zwarte gat Sgr A*, gemaakt door de Event Horizon Telescope (EHT).

De samenstelling van het heelal, de vermoedelijke oorsprong ervan en de belangrijkste dominante kracht, de zwaartekracht, maken deel uit van de popcultuur van onze beschaving. Maar hoe wordt (en werd) het heelal bestudeerd? We vertrouwen op astronomische waarnemingen, bijvoorbeeld van sterren en andere massieve, compacte objecten in het heelal, om onze wetenschappelijke kennis op te bouwen, gebaseerd op de vergelijking van die waarnemingen met voorspellingen gebaseerd op een aantal fysische veronderstellingen, die een kosmologisch model vormen. De gemiddelde leek is bijvoorbeeld bekend met het begrip "Big Bang".

Kosmologen hebben gevonden dat het heelal ongeveer 13,77 miljard jaar geleden is ontstaan uit een zeer dichte en hete toestand. Het jonge heelal koelde af terwijl het uitdijde, en het dijt nu nog steeds sneller uit. Ons heelal is opgebouwd uit twee hoofdbestanddelen: een onbekende materie die gravitationeel op elkaar inwerkt, genaamd donkere materie, en de stof die verantwoordelijk is voor de versnelde uitdijning van het heelal, donkere energie. De materie waar wij allemaal van gemaakt zijn, de baryonische materie, maakt slechts ongeveer 5% uit van de inhoud van het heelal. Het is dan ook niet verwonderlijk dat begrippen als donkere energie en donkere materie omgeven zijn door een waas van mysterie die de belangstelling van het publiek wekt. Wetenschappelijk gezien is het grootste deel van de kennis over de samenstelling van

¹³Zie bijvoorbeeld <https://www.quora.com/Why-do-you-love-to-gaze-at-the-stars-of-the-night-sky>.

het heelal verkregen door waarnemingen van de kosmische achtergrondstraling (Cosmic Microwave Background): de reststraling die in het vroege heelal is uitgezonden toen de eerste neutrale atomen werden gevormd en fotonen eindelijk konden verstrooien en zich vrij door het heelal konden bewegen. Uit deze waarnemingen hebben we ook geconcludeerd dat ons heelal geometrisch vlak is.

Naarmate het heelal afkoelde en uitdijde, begon de zwaartekracht een grotere rol te spelen en ontstonden de huidige structuren die in ons heelal waar te nemen zijn: van de eerste sterren tot de eerste sterrenstelsels, en van de eerste sterrenstelsels tot de eerste clusters van sterrenstelsels. Deze objecten vormden de zogenaamde grootschalige structuur van het heelal, die eruitziet als een web. De zwaartekracht is echter niet de enige speler bij de vorming van de grootschalige structuur van het heelal. Er moet een onderliggende Gaussische verdeling van primordiale dichtheidsverstoringen bestaan die de oorsprong van de huidige structuur kan verklaren. Tot nu toe wordt aangenomen dat inflatie, een exponentiële expansiefase in het oerheelal, het mechanisme is voor de productie van deze primordiale kiemen. Het eenvoudigste model van inflatie is verenigbaar met de veronderstelling van een Gaussische verdeling van de primordiale dichtheidsverstoringen, en verklaart tevens waarom de kosmologische waarnemingen erop wijzen dat het heelal causaal verbonden en vlak is. Het inflatieparadigma vormt samen met de aanname dat ons heelal voor het grootste deel uit donkere materie en donkere energie bestaat, de basis van het standaard kosmologisch model.

In de komende jaren zullen naar verwachting steeds meer kosmologische waarnemingen beschikbaar komen om het heelal te bestuderen. In het bijzonder zal gebruik worden gemaakt van de informatie die is gecodeerd in de grootschalige structuur van het heelal. Daartoe zijn nieuwe missies en experimenten gericht op het maken van grote verzamelingen van gegevens van sterrenstelsels, die informatie zullen bevatten over hun posities aan de hemel, hun vormen en ook hun roodverschuivingen. Met de catalogi zullen kosmologen de samenstelling van het heelal bestuderen aan de hand van waarnemingen zoals clustering van sterrenstelsels (GC) en zwakke lenseffecten (WL). Een van deze experimenten, de middelgrote *Euclid* missie van de Europese ruimtevaartorganisatie (European Space Agency), waarvan de satelliet naar verwachting binnenkort zal worden gelanceerd, is erop gericht de fysische oorsprong van de versnelde uitdijning van het heelal te begrijpen en inzicht te krijgen in de aard van donkere materie. De *Euclid* missie zal ook de begincondities bestuderen die aan de basis lagen van het ontstaan van het heelal en die verantwoordelijk waren voor de vorming van de structuur van het kosmische web. Daartoe zal *Euclid* de grootschalige structuur van het heelal in kaart brengen door een van de grootste catalogi van sterrenstelsels ooit te maken.

Wat is de methodologie die wordt gebruikt om astrofysische waarnemingen te vergelijken met fysische modellen om conclusies te trekken over het heelal waarin we leven? Kosmologen hebben in de laatste 30 jaar gebruik gemaakt van statistiek, in het bijzonder van een Bayesiaanse statistische benadering, om verschillende kos-

mologische modellen te toetsen aan data. Bayesiaanse statistiek is gebaseerd op het theorema van Bayes. Dit theorema zegt ons dat onze graad van geloof kan worden gecodeerd in een waarschijnlijkheidsverdeling. De Bayesiaanse statistiek gaat ervan uit dat de kansverdeling van een aantal parameters van een model, gegeven een aantal geobserveerde gegevens, recht evenredig is met het product van de kansverdeling die de waarschijnlijkheid geeft van het observeren van de gegevens gegeven het model M en de parameterwaarden maal de kansverdeling die een aantal *a priori* bekende informatie over de theorie of het experiment codeert. In de laatste decennia is het succes van de Bayesiaanse statistiek toegepast op de kosmologie vooral te danken aan de exponentiële toename van de rekenkracht van computers die massale numerieke inferentie voor het eerst haalbaar maakte.

Bayesiaanse statistiek kan worden gebruikt voor een groot aantal wetenschappelijke vraagstukken. Het stelt ons bijvoorbeeld in staat om, gegeven een model en enkele gegevens, de best passende waarden van de parameters van dat model te bepalen. Daarnaast kunnen wij met deze statistische benadering ook nagaan of een alternatief kosmologisch model volgens de statistiek de voorkeur geniet ten opzichte van het standaard kosmologisch model. Bovendien zijn Bayesiaanse statistieken ook nuttig in voorspellingsanalyses, waarin we willen testen hoe gevoelig een toekomstig experiment zal zijn voor de mogelijke detectie van een nieuw waarneembaar object of zelfs, gegeven de experimentele opzet van een toekomstige missie, of er een kans is dat sommige uitbreidingen van het standaard kosmologisch model statistisch kunnen worden uitgesloten.

Dit proefschrift is gewijd aan de Bayesiaanse statistische analyses van uitbreidingen van het standaard kosmologisch model met behulp van verschillende astronomische datasets, en aan de voorspelling van nieuwe waarneembare gegevens of experimenten. Het gebruik van deze methodologie is de rode draad van het proefschrift. Na hoofdstuk 1, waarin we de belangrijkste concepten van de kosmologie en de basisprincipes van de Bayesiaanse statistiek introduceren, is het proefschrift verdeeld in drie verschillende delen, afhankelijk van het doel waarvoor de Bayesiaanse statistische methodologie werd gebruikt:

- Het eerste deel richt zich op datawetenschap en inflatie, en het heeft als doel inflatiemodellen te begrenzen met behulp van geavanceerde inferentietechnieken en voorspellingsinstrumenten. **Hoofdstuk 2** toont de allereerste resultaten van de reconstructie van de geluidssnelheid van het veld dat verantwoordelijk is voor de inflatie, met behulp van de nieuwste gegevens over de kosmische achtergrondstraling (CMB) van Planck 2018 en moderne algoritmes (Gaussische processen). **Hoofdstuk 3** is gewijd aan de voorspelling van een bepaalde klasse van inflatiemodellen met één veld, bekend als -attractoren, voor een toekomstig CMB fase-IV-experiment met behulp van een model-afhankelijke, alternatieve benadering voor de verzameling van de inflatieparameters op basis van de huidige beperkingen die verkregen zijn met gegevens over de kosmische achtergrondstraling en de grote schaal structuur.

- Het tweede deel van het proefschrift is gewijd aan de nieuwe interesse in de cross-correlaties van gravitatie-golf (GW) fysica en grote schaal structuur waarnemingen; en in het bijzonder, clustering van sterrenstelsels. In deze twee projecten bestuderen we hoe we de informatie in deze nieuwe waarneemgegevens kunnen exploiteren door hun gedrag en mogelijke detectie te voorspellen met behulp van toekomstige experimenten en Bayesiaanse statistiek. In **hoofdstuk 4** bestuderen we de ruis van overlappende GW signalen die de astrofysische zwaartekrachtgolfachtergrond vormen, en hoe we de correlatie tussen de structuur van die achtergrond en de clustering van melkwegstelsels kunnen gebruiken om zowel astrofysische als kosmologische informatie te verkrijgen. Anderzijds onderzoeken we in **hoofdstuk 5** hoe machine learning technieken kunnen worden gebruikt om de propagatie van tensorverstoringen te reconstrueren door de ruimtelijke correlatie tussen individueel herkenbare GW bronnen en melkwegstelsels te combineren.
- Het derde deel van dit proefschrift is gewijd aan de *Euclid* missie en de taken van het Euclid Consortium: doelstellingen, hoofdkenmerken van de survey en de primaire *Euclid*-observatiesondes. In het bijzonder wordt in **hoofdstuk 6** aandacht besteed aan cruciale software voor de data-analyse voor de missie: de code *Cosmological Likelihood for Observables in Euclid*, ook bekend als CLOE. Dit Bayesiaanse analyse-instrument is ontworpen om de parameterwaarden van een model te kunnen afleiden uit de toekomstige metingen van de *Euclid* missie. In dit hoofdstuk beschrijven we het geïmplementeerde kosmologische recept van de primaire sondes, evenals de beschrijving van de statistische waarschijnlijkheid van *Euclid* data. De resultaten met betrekking tot de structuur van CLOE en hoe goed de code presteert bij het beperken van kosmologische parameters zijn een voorproefje van wat we hopen te leren van de *Euclid* missie, ik een groot deel van mijn tijd heb besteed.