Exoplanet Analysis*

Dingding Wang

2022-04-07

Abstract

The discovery of numerous exoplanet systems containing diverse populations of planets orbiting very close to their host stars challenges the planet formation theories based on the solar system. This paper proposes a general model that uses statistical techniques to explain the relationship between exoplanets detected as of April 2022 and the major features of their host stars. The main goal is to establish a mathematical relationship between a set of variables to better describe the physical characteristics of the host star and the planet itself.

keywords: Linear regression, Bayesian Inference, Cross-sectional data, Exoplanets, Orbital properties, Dynamical evolution

Contents

1 Introduction

An extrasolar planet (or exoplanet) is a planet that orbits a star other than the sun and therefore belongs to an extrasolar planetary system. For a long time, the only known planetary system is the solar system, which

*Code and data are available at: https://github.com/estherxwang/EXOPLANET-Analysis.

greatly limits the research on this problem. Since the discovery of the first exoplanet orbiting a sun-like star in 1995 (which won the 2019 Nobel Prize in Physics), the number of known exoplanets has exploded, fueling the field's rapid development and making it one of the most active frontiers in astronomy today. One of the most surprising results of exoplanet research is that most of the discovered planetary systems have surprisingly different properties and orbits than the planets in our solar system. The findings challenge traditional theories of planetary formation and offer an unprecedented opportunity to answer questions about the origins of planetary systems. Over the past decade, NASA's Kepler satellite has revolutionized research by hunting for thousands of exoplanets using the transiting method, or planetary occultation of stars. These massive discoveries have led to a major trend in exoplanet research, which is to systematically study the statistical properties of the distribution of planets and their host stars using large samples of observational data, and to further understand the physical mechanism and environment of planet formation and evolution.

A few years ago, a team of NASA scientists made an exciting announcement: Kepler had discovered 1,284 new exoplanets. This is the largest number of planets ever announced at one time. Prior to this announcement, NASA had only confirmed the existence of 984 exoplanets in our galaxy, most of which were larger than the mass of Jupiter (Schneider, 2009). Detecting an exoplanet is a very difficult task because they don't emit any electromagnetic radiation of their own and are completely obscured by their extremely bright host star, meaning that ordinary telescope observation techniques are not available. Therefore, to search for exoplanets, a variety of techniques such as radial velocity, pulsar timing, astrometry, gravitational lensing, spectroscopy and photometry are used (DePater and Lissauer, 2001). The main aim of any method would be to detect the effects of exoplanets on their star systems. In addition to these findings, it is important to look for models that can explain the origin, formation and possible migration of these objects. For example, Rice and Armitage(2005) investigated how statistical distributions of extrasolar planets could be combined with knowledge of host star metallicity to constrain the migration history of gas giant planets. In addition, in a series of papers (Udry et al., 2003; Santos and so on. , 2003; Eggenberger et al., 2004; (Halbwachs et al., 2005) studied the appearance nature of planet host stars and the distribution characteristics of elements in different orbits of exoplanet systems. In this work, we analyzed cross-sectional data of exoplanets detected prior to June 2008 using linear regression techniques. The purpose of this analysis is to verify the relationship between the host star and its orbiting planet. For example, if the mass of a planet depends heavily on the type of star, it affects the formation stage of the planet.

The exoplanet catalog was created in February 1995 in order to facilitate the progress of the new field named Exoplanetology through the publication of recent detections and their associated data. The catalog is interactive and it is available in the webpage: http://exoplanet.eu. With the collected data, analysis can be conducted to identify and analysis dynamical data of confirmed exoplanets.

In this paper, we seek to investigate the properties of exoplanets and propose a general model that uses statistical techniques to explain the relationship between exoplanets detected as of April 2022 and the major features of their host stars. We will be looking at the set of properties of exoplanets such as mass and radius first, followed by proper motion. Then, we will proceed by building a linear mixed model with mass of planet as the response and radius, mass, temperature, age and other predictors. We found that ...

The remaining part of the paper is divided as follows: Section 2 explains where our data comes from and gives a general idea of the variables present in the dataset. Section 3 shows the model that explains how mass of exoplanet is affected by different factors. Section 4 explains the key findings of the data analysis and model. Section 5 expands on what is found and why it is important. Section 5.4 contains supplementary graphs that support the arguments in the discussion section.

2 Data

2.1 Data Source

In paper we run our analysis in `R` (R Core Team 2020).

2.2 Load Libraries

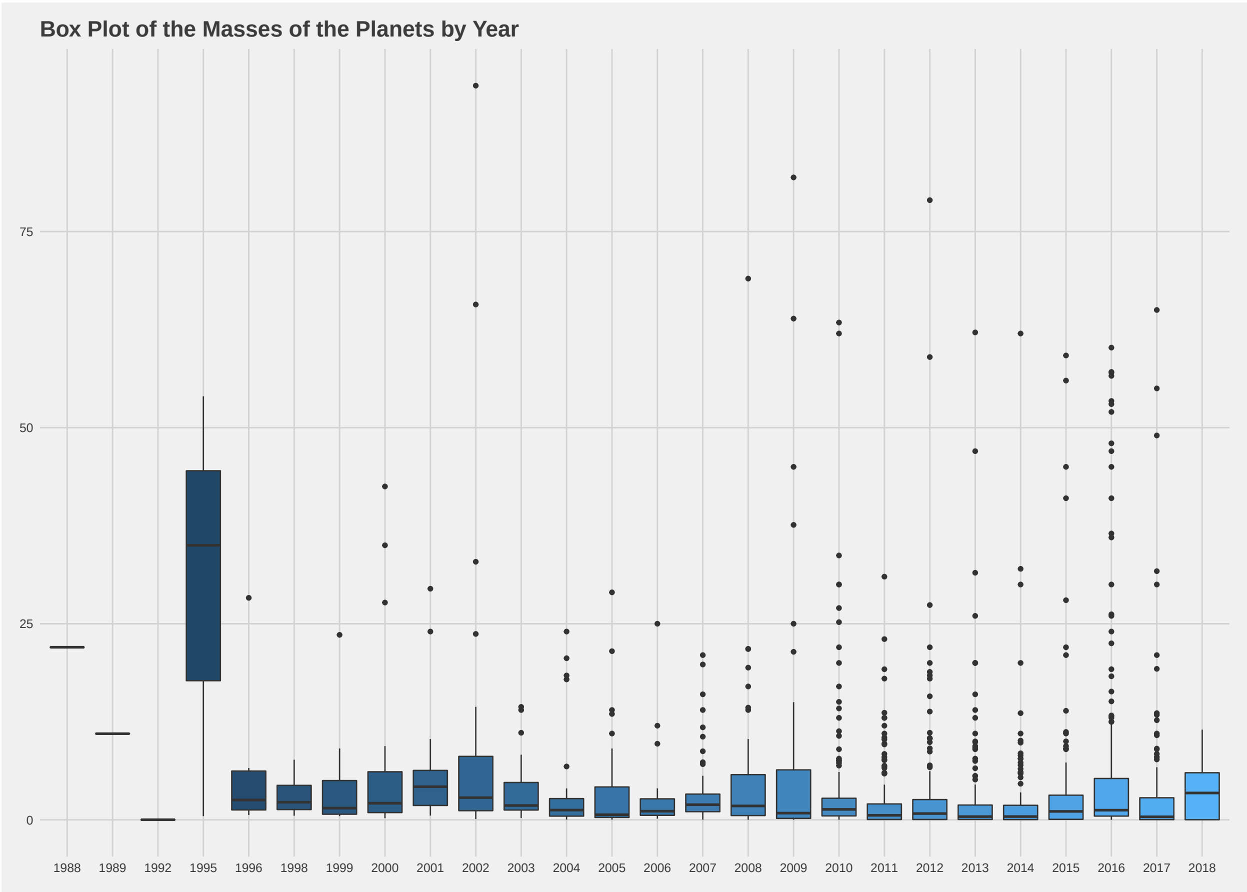
Loading a few key libraries

2.3 Load Data

Load the data and take a look at a few observations

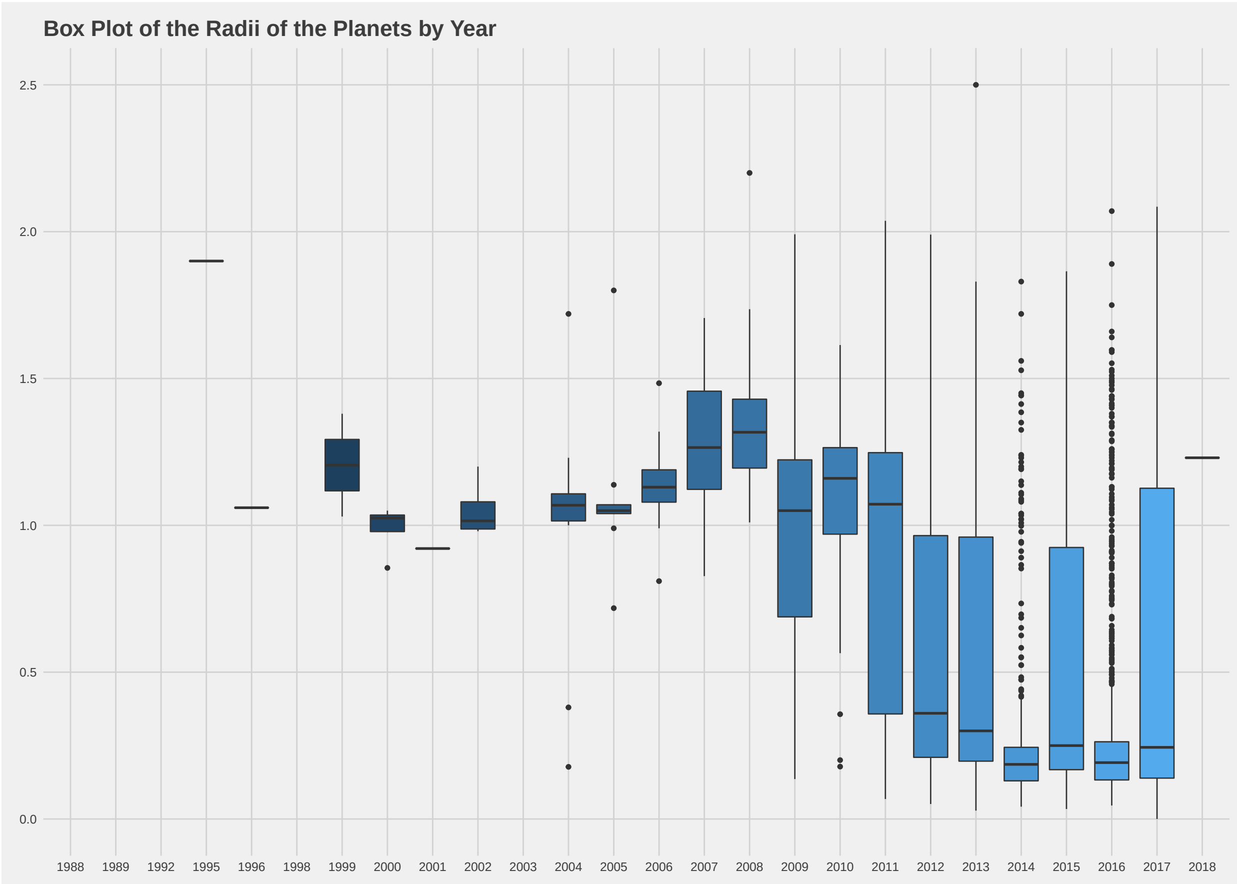
2.4 Data Visualizations

2.4.1 Planet Mass by Year



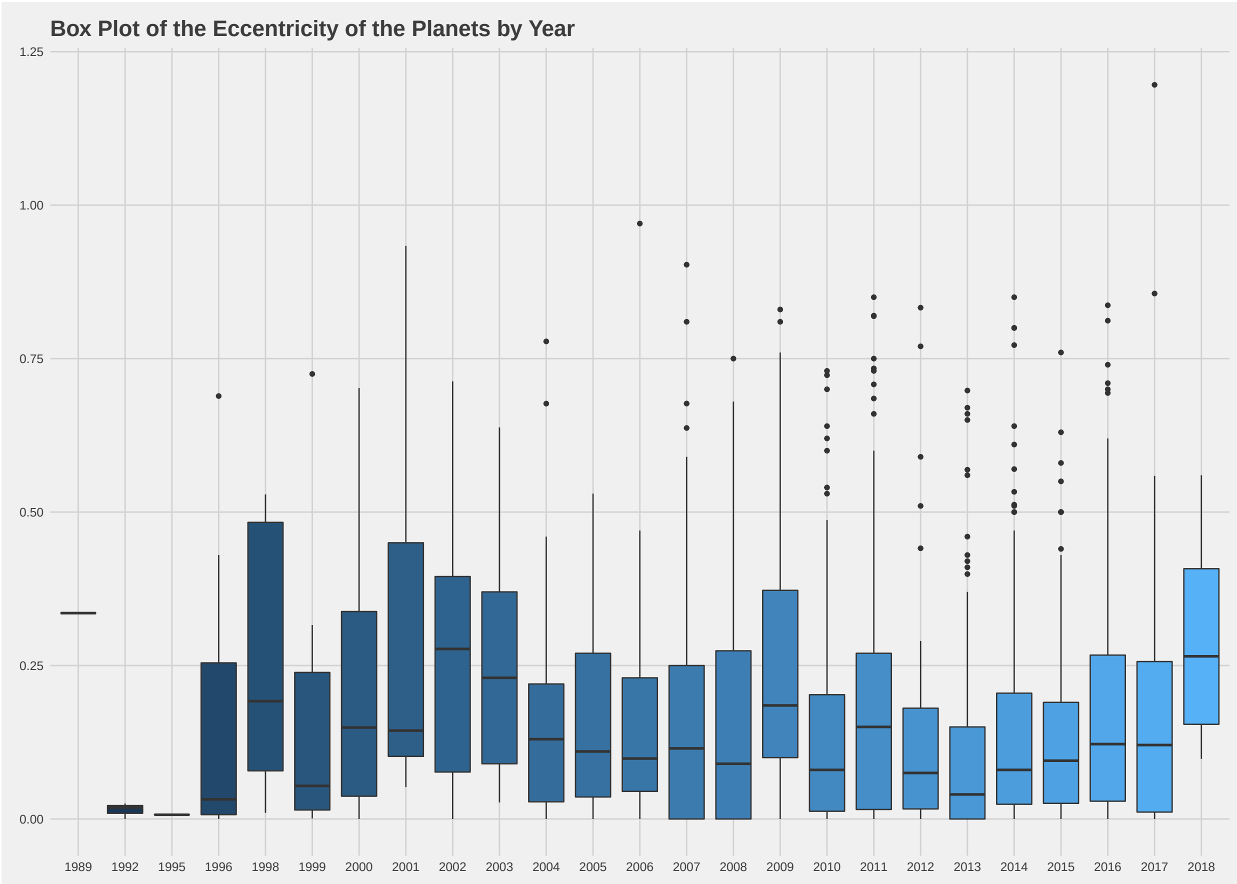
1995 was a significant year with seeing the biggest exoplants on average and then the planets went on a diet. There was not much of a change in the average masses of the planets discovered throught the years.

2.5 Planet Radius by Year



The average radius of the planets discovered through the years has been decreasing. A hypothesis is that maybe with better instruments we are able to measure the radii better.

2.6 Planet Eccentricity by Year



3 Model

4 Results

5 Discussion

5.1 First discussion point

5.2 Second discussion point

5.3 Third discussion point

5.4 Weaknesses and next steps

Appendix

A Additional details

References

R Core Team. 2020. *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.