



UNIVERSIDAD DE ANTIOQUIA
INSTITUTO DE FÍSICA
FACULTAD DE CIENCIAS EXACTAS Y NATURALES
GRUPO DE FÍSICA Y ASTROFÍSICA COMPUTACIONAL FACOM

Ajuste de la Función de Distribución de Probabilidad de las masas planetarias

TRABAJO DE INVESTIGACIÓN PARA OPTAR AL TÍTULO DE
MAGISTER EN FÍSICA

FABIO ONIER OSORIO PELÁEZ

Asesor: Pablo Andrés Cuartas Restrepo, PhD.

Co-asesor: Liliam Cardeño, PhD.

Medellín, 2022



**UNIVERSIDAD
DE ANTIOQUIA**
1 8 0 3

Universidad de Antioquia



Universidad de Antioquia
Instituto de Física
Facultad de Ciencias Exactas y Naturales
Grupo de Física y Astrofísica Computacional FCom

Ajuste de la Función de Distribución de Probabilidad de las masas planetarias

TRABAJO DE INVESTIGACIÓN PARA OPTAR AL TÍTULO DE
MAGISTER EN FÍSICA

FABIO ONIER OSORIO PELÁEZ

Asesor: Pablo Andrés Cuartas Restrepo, PhD.

Co-asesor: Liliam Cardeño, PhD.

.....
Pablo Andrés Cuartas Restrepo, PhD.

Medellín, 2022

Resumen

En este trabajo de investigación se estudió la función de distribución de las masas de los exoplanetas confirmados y reportados en los catálogos de exoplanetas, así como de las masas generadas a partir de un código de síntesis planetaria. Se proponen varias Funciones de Distribución de Probabilidad (FDP) y se desarrolla el proceso completo del método de bondad de ajuste basado en estimadores de máxima verosimilitud, prueba χ^2 y otros métodos de comparación para determinar a cuál de las FDP propuestas se ajustan mejor las masas planetarias. De los diferentes métodos de detección de exoplanetas solamente dos permiten estimar sus masas: el método de **Velocidad Radial** (VR) y el del **Tránsito Planetario** (TP). El primero permite estimar un valor **mínimo** de la masa del exoplaneta y el segundo estima el valor **total** de esta. Junto al análisis estadístico de las masas reportadas se desarrolla un código de síntesis planetaria, es decir, generación por computador de sistemas planetarios con el objetivo de obtener masas planetarias sintetizadas y realizar el mismo análisis estadístico con estas. Los resultados de ambos análisis se comparan entre sí para evaluar qué tanto se acercan los resultados del modelo computacional de formación planetaria a los de los datos observacionales. Los resultados de los ajustes de los datos de masas observacionales difieren de aquellos correspondientes de las masas sintetizadas. Finalmente se hace un análisis de aquellos factores que favorecen o desfavorecen el empleo de algunas pruebas de bondad de ajuste de acuerdo al tipo de datos con lo que se pretende trabajar.

Palabras clave

Exoplanetas, velocidad radial, tránsito planetario, función de densidad de probabilidad, estimación por máxima verosimilitud.

Para Jerónimo.

*No quiero que seas como yo ni que lo intentes.
Deseo que seas el mejor TÚ que puedas llegar a ser.*

Índice general

Resumen	I
1. Introducción	1
2. Formación y detección de planetas	5
2.1. Modelos de Formación Planetaria	5
2.1.1. La estructura y evolución del disco protoplanetario	6
2.1.2. La formación de planetecimales	8
2.1.3. Acreción de gas	10
2.1.4. Migración orbital	12
2.1.5. Interacciones N cuerpos	14
2.2. Estimación de la masa de los exoplanetas	14
2.2.1. Masas medidas por velocidad radial	14
2.2.2. Masas medidas por tránsito planetario	16
3. Síntesis de poblaciones planetarias	21
3.1. Modelo de Berna de formación y evolución planetaria	22
3.1.1. Generación I	23
3.1.2. Generación Ib	23
3.1.3. Generación II	24
3.1.4. Generación III	24
3.2. PSyCo: Planetary Synthesis Code	25
3.2.1. Estructura y evolución del disco protoplanetario	25
3.2.2. PSyCo: crecimiento de núcleos y migración tipo I	26
3.2.3. PSyCo: formación de planetas gigantes	27
3.2.4. PSyCo: migración tipo II	28
4. Métodos estadísticos para el ajuste de masas planetarias a una FDP	31
4.0.1. Estimadores de Máxima Verosimilitud	32
4.0.2. Prueba de bondad de ajuste	33
4.0.3. Prueba Chi-Cuadrado	34
4.0.4. Prueba de Kolmogorov-Smirnov	34
5. Resultados	37
5.1. Resultados de PSyCo	42

Conclusiones	45
A. Elementos de Estadística	53
A.1. Definición de conceptos	53
A.2. Pruebas de hipótesis	54
A.2.1. Histograma	55
Agradecimientos	57
Referencias	67
Abreviaciones	69

Índice de figuras

2.1.	<i>51 Pegasi b</i> fue el primer exoplaneta detectado y confirmado. Los puntos de la gráfica corresponden a medidas reales de la velocidad radial de la estrella 51 Pegasi [The Planetary Society, 2022].	15
2.2.	Esquema de la geometría básica en la que se basa el método de la velocidad radial.	16
2.3.	Variaciones en el flujo $\frac{\Delta F}{F}(\lambda)$, producidas por la opacidad, dependiendo de la longitud de onda durante un tránsito planetario. Imagen tomada de [de Wit and Seager, 2013a].	17
2.4.	Fundamentos del espectro de transmisión (la atmósfera planetaria ha sido aumentada para mejorar la visibilidad) [de Wit and Seager, 2013a].	18
4.1.	Izquierda: ilustración del estadístico de Kolmogorov–Smirnov. La línea roja muestra la función de distribución acumulada teórica, la azul la función de distribución acumulada empírica, y la flecha negra es el estadístico K-S. Derecha: ilustración del estadístico de Kolmogorov–Smirnov entre dos muestras. Las líneas roja y azul muestran la función de distribución acumulada empírica de dos muestras, y la flecha negra es el estadístico K-S.	35
5.1.	Histogramas de los dos grupos de masas observadas. <i>Izquierda</i> : masas medidas por tránsito planetario. <i>Derecha</i> : masas medidas por velocidad radial.	39
5.2.	Histogramas de masas observadas comparadas con las gráficas de las FDP usadas en esta investigación. <i>Izquierda</i> : masas medidas por tránsito planetario. <i>Derecha</i> : masas medidas por velocidad radial.	40
5.3.	Áreas bajo las curvas de las diferentes FDP comparadas con los histogramas de las masas observadas. <i>Izquierda</i> : masas medidas por tránsito planetario. <i>Derecha</i> : masas medidas por velocidad radial.	41
5.4.	Comparación de la FDAE de los conjuntos de masas con las FDA de las distribuciones propuestas para el ajuste. <i>Izquierda</i> : masas medidas por TP. <i>Derecha</i> : masas medidas por VR	42
5.5.	Histograma de las masas sintetizadas con PSyCo	43
5.6.	Histograma de las masas sintetizadas con PSyCo contra las FDP propuestas	44
5.7.	Comparación de las áreas bajo las curvas de las FDP propuestas vs el histograma de masas sintetizadas	45
5.8.	Comparación de la FDAE de las masas sintetizadas con las FDA de las distribuciones propuestas.	46

5.9.	Gráficos de dispersión de las masas observadas y sus radios orbitales. <i>Izquierda:</i> masa por tránsito planetario. <i>Derecha:</i> masas por velocidad radial.	46
5.10.	Síntesis planetaria con PSyCo. Masas vs semiejes mayores en una de las corridas del modelo.	47
5.11.	Evolución de la masa y la posición de planetas en el modelo PSyCo para diferentes masas estelares: (a). $0,1M_{\odot}$, (b). $0,3M_{\odot}$, (c). $0,5M_{\odot}$, (d). $0,8M_{\odot}$, (e). $1,0M_{\odot}$ y (f). $1,2M_{\odot}$	47
A.1.	(a). Histograma de densidades de 1000 números generados aleatoriamente a partir de una distribución normal. (b) Gráfica de una FDP normal. (c) Histograma de densidades de (a) y la FDP de (b) superpuestos [Dekermenjian, 2020].	55

Índice de tablas

2.1.	Valores de A_λ y B de acuerdo con el proceso de extinción.	19
5.1.	Contenidos de los más grandes catálogos de exoplanetas.	38
5.2.	Estimadores de los parámetros de las distribuciones utilizadas para el ajuste de los datos.	39
5.3.	Valores del estadístico de prueba χ^2 calculado para los dos conjuntos de datos y las tres distribuciones propuestas para el ajuste.	40
5.4.	Valores del <i>valor-p</i> calculado para los dos conjuntos de datos y las tres distribuciones propuestas para el ajuste.	42
5.5.	Valores del estadístico D de la prueba Kolmogorov-Smirnov para las tres distribuciones propuestas.	42
5.6.	Valores p de la prueba Kolmogorov-Smirnov para las tres distribuciones propuestas.	43
5.7.	Estimadores, valores del estadístico de prueba χ^2 y valores p de las pruebas χ^2 , para las distribuciones utilizadas en el ajuste de los datos de PSyCo.	43
5.8.	Estadístico D y <i>valor-p</i> de la prueba de Kolmogorov-Smirnov para evaluar el ajuste de las masas sintetizadas a las distribuciones propuestas	44

Introducción

La investigación en el área de las ciencias planetarias inició, como habría de esperarse, tratando de explicar las características de nuestro sistema solar: los diferentes tipos de planetas clasificados según su masa, su distribución espacial y otros aspectos observados como la presencia o ausencia de una atmósfera, agua en estado líquido o sólido en su superficie y existencia o no de formas de vida en ellos. La confirmación del primer exoplaneta (planeta que orbita a una estrella diferente al Sol) [Mayor and Queloz, 1995] dió paso a nuevas líneas de investigación para explicar, por ejemplo, planetas gigantes gaseosos orbitando cerca de su estrella. Estos descubrimientos obligaron a pensar en la concurrencia de diferentes fenómenos físicos y químicos implicados en la formación de planetas individuales y sistemas planetarios completos. Uno de estos fenómenos es, por ejemplo, el de la migración planetaria: los planetas no necesariamente terminan su proceso de formación la misma posición orbital donde empezaron.

A la fecha del 16 de agosto de 2022 se han confirmado entre 5000 y 5100 (dependiendo del catálogo que se consulte) exoplanetas y cada nueva confirmación trae consigo nuevos interrogantes acerca de los procesos de formación planetaria. Estas preguntas están relacionadas con la clasificación del planeta según su masa y tamaño (rocoso, gigante gaseoso o helado), posición orbital, habitabilidad, composición atmosférica, procesos de migración planetaria, presencia o no de satélites naturales alrededor, entre otras. Ante la falta de nuevas técnicas, instrumentos y métodos para obtener datos relevantes que permitan responder estas preguntas, la mejor estrategia de investigación actual yace en la síntesis planetaria.

Los códigos de síntesis planetaria pretenden simular las condiciones físicas reales de los procesos de formación de sistemas planetarios como:

- masa y metalicidad de la estrella,
- relación entre las masas del disco protoplanetario y la estrella,
- perfil de densidad de masa del disco protoplanetario,
- tasas de fotoevaporación del disco en su interacción con la radiación de la estrella,
- viscosidad del disco,
- posición de los planetecimales,
- tasas de acreción de gas y polvo según la posición orbital,

-
- interacción entre los planetecimales y del disco,
 - procesos de colisión entre planetecimales de forma concurrente con los procesos de acreción.

Estas consideraciones permiten comprender, de manera general, la diversidad de sistemas planetarios observados y responder varias de las preguntas generadas por estas observaciones. De esta forma, la observación y medición de diferentes propiedades de los exoplanetas y los resultados de los códigos de síntesis planetaria se convierten en líneas de investigación mutuamente complementarias que, además, se retroalimentan constantemente: las observaciones imponen restricciones los posibles resultados de la síntesis planetaria y los resultados de síntesis planetaria logran dar cuenta de algunas de las observaciones; las nuevas observaciones sugieren incorporar nuevos micro-modelos de procesos de formación para dar cuenta de estas observaciones.

Este trabajo de investigación pretende responder la pregunta de ¿cómo es la distribución de las masas de los planetas en el universo? o, expresada de otra manera ¿cuál es la función de densidad de probabilidad que describe la distribución de masas planetarias en el universo? Para ello, deben hacerse algunas consideraciones: la primera tiene que ver con la cantidad de datos observacionales y el origen de los mismos; debido a que hasta ahora solo ha sido posible detectar exoplanetas dentro de la Vía Láctea y que no hay razón para sospechar que los procesos físicos implicados en la formación de planetas sea diferente en las demás galaxias, el conjunto de datos observacionales se asumirá como una muestra de la población de planetas en el universo. La segunda consideración es consecuencia de la primera: debido a que los procesos físicos de formación de planetas siguen las mismas leyes en todo el universo, los datos generados por síntesis planetaria serán asumidos como representativos a nivel universal, lo que nos permite considerar sus resultados como una muestra válida proveniente de una población que sería el universo. Adicionalmente, los códigos de síntesis planetaria permiten obtener una cantidad de datos mayor a la de las observaciones, lo que arroja resultados estadísticos más robustos y confiables.

Conocer la distribución de las masas de los planetas permite mejorar los modelos de síntesis planetaria. Los códigos de síntesis planetaria permiten conocer mejor los procesos y subprocesos que favorecen la formación de ciertos tipos de planetas en lugar de otros y, a su vez, hacer mejores estimaciones sobre la probabilidad de hallar planetas en zona de habitabilidad; estas estimaciones permiten, por ejemplo, restringir las estrellas candidatas para la búsqueda de exoplanetas. Trabajos anteriores en esta línea han sido realizados por [Ananyeva et al., 2020, Neil and Rogers, 2020, Ivanova et al., 2019, Marcy et al., 2005, Jorissen et al., 2001, Uzsoy et al., 2021, Marcy et al., 2003, 2010].

El objetivo general de esta investigación es la determinación de la función de densidad de probabilidad de las masas planetarias. Para lograrlo se implementaron dos estrategias: la primera consistió en tomar los datos de las masas reportadas de exoplanetas, estimadas mediante los métodos de velocidad radial y tránsito planetario, para determinar el ajuste de datos a varias FDP propuestas. La segunda estrategia consistió en tomar los datos de planetas sintetizados mediante PSyCo y ajustarlas a las mismas funciones de distribución.

Teniendo en cuenta lo anterior, el siguiente trabajo de investigación está construido de la siguiente manera:

- En el capítulo 2 se exponen las características comunes existentes entre los diferentes modelos

de formación planetaria y sus etapas. Así mismo, se explican los métodos de estimación de masas de exoplanetas a partir de los datos de detección por velocidad radial y tránsito planetario.

- En el capítulo 3 se explican las diferentes generaciones en las cuales se han clasificado los modelos de síntesis planetaria y las características de PSyCo (Planet Synthesis Code), el código de síntesis planetaria desarrollado durante esta investigación.
- En el capítulo 4 se explican los métodos estadísticos utilizados en las pruebas de bondad de ajuste de los datos observacionales y sintéticos a las diferentes funciones de densidad de probabilidad propuestas.
- En el capítulo 5 se muestran los resultados obtenidos en las pruebas de bondad de ajuste de los conjuntos de datos utilizados.
- Finalmente, se exponen las conclusiones a las que se llegó a partir de los resultados y se proponen líneas de trabajo a futuro que podrán avanzar y profundizar en esta área de investigación en ciencias planetarias.

Formación y detección de planetas

Las *ciencias planetarias* son un área interdisciplinar de investigación en astrofísica que involucra la participación de diferentes ramas de las ciencias exactas y naturales como la biología, química, geología, estadística y computación. Dentro de los diferentes tópicos de investigación en ciencias planetarias se encuentran la formación de sistemas planetarios, planetas rocosos, gigantes gaseosos, estructura interna de los planetas, condiciones de habitabilidad, métodos de detección de exoplanetas, entre otros.

Este trabajo de investigación se centra en el análisis de la distribución de las masas planetarias a partir de los datos reportados de exoplanetas confirmados. Debido a las limitaciones observacionales (distancia a la que se encuentran las estrellas, luminosidad, condiciones atmosféricas, instrumentos de detección y análisis de luz) que tiene la detección de exoplanetas, los datos obtenidos tienen un sesgo difícil de evadir. Teniendo en cuenta esto, se hacen necesarios elementos adicionales que permitan, de alguna manera, resolver parte del sesgo o, incluso, sacar provecho del mismo para validar modelos teóricos.

2.1. Modelos de Formación Planetaria

Gracias a los avances logrados en los métodos de detección de exoplanetas, somos la primera generación de seres humanos con la capacidad tecnológica para responder la pregunta de si existen o no otros planetas orbitando otras estrellas [Mayor and Queloz, 1995]. Estos avances nos llevan a pensar que estrellas con sistemas planetarios parecen ser la regla en el universo y no la excepción. A pesar de estos importantes avances aún no son completamente claros, desde el punto de vista teórico, los procesos y mecanismos mediante los cuales se forman los diferentes tipos de planetas y las distribuciones espaciales que tienen dentro de sus propios sistemas planetarios. De hecho, los diferentes modelos de formación planetaria propuestos no logran dar cuenta de algunas de las propiedades de los exoplanetas observados, bien sea porque están fuera del alcance de los modelos o bien porque los contradicen. Es por esta razón que los avances en los métodos y técnicas observacionales en la búsqueda de exoplanetas son importantes puesto que la detección de un número mayor de estos y su caracterización nos imponen cada vez más restricciones a los modelos de formación de planetas, de tal manera que las predicciones sean más consistentes con las observaciones. Dos de estas observaciones que imponen restricciones a los modelos de formación y síntesis de poblaciones planetarias son las relaciones *masa–semi eje mayor* y *masa–radio planetario*.

Podríamos decir que existen modelos *globales* [Mordasini et al., 2015] de formación planetaria

en cuanto que todos ellos describen las mismas etapas del proceso; las diferencias entre los modelos radican en los mecanismos y submodelos específicos de cada etapa y los tiempos de duración de las mismas. De manera general, estas etapas y mecanismos son [Raymond and Morbidelli, 2022], [André et al., 2014]:

- la estructura y evolución del disco protoplanetario;
- la estructura y evolución del disco de sólidos (polvo, guijarros y planetecimales);
- acreción de sólidos que llevan al crecimiento del núcleo sólido planetario;
- acreción de H/He que lleva a la formación de una envoltura gaseosa alrededor del planeta;
- migración orbital como consecuencia del intercambio de momento angular;
- interacciones de N-cuerpos entre los planetas y protoplanetas.

Para cada una de estas etapas se han desarrollado diferentes submodelos de los procesos físicos [Armitage, 2020] que las gobiernan y todos ellos logran dar cuenta de la mayoría de las observaciones, pero no siempre las propiedades planetarias modeladas son las mismas. “*Adicionalmente, los submodelos pueden describir la estructura interna de los núcleos con envolturas, la estructura de la atmósfera planetaria, la interacción entre planetecimales con guijarros y la envoltura de protoplanetas, la evolución de la estrella o la pérdida de la envoltura gaseosa durante las fases de evolución como, por ejemplo, el escape atmosférico*” [Jin et al., 2014, Mordasini, 2018].

Los submodelos empleados en cada una de las etapas de la formación y evolución de sistemas planetarios son de dimensiones espaciales bajas. En su mayoría, emplean ecuaciones de 1D espacial debido a que modelos en 2 y 3 dimensiones son muy costosos computacionalmente. Algunos de estos submodelos omiten ciertos procesos físicos relacionados con la dinámica de fluidos, interacción del gas del disco con el campo magnético de la estrella o entre los sólidos. El objetivo de estos submodelos es arrojar luces sobre lo que ocurre en 2D y 3D a partir de los resultados en 1D.

2.1.1. La estructura y evolución del disco protoplanetario

El punto de partida para la formación de cualquier sistema planetario es la formación de la estrella alrededor de la cual orbitarán los planetas. La teoría más ampliamente aceptada es la de la **Hipótesis nebular** según la cual las estrellas nacen en el interior de nebulosas moleculares gigantes, donde una sección de la nebulosa comienza un colapso gravitacional sobre sí misma. No todo el gas que está colapsando logra llegar hasta el centro para formar la estrella: más o menos el 1% de la masa de gas y polvo originales del colapso se queda girando alrededor de la protoestrella en el mismo sentido de rotación de ésta. Esta rotación conjunta es la responsable que la masa de gas y polvo que no cae a la estrella comience a caer hacia un plano coincidente con el plano ecuatorial de la protoestrella formando un disco circumstelar conocido como **disco protoplanetario**.

Los modelos de disco describen la evolución de su perfil de densidad de masa, temperatura del gas, presión y estructura vertical a partir de un factor de escala vertical. El modelo de disco más utilizado es el de **disco delgado**, en el que la estructura horizontal (radial) es varios órdenes de magnitud mayor que la estructura vertical.

La manera más simple de parametrizar un modelo de disco de gas es usar una aproximación de ley de potencias inspirada en el modelo de **Nebulosa Solar de Masa Mínima** (MMSN o Minimun Mass Solar Nebula, en inglés) [Weidenschilling, 1977, Hayashi, 1981]. En el modelo MMSN la densidad superficial (inicial) Σ de gas es una función de la distancia r a la estrella y tiene la forma

$$\Sigma(r) = \Sigma_0 \left(\frac{r}{1\text{UA}} \right)^{-3/2}. \quad (2.1)$$

En síntesis de poblaciones, la constante de normalización Σ_0 varía como un Monte Carlo con el objetivo de representar discos de diferentes masas. En particular, para el MMSN, $\Sigma_0 \approx 2400 \text{ g/cm}^2$. En estos modelos, la temperatura T del disco también está dada por una ley de potencias. Por ejemplo, asumiendo un disco ópticamente delgado y una luminosidad estelar en la escala de secuencia principal en la forma $L \propto M_*^4$ [Ida and Lin, 2004b],

$$T(r) = 280\text{K} \left(\frac{r}{1\text{UA}} \right)^{-1/2} \left(\frac{M_*}{M_\odot} \right). \quad (2.2)$$

Este modelo de temperatura no tiene en cuenta discos ópticamente gruesos, calentamiento por viscosidad ni evolución temporal porque las estrellas ya no están en la secuencia principal cuando tienen discos de gas alrededor. Otros modelos [Chambers, 2009] consideran estos efectos y constituyen una mejora a los de MMSN.

Algo más complejo, pero aún en la aproximación 1D, es la ecuación de la evolución viscosa clásica ([Lüst, 1952, Lynden-Bell and Pringle, 1974]) para la densidad superficial del gas en función del tiempo t y de la distancia r a la estrella:

$$\frac{\partial \Sigma}{\partial t} = \frac{1}{r} \frac{\partial}{\partial r} \left[3r^{1/2} \frac{\partial}{\partial r} (r^{1/2} \nu \Sigma) \right] - \dot{\Sigma}_{\text{phot}}(r) - \dot{\Sigma}_{\text{planet}}(r), \quad (2.3)$$

donde la viscosidad ν está escrita en términos del parámetro α como $\nu = \alpha c_s H$, c_s es la rapidez del sonido y H es el factor de escala vertical [Shakura and Sunyaev, 1973]. Adicional al término de la evolución de la viscosidad se han tenido en cuenta la pérdida de masa por fotoevaporación $\dot{\Sigma}_{\text{phot}}(r)$ [Alexander et al., 2014] y por acreción de gas de los planetas $\dot{\Sigma}_{\text{planet}}(r)$.

Como condición inicial para esta ecuación se asume una densidad de gas consistente en un decrecimiento cerca de la estrella debido a la cavidad magnetosférica estelar, una ley de potencias en la parte principal y un decrecimiento exponencial más allá de un radio característico [Lynden-Bell and Pringle, 1974]. Esta densidad superficial inicial es entonces

$$\Sigma_g(t=0, r) = \Sigma_0 \left(\frac{1}{1\text{UA}} \right)^{p_g} \exp \left[- \left(\frac{r}{R_{\text{out}}} \right)^{2+p_g} \right] \left(1 - \sqrt{\frac{r}{R_{\text{in}}}} \right). \quad (2.4)$$

en donde R_{out} es el radio externo al radio “característico”, R_{in} el radio interno y p_g el exponente de la ley de potencias. Algunas observaciones [Andrews et al., 2010] indican que $p_g \approx -1$. Los cuatro parámetros en esta ecuación pueden ser tratados como variables aleatorias Monte Carlo en síntesis de poblaciones.

Suponiendo que el polvo se convierte en planetecimales con una eficiencia plena, la densidad

superficial inicial Σ_p de éstos estaría dada por [Mordasini et al., 2009b]

$$\Sigma_p(t = 0, r) = f_{dg}\eta_{ice}\Sigma_g(t = 0, r), \quad (2.5)$$

donde f_{dg} es la razón polvo-gas (\approx a la fracción Z de la masa de elementos pesados) que, para el Sol, es aproximadamente 0,0149 aunque también es una variable Monte Carlo que representa las diferentes metalicidades de las estrellas; η_{ice} representa la disminución de la densidad superficial de sólidos en la línea de hielos.

De acuerdo con algunas observaciones [Panić et al., 2009] y resultados teóricos [Birnstiel et al., 2012] podría presentarse una redistribución radial significativa de sólidos en forma de guijarros, lo que llevaría a distribuciones más concentradas de sólidos que la indicada por la ecuación 2.5.

La evolución temporal del disco se encuentra resolviendo la ecuación 2.3 teniendo en cuenta la evolución viscosa y la fotoevaporación. Otros modelos parametrizados como el de [Ida and Lin, 2004b] usan, en su lugar, una ecuación de la forma

$$\dot{\Sigma}_g(r) = -\frac{\Sigma_g}{\tau_{disk}} + \dot{\Sigma}_{phot}, \quad (2.6)$$

donde el primer término lleva a un decaimiento exponencial y el segundo imita los efectos de la fotoevaporación. La escala de tiempo τ_{disk} del disco puede ser tratada como una variable Monte Carlo.

La densidad superficial de sólidos decrece con la zona de alimentación de acuerdo con la cantidad de masa que acreta el planeta, suponiendo que la densidad superficial es uniforme dentro de esta zona: [Thommes et al., 2003]

$$\dot{\Sigma}_p = -\frac{(3M_*)^{1/3}}{6\pi a_p^2 B_L M_p^{1/3}} \dot{M}_c, \quad (2.7)$$

donde B_L es el ancho de la zona de alimentación en esferas de Hill, M_p la masa del planeta, a_p su semieje mayor y \dot{M}_c la tasa de acreción de planetecimales del planeta.

2.1.2. La formación de planetecimales

A medida que el disco evoluciona dinámicamente comienzan a formarse acumulaciones de materia cada vez más grandes: desde granos de micrómetros hasta alcanzar objetos de miles de kilómetros de diámetro. Aunque los procesos de la evolución de estos granos de polvo hasta alcanzar tamaños entre los 10 km y 100 km de diámetro aún es motivo de debate [Cuzzi et al., 2008], la mayoría de investigadores están de acuerdo en que estos granos interactúan con otros y con el gas. En esta interacción empiezan a acretarse a través de procesos colisionales no destructivos hasta formar lo que se conoce como **planetecimales**: cuerpos rocosos con tamaños desde metros hasta los cientos de kilómetros. La cantidad y el tamaño de los planetecimales formados depende de características del disco tales como su estructura y densidad superficial [Hayashi, 1981], viscosidad [Ida and Makino, 1993], composición química, volatilidad, procesos de fotoevaporación causados por la radiación de la protoestrella y rayos cósmicos externos, entre otros.

En modelos basados en planetecimales se asume que el crecimiento de la masa M_c del núcleo

sólido se da a través de la acreción de otros sólidos y este crecimiento está dado por [Safronov, 1969]

$$\dot{M}_c = \Omega \Sigma_p R_{capture}^2 F_G, \quad (2.8)$$

donde Ω es la frecuencia kepleriana, Σ_p es la densidad superficial promedio de planetecimales en la zona de alimentación; $R_{capture}$ es el radio de captura de planetecimales que, en general, es mayor que el radio del núcleo debido al arrastre del gas ([Podolak et al., 1988, Arnold et al., 2006]) y F_G es el factor de enfoque gravitacional ([Nakazawa et al., 1989, Greenzweig and Lissauer, 1992]. \dot{M}_c también depende de otras cantidades como las velocidades aleatorias v_{pls} de los planetecimales y, en el caso ideal de un sistema de dos cuerpos, estaría dada por $1 - (v_{esc}/v_{pls})^2$ donde v_{esc} es la velocidad de escape respecto a los planetecimales [Safronov, 1969].

Una visión de cómo la tasa de acreción del núcleo depende de otros parámetros se puede lograr considerando la escala del tiempo de acreción τ_c :

$$\dot{M}_c = \frac{M_c}{\tau_c}. \quad (2.9)$$

Una expresión aproximada de la escala de tiempo de acreción en el régimen oligárquico fue derivada por [Ida and Lin, 2004b] a partir del trabajo de [Kokubo and Ida, 2002]. Este régimen ocurre después del crecimiento desbocado, cuando los planetecimales han alcanzado tamaños entre los 100 km y los 1000 km, dependiendo de la distancia orbital [Ormel et al., 2010]. En este régimen las velocidades aleatorias de los planetecimales se ven modificadas por las perturbaciones viscosas de los planetecimales y el arrastre del gas aún presente en el disco. Esta escala de tiempo de acreción está dada por

$$\tau_c = 1,2 \times 10^5 \text{yr} \left(\frac{\Sigma_p}{10 \text{g cm}^{-2}} \right)^{-1} \left(\frac{a_p}{1 \text{UA}} \right)^{1/2} \left(\frac{M_c}{M_\oplus} \right)^{1/3} \left(\frac{M_*}{M_\odot} \right)^{-1/6} \times \left[\left(\frac{\Sigma_g}{2400 \text{g cm}^{-2}} \right)^{-1/5} \left(\frac{a_p}{1 \text{UA}} \right)^{1/20} \left(\frac{m}{10^{28} \text{g}} \right)^{1/15} \right]^2. \quad (2.10)$$

En esta ecuación Σ_g es la densidad superficial de gas en la posición a_p del planeta y m es la masa del planetesimal. Adicionalmente, esta ecuación permite explicar, entre otras cosas, por qué la teoría de acreción de núcleos predice más planetas gigantes cuando la metalicidad [Fe/H] de la estrella es mayor.

Durante estas etapas también empiezan a presentarse colisiones entre planetecimales. Algunas de estas colisiones son inelásticas y su efecto es que los cuerpos terminan acretados; otras colisiones fragmentan los cuerpos de tal forma que no vuelven a formarse como uno solo y otras causan dispersión de los mismos. El resultado de estas colisiones es que aquellos cuerpos de mayor masa, llamados **embriones planetarios**, comienzan a acretar y crecer más rápido que los demás en alguno de dos mecanismos: en el primero, conocido como crecimiento *desbocado*, ([Greenberg et al., 1978, Wetherill and Stewart, 1993, Kokubo and Ida, 1996]) el planetesimal se mueve con una velocidad de escape mayor que la velocidad relativa de los cuerpos que lo rodean; en el segundo, conocido como crecimiento *oligárquico* [Kokubo and Ida, 1998], la velocidad de dispersión de los planetecimales se vuelve del orden de la velocidad de escape y los planetecimales más pequeños no se acretan entre ellos sino que se acretan al cuerpo más masivo a causa del enfoque gravitacional

[Kokubo and Ida, 2000]. A los cuerpos sobrevivientes a esta etapa también se les da el nombre de **protoplanetas**.

Existe otro modelo, más eficiente, que explica el crecimiento de los planetecimales hasta convertirse en protoplanetas: la acreción de **guijarros** (en inglés, **pebbles**). Estos guijarros son cuerpos del orden de los cm de diámetro que se acretan mutuamente por su interacción gravitacional. Una vez que se ha formado un planetesimal este sigue dentro del disco de gas y guijarros acretando individualmente a estos últimos [Ormel and Klahr, 2010, Lambrechts and Johansen, 2012, 2014]. Este modelo permite explicar la formación de cuerpos de hasta decenas de veces la masa de la Tierra (*súper* Tierras) que conforman los núcleos de los gigantes gaseosos [Lambrechts and Johansen, 2012]. Esta acreción de guijarros no puede ocurrir indefinidamente porque cuando un planeta se vuelve suficientemente masivo comienza a abrir una brecha (*gap*, en inglés) en el disco y se dice que ha alcanzado su **masa de aislamiento** [Morbidelli and Nesvorny, 2012].

2.1.3. Acrción de gas

Luego de haberse formado los protoplanetas estos adquieren la masa suficiente (masa crítica) para empezar a acretar el gas que los rodea y formar una envoltura a su alrededor. A medida que van acretando más gas se vuelven más masivos y, por tanto, adquiriendo mayor capacidad de acreción por interacción gravitacional con el gas del disco y algunos terminan formando una atmósfera. Dependiendo de la distancia orbital a la que se encuentren estos cuerpos tendrán más o menos gas disponible para formar atmósferas grandes hasta convertirse en gigantes gaseosos. Los cuerpos relativamente cercanos a la estrella no tienen mucho gas disponible y terminan como planetas rocosos, algunos con atmósferas. Otros, en tanto, debido a su lejanía de la estrella, no tienen gas disponible para formar grandes atmósferas sino que encuentran a su alrededor hielos que acretan hasta convertirse en planetas tipo Neptuno [Hayashi, 1981]. Mientras esto ocurre se reduce la cantidad de gas disponible en el disco, bien sea porque ha sido acretado por los protoplanetas o bien sea por efecto de fotoevaporación y la actividad estelar.

En los modelos globales de formación planetaria se tienen dos aproximaciones a la tasa de acreción de gas. En algunos de ellos [Alibert et al., 2005, Mordasini et al., 2012b] se calcula la estructura interna del protoplaneta mediante las ecuaciones de [Bodenheimer and Pollack, 1986]

$$\frac{\partial m}{\partial r} = 4\pi r^2 \rho \quad (2.11)$$

$$\frac{\partial P}{\partial r} = -\frac{Gm}{r^2} \rho \quad (2.12)$$

$$\frac{\partial l}{\partial r} = 4\pi r^2 \rho \left(\varepsilon - P \frac{\partial V}{\partial t} - \frac{\partial u}{\partial t} \right) \quad (2.13)$$

$$\frac{\partial T}{\partial r} = \frac{T}{P} \frac{\partial P}{\partial r} \nabla(T, P) \quad (2.14)$$

que corresponden a la conservación de la masa, equilibrio hidrostático, conservación de la energía y ecuaciones de transporte de la energía. En estas, r es el radio medido desde el centro del planeta, m es la masa encerrada en ese radio, P es la presión, ρ la densidad y G es la constante de gravitación universal. El gradiente ∇ depende de si el transporte de energía es por convección o radiación. La ecuación de la energía es la única dependiente del tiempo t y controla la evolución

temporal. $V = 1/\rho$ es el volumen específico, u la energía interna, ε es una fuente de energía como calentamiento radiogénico o por impactos y l es la luminosidad intrínseca.

Estas ecuaciones de estructura se resuelven con diferentes condiciones de frontera que dependen de la fase de formación en la que se encuentra el protoplaneta [Bodenheimer et al., 2000, Mordasini et al., 2012b]. En la primera fase, llamada “nebular”, la envoltura gaseosa está inicialmente en equilibrio térmico y de presión con el gas del disco y luego sufre un enfriamiento radiativo que la hace contraer. Al contraerse deja una brecha entre la envoltura y el gas del disco que luego es vuelto a llenar por nuevo gas de la nube. Durante esta fase, la tasa de acreción está regulada por la escala de tiempo (Kevin-Helmholtz) de enfriamiento de la envoltura.

Cuando el núcleo alcanza una masa de alrededor de $10 M_{\oplus}$ la contracción de la envoltura gaseosa es tan rápida que el disco protoplanetario ya no es capaz de suministrar gas a una tasa que permita que la envoltura y el disco se mantengan en contacto y el radio externo del planeta se desliga de la nube de gas. En esta segunda fase el radio r de las ecuaciones queda libre y se calcula resolviendo las ecuaciones de estructura mientras que la tasa de acreción queda dependiendo de procesos en el disco protoplanetario y no de la contracción de la envoltura gaseosa. Esta tasa limitada por el disco puede estar dada por la tasa de acreción de Bondi [D’Angelo and Lubow, 2008, Mordasini et al., 2012b]

$$\dot{M}_{e,Bondi} \approx \frac{\Sigma_g}{H} \left(\frac{R_H}{3} \right)^3 \Omega, \quad (2.15)$$

donde Σ_g , H , R_H y Ω son la densidad superficial promedio de gas en la zona de alimentación, el factor de escala vertical, el radio de la esfera de Hill del planeta y la frecuencia de Kepler en la posición del planeta, respectivamente.

La segunda aproximación a la tasa de acreción [Ida and Lin, 2004b, Hasegawa and Pudritz, 2012, Ndugu et al., 2018] consiste en calcular esta con base en ajustes a la escala de tiempo de Kevin-Helmholtz. En esta aproximación la tasa de acreción debida a la contracción de la envoltura está dada por

$$\dot{M}_{e,KH} = \frac{M_p}{\tau_{KH}} \quad (2.16)$$

donde la escala de tiempo de enfriamiento de Kevin-Helmholtz de la envoltura está parametrizada como [Ikoma et al., 2000]

$$\tau_{KH} = 10^{p_{KH}} \text{yr} \left(\frac{M_p}{M_{\oplus}} \right)^{q_{KH}} \left(\frac{\kappa}{1 \text{g cm}^{-2}} \right); \quad (2.17)$$

p_{KH} y q_{KH} son parámetros que se obtienen ajustando la tasa de acreción obtenida con cálculos de estructura interna. De acuerdo con esta ecuación la tasa de acreción de gas es una función creciente de la masa y se vuelve importante cuando τ_{KH} es comparable o menor que el tiempo de vida del disco.

Hay pocos planetas en el rango de masas intermedias entre $10 M_{\oplus}$ y $100 M_{\oplus}$ debido a que las escalas de tiempo para acretar esta masa de gas en acreción desbocada son más cortos que el tiempo de vida del disco, por lo que es improbable que el disco desaparezca exactamente en los momentos de masa intermedia.

2.1.4. Migración orbital

La migración orbital consiste en que el planeta cambia su radio orbital, bien sea aumentándolo o disminuyéndolo y se conocen varios mecanismos por los cuales puede ocurrir este fenómeno. El principal mecanismo por el que se presenta la migración orbital, al menos para los planetas rocosos, se debe a las ondas de densidad gravitacional [Goldreich and Tremaine, 1979, 1980, Raymond and Morbidelli, 2022] que se forman en el disco. La onda exterior sigue al planeta y lo ralentiza, generando un torque negativo. La parte interior de la onda conduce al planeta y genera un torque positivo. El desbalance entre estos dos torques de signo contrario es el principal mecanismo que da origen a la migración orbital. En cualquier disco de simetría radial cuyo perfil de densidad es cualquier ley de potencias radial el torque negativo vence [Ward, 1986] y el efecto es que el planeta pierde momento angular y migra hacia la estrella. Esta es la migración **tipo I** y solamente la tienen los planetas de masas pequeñas.

El momento angular J de un planeta de masa M_p que orbita a una estrella de masa M_* en un semieje mayor a_p y la tasa de migración da/dt causada por un torque total $\Gamma_{tot} = dJ/dt$ son

$$J = M_p \sqrt{GM_* a_p} \quad \frac{da}{dt} = 2a_p \frac{\Gamma_{tot}}{J}. \quad (2.18)$$

El cálculo del torque causante de la migración del planeta requiere, entre otras cantidades, la densidad superficial de gas, el exponente de la ley de potencias local de la temperatura del disco p_T y de la densidad superficial del gas p_Σ , las cuales son producidas por el modelo del disco. La escala de tiempo de la migración en la aproximación isotérmica usada por [Ida and Lin, 2008] está dada por

$$\tau_{tipoI} = \frac{1}{2,728 + 1,082p_\Sigma} \left(\frac{c_s}{a_p \Omega} \right)^2 \frac{M_*}{M_p} \frac{M_*}{a_p^2 \Sigma_g} \Omega^{-1} \quad (2.19)$$

y la tasa de migración es

$$\dot{a}_p = \frac{a_p}{\tau_{tipoI}}, \quad (2.20)$$

lo que muestra que la migración acelera el crecimiento de un planeta para que se vuelva más masivo. El torque total está dado por [Paardekooper et al., 2010b]

$$\Gamma_{tot} = \frac{1}{\gamma} (C_0 + C_1 p_\Sigma + C_2 p_T) \Gamma_0, \quad (2.21)$$

con

$$\Gamma_0 = \left(\frac{q}{h} \right)^2 \Sigma_g a_p^4 \Omega^2, \quad (2.22)$$

donde γ es la razón de las capacidades caloríficas, $h = H/a_p$ es la razón de aspecto local del disco, $q = M_p/M_*$, Σ_g es la densidad superficial del gas en la posición del planeta y Ω es la frecuencia kepleriana. Las constantes C_i dependen del subrégimen Tipo I y sus valores numéricos se pueden encontrar listados, por ejemplo, en [Dittkrist et al., 2014].

Cuando el planeta se mueve en el disco lo hace atravesando el gas, lo que genera que porciones de este pasen del interior de la órbita al exterior y viceversa causando torques adicionales. Para discos con perfiles de densidad radial (la ecuación 2.4 muestra un perfil de densidad) proporcionales a $1/r^{3/2}$ estos torques se anulan entre sí [Masset et al., 2006]; para perfiles menos profundos, gana

el torque positivo y esto frena la migración tipo I [Paardekooper et al., 2010a]. Dependiendo de la pendiente del perfil de densidad del disco el torque de corrotación positivo puede vencer al negativo y revertir la migración [Masset et al., 2006]. Bajo estas condiciones existen algunas locaciones dentro del disco donde se forman *trampas* planetarias y se detiene la migración. También, la viscosidad del disco y fenómenos de calentamiento y enfriamiento del gas alrededor del planeta generan retroalimentaciones y retrocesos de la migración.

En el caso de los planetas más masivos (masas superiores a las 10 masas terrestres) la acreción de gas por parte del núcleo genera una atmósfera cuyos gases se “desconectan” del gas del disco y forman una brecha (*gap*, en inglés) de tal manera que se forman anillos de gas a ambos lados de la órbita del planeta. Estos anillos, debido a su viscosidad, generan torques sobre el planeta y el desbalance entre estos torques viscosos dan origen a una migración que hace que el planeta reduzca su radio orbital. Esta es la migración **tipo II**. Cuando el planeta se mueve en esta migración, el gas del anillo interior llena la brecha detrás del planeta generando vórtices que, a su vez, retroalimentan la migración hacia dentro. Eventualmente se presenta un balance entre los torques sobre el planeta y este terminará en una órbita estable (la *trampa*). En algunos casos este balance se presenta cerca de la estrella a distancias menores a 1 UA. Los planetas gigantes que orbitan a sus estrellas en estas distancias se conocen como *Júpiter calientes* y son bastante comunes. Otros gigantes gaseosos llegan a su trampa en distancias superiores a las 3 UA de su estrella y son conocidos como gigantes tibios o fríos: Júpiter y Saturno están entre estos. La gran variedad de tipos de gigantes hace que los estudios mediante síntesis planetarias acerca de la migración tipo II aún se sigan desarrollando para arrojar mejores luces sobre de los posibles procesos que regulan estos cambios orbitales. En los modelos globales actuales, al insertar el torque por viscosidad en la ecuación 2.18, la tasa de migración tipo II está dada por

$$\dot{a}_p = 3\text{sign}(a_p - R_m)\alpha \frac{\Sigma_{g,m} R_m^2}{M_p} \frac{\Omega_m}{\Omega} \left(\frac{H_m}{a_p}\right)^2 \Omega_m^1, \quad (2.23)$$

donde las cantidades con subíndice m se evalúan en el radio de máximo acople viscoso [Lynden-Bell and Pringle, 1974]. La posición de R_m se puede estimar como en [Ida and Lin, 2004b]

$$R_m = 10\text{UA} \exp\left(\frac{2t}{5\tau_{\text{disco}}}\right) \quad (2.24)$$

o como resultado de resolver la ecuación de la evolución de la densidad superficial de gas (ecuación 2.3).

Otros modelos consideran que en la migración tipo II el planeta sigue el movimiento del gas del disco excepto cuando su masa es muy grande comparada con la masa local del disco, en cuyo caso se ralentiza debido a su inercia [Alexander and Armitage, 2009] o que esta migración no depende de la viscosidad del disco sino de los torques. Aún hay mucho por explorar en cuanto a submodelos que expliquen de mejor manera las observaciones de las distancias orbitales de los planetas.

¹La función sign, sgn o signo es una función por tramos definida como $\text{sgn}(x) = -1$ si $x < 0$, $\text{sgn}(x) = 0$ si $x = 0$ y $\text{sgn}(x) = 1$ si $x > 0$

2.1.5. Interacciones N cuerpos

La formación concurrente de varios protoplanetas afecta el crecimiento de los mismos de varias maneras: compiten por la acreción de gas y sólidos, incrementan la velocidad de dispersión de los planetecimales reduciendo potencialmente la tasa de acreción de sólidos de los protoplanetas vecinos; alteran la densidad superficial de planetecimales y reducen el flujo radial de guijarros [Mordasini, 2018]. En el caso en que la amortiguación del gas del disco es insuficiente, las interacciones gravitacionales entre protoplanetas conducen a excitaciones de las excentricidades y esto resulta en alteración de las órbitas, colisiones y eyecciones. La migración también se puede ver afectada porque algunos planetas pueden capturarse en movimientos resonantes promedio que conducen a migraciones colectivas resonantes.

Algunos de los primeros modelos de síntesis planetaria usaron la aproximación de 1 embrión por disco que daban como resultado configuraciones compactas de pocos planetas [Mayor et al., 2011]. En el modelo de Berna se agregó un integrador explícito de N cuerpos al inicio de los submodelos existentes para calcular las colisiones concurrentes entre estos durante la formación de los protoplanetas. Con el objetivo de mantener cortos los tiempos de cómputo en la síntesis planetaria se introdujeron entre 20 y 50 embriones de poca masa ($0,01 M_{\oplus} - 0,1 M_{\oplus}$) en el disco, los cuales interactúan newtonianamente mediante la Segunda Ley escrita en coordenadas heliocéntricas. Las interacciones gravitacionales con el gas del disco (migración, amortiguación de excentricidades) se agregaron como fuerzas adicionales en el integrador [Cresswell and Nelson, 2008].

Una aproximación diferente fue hecha por [Ida and Lin, 2010] y [Ida et al., 2013] quienes desarrollaron una aproximación semi analítica para describir las interacciones gravitacionales de varios protoplanetas de forma estadística basada en escalas de tiempo de cruces orbitales, incluyendo el efecto de captura resonante para planetas migrantes. La ventaja de esta aproximación es que el tiempo de cómputo es de órdenes de magnitud inferior al de integración directa de N cuerpos, mientras que produce distribuciones de excentricidades y semiejes mayores de planetas en interacción que están de acuerdo con las simulaciones de N cuerpos. [Mordasini, 2018].

2.2. Estimación de la masa de los exoplanetas

Existen diferentes métodos de detección de exoplanetas, de los cuales solamente dos permiten hacer una *estimación* de la masa: el método de la **Velocidad Radial** (VR) y el método del **Tránsito Planetario** (TP). Ambas estimaciones serán explicadas a continuación.

2.2.1. Masas medidas por velocidad radial

El método de la VR se basa en el efecto Doppler producido por el movimiento de la estrella observada y que es causado por la presencia de un cuerpo en órbita alrededor de ella. Este efecto Doppler se ha estudiado mediante el uso de espectrógrafos en diferentes observatorios durante los últimos 20 años, tales como el ELODIE [Baranne et al., 1996], CORALIE [Queloz et al., 2001], SOPHIE (Spectrographe pour l’Observation des Phénomènes des Intérieurs stellaires et des Exoplanètes) [Bouchy and Team, 2006], MARVELS (Multi-object Apache Point Observatory Radial Velocity Exoplanet Large-area Survey)[Wang et al., 2012], HARPS (High Accuracy Radial

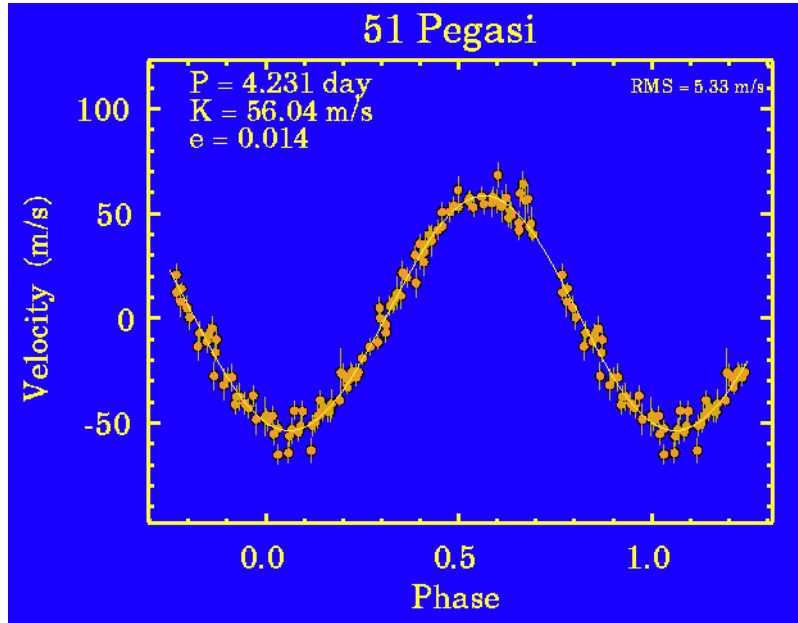


Figura 2.1: *51 Pegasi b* fue el primer exoplaneta detectado y confirmado. Los puntos de la gráfica corresponden a medidas reales de la velocidad radial de la estrella *51 Pegasi* [The Planetary Society, 2022].

velocity Planet Searcher) [Pepe et al., 2000] y el ESPRESSO (Echelle SPectrograph for Rocky Exoplanets and Stable Spectroscopic Observations) [Pasquini et al., 2009]. Con las medidas de velocidad radial se crean gráficas como la que se muestra en la figura 2.1 [The Planetary Society, 2022] De estas gráficas se toman los valores máximo y mínimo de la VR y se calcula la *semiamplitud* de la VR como

$$K_1 = \frac{v_{r,max} - v_{r,min}}{2}, \quad (2.25)$$

que es el valor usado en la estimación de la masa del exoplaneta.

Teniendo en cuenta que la estrella no permanece en reposo en una posición fija sino que tanto ella como el planeta orbitan alrededor del **centro de masa** (CM) común del sistema conjunto; que este movimiento de los dos cuerpos ocurre en un **plano orbital** que, en general, forma un ángulo i con el plano del cielo (figura 2.2), la solución analítica del *problema de los dos cuerpos* para el sistema estrella-planeta arroja una expresión para la semiamplitud de la **componente radial** de la velocidad de la estrella, la cual es [Lovis et al., 2010]

$$K_1 = \sqrt{\frac{G}{(1-e^2)}} m_p \sin i (m_* + m_p)^{1/2} a^{-1/3}. \quad (2.26)$$

Esta semiamplitud, reescrita en unidades de SI y en términos del período orbital P del planeta alrededor de la estrella (que es igual al período orbital de la estrella alrededor del CM común), es

$$K_1 = \frac{28,4329 \text{ ms}^{-1}}{\sqrt{1-e}} \frac{m_p \sin i}{M_{Jup}} \left(\frac{m_* + m_p}{M_\odot} \right)^{-2/3} \left(\frac{P}{1 \text{ yr}} \right)^{-1/3}, \quad (2.27)$$

donde m_* es la masa de la estrella observada, en masas solares; m_p es la masa del exoplaneta y e es la excentricidad de su órbita.

En la ecuación 2.26 la masa m_p aparece ligada al seno del ángulo i de inclinación orbital. Debido

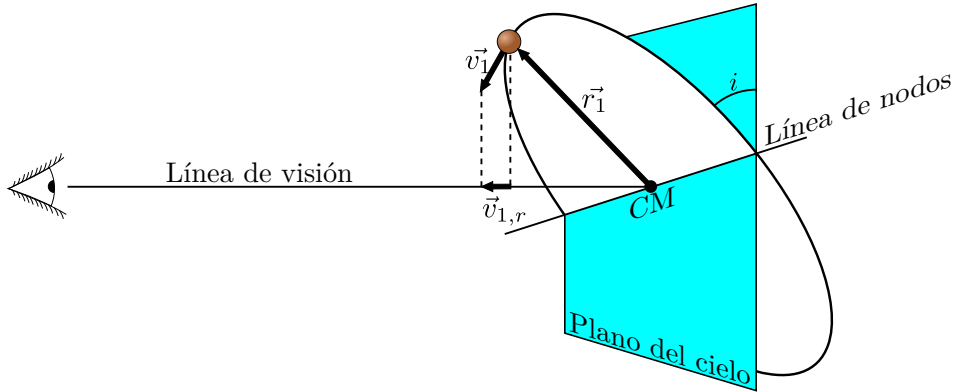


Figura 2.2: Esquema de la geometría básica en la que se basa el método de la velocidad radial.

a que este ángulo es, en principio, desconocido, el término $m_p \sin i$ nos da una masa *mínima* del explaneta y no su masa *real*.

La masa de los exoplanetas medida con el método de la VR tiene dos problemas de base:

- no es la masa *real* de los planetas sino una masa *mínima* y
- por el momento, la precisión de los espectrógrafos limitan las masas mínimas observables, de modo que no se puede abarcar un rango de masas que incluya, por ejemplo, planetas enanos.

De esta manera las masas observadas por VR contienen un sesgo que hace que esta muestra no constituya una representación confiable de la población de planetas en general. Se hace necesario, entonces, buscar estrategias para evadir tal sesgo o validar los datos de masas sintéticas agregando las condiciones de sesgo observacionales de tal forma que los resultados sean equivalentes a las observaciones reales.

2.2.2. Masas medidas por tránsito planetario

A partir de los datos de un explaneta TP es posible obtener una estimación de su masa *total* mediante el método **MassSpec** [de Wit and Seager, 2013a].

El método MassSpec usa los datos del espectro de transmisión de la atmósfera del explaneta cuando la luz de su estrella huésped pasa a través de ésta durante el *tránsito primario* (cuando el explaneta pasa entre la estrella y la Tierra). Este método se fundamenta en que la masa de un planeta afecta el espectro de transmisión a través del perfil de presión de su atmósfera y, como consecuencia, su perfil de absorción. A esta variación de la presión atmosférica con la altura se le conoce como **factor de escala** (*height scale* en inglés [American Meteorological Society, 2022]) H y está definida como

$$H = \frac{\kappa T}{\mu g}, \quad (2.28)$$

donde κ es la constante de Boltzmann y T , μ y g son la temperatura, masa molecular media y aceleración de la gravedad *locales*, es decir, que dependen de la altitud. Expresando g en términos de la masa M_p del planeta y su radio R_p a través de la ley de gravitación universal

$$g = G \frac{M_p}{R_p^2}, \quad (2.29)$$

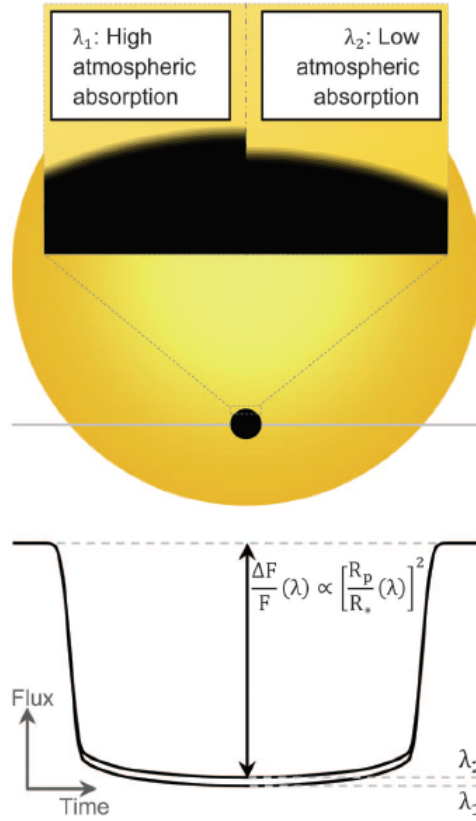


Figura 2.3: Variaciones en el flujo $\frac{\Delta F}{F}(\lambda)$, producidas por la opacidad, dependiendo de la longitud de onda durante un tránsito planetario. Imagen tomada de [de Wit and Seager, 2013a].

la ecuación 2.28 se reescribe como

$$M_p = \frac{\kappa T R_p^2}{\mu G H}. \quad (2.30)$$

La ecuación 2.30 requiere restringir los valores del radio planetario, temperatura atmosférica, masa molecular media y factor de escala.

La temperatura promedio de la atmósfera planetaria se puede obtener a partir de la observación en infrarrojo del tránsito secundario, es decir, cuando el planeta pasa detrás de la estrella. El radio del planeta puede estimarse a partir del hecho de que el espectro de transmisión de un planeta depende de la longitud de onda que atraviesa su atmósfera. De esta manera, una longitud de onda λ_1 con una alta absorción atmosférica registrará una mayor caída en el flujo de luz relativo $\frac{\Delta F}{F}(\lambda)$ de luz de la estrella. Tal caída será menor para una longitud de onda λ_2 con una absorción atmosférica menor. De esta forma, para λ_1 el planeta parecerá tener un radio más grande que para λ_2 (ver figura 2.3). Teniendo en cuenta esto el radio aparente $R_p(\lambda)$ del planeta se puede estimar con la ecuación 2.31 como

$$R_p(\lambda) = \sqrt{\frac{\Delta F}{F}(\lambda)} R_*. \quad (2.31)$$

De acuerdo con [de Wit and Seager, 2013a], el radio aparente del planeta está relacionado directamente con sus propiedades atmosféricas debido a sus efectos en su opacidad y tal relación

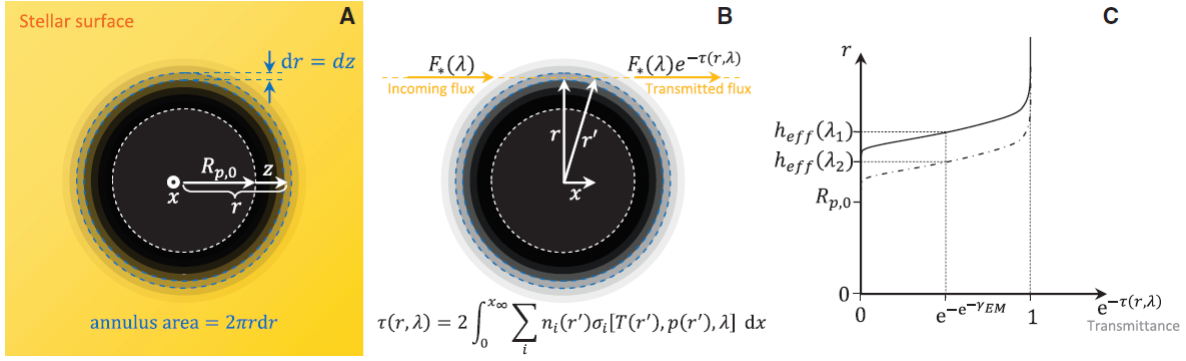


Figura 2.4: Fundamentos del espectro de transmisión (la atmósfera planetaria ha sido aumentada para mejorar la visibilidad) [de Wit and Seager, 2013a].

está dada por

$$\begin{aligned}\pi R_p^2(\lambda) &= \pi [R_{p,0} + h_{eff}(\lambda)]^2 \\ &= \int_0^\infty 2\pi r (1 - e^{-\tau(r,\lambda)}) dr,\end{aligned}\quad (2.32)$$

donde $R_{p,0}$ es un radio planetario de referencia que se toma como la distancia radial en la cual el cuerpo es ópticamente grueso para todo el rango espectral de interés; h_{eff} es la altura atmosférica efectiva y $e^{-\tau(r,\lambda)}$ es la transmitancia del planeta en el radio r (figura 2.4 donde (A) muestra la geometría del tránsito vista por un observador que presenta las áreas de los anillos atmosféricos que afectan el espectro de transmisión; (B) muestra la vista lateral del flujo transmitido a través de un anillo atmosférico de radio r ; (C) muestra la transmitancia en función del radio en longitudes de onda con alta y baja absorción atmosférica λ_1 (líneas continuas) y λ_2 (líneas discontinuas), respectivamente. Debido a la mayor absorción atmosférica en λ_1 , el planeta parecerá más grande que en λ_2 como consecuencia del anillo atmosférico opaco más extenso [$h_{eff}(\lambda_1) > h_{eff}(\lambda_2)$] que se traduce en una caída adicional de flujo). $\tau(r, \lambda)$ es el trayecto oblicuo de la profundidad óptica definido como

$$\tau(r, \lambda) = 2 \int_0^{x_\infty} \sum_i n_i(r') \sigma_i [T(r'), p(r'), \lambda] dx, \quad (2.33)$$

donde $r' = \sqrt{r^2 + x^2}$, $n_i(r')$ es la densidad de partículas en la atmósfera y $\sigma_i [T(r'), p(r'), \lambda]$ es el área eficaz de extinción de la i -ésima componente de atmosférica a la distancia r' [Seager, 2010]. De esta manera, las propiedades de la atmósfera de un planeta [$n_i(z), T(z), p(z)$] están embebidas en su espectro de transmisión mediante $\tau(r, \lambda)$ (ecuaciones 2.32 y 2.33).

La integral en la ecuación 2.32 se puede separar en el radio de referencia debido a que el planeta es opaco en todas las λ para cualquier radio menor a éste y escribirla como

$$\begin{aligned}[R_{p,0} + h_{eff}(\lambda)]^2 &= R_{p,0}^2 + R_{p,0}c, \\ c &\triangleq 2 \int_0^\infty (1+y)(1 - e^{-\tau(y,\lambda)}) dy, \\ y &= z/R_{p,0},\end{aligned}\quad (2.34)$$

Proceos de extinción — Régimen registrado	A_λ	B
Dispersión Rayleigh	$\sqrt{2\pi R_{p,0} \bar{H}} \sum_i n_{i,0} \sigma_i(\lambda)$	$\frac{H}{R_{p,0}}$
Absorción inducida por colisión	$\sqrt{\pi R_{p,0} \bar{H}} \sum_i K_i(T) n_{i,0}^2$	$\frac{H}{2R_{p,0}}$
Absorción molecular		
Régimen Doppler ($p < a_\kappa$)	$\sqrt{2\pi R_{p,0} \bar{H}} \Lambda_\kappa \frac{a_\kappa}{b_\kappa} p_0$	$\frac{H}{R_{p,0}}$
Régimen de transición de Voigt a Doppler ($p^2 < b_\kappa$)		
if $p \sim a_\kappa$	$\sqrt{\frac{4}{3}\pi R_{p,0} \bar{H}} \Lambda_\kappa \frac{2\sqrt{a_\kappa}}{b_\kappa} p_0^{1,5}$	$\frac{H}{1,5R_{p,0}}$
if $p \gg a_\kappa$	$\sqrt{\pi R_{p,0} \bar{H}} \Lambda_\kappa \frac{1}{b_\kappa} p_0^2$	$\frac{H}{2R_{p,0}}$
Régimen de Voigt		
if $p^2 \sim b_\kappa \gg a_\kappa^2$	$\sqrt{2\pi R_{p,0} \bar{H}} \Lambda_\kappa \frac{2}{\sqrt{b_\kappa}} p_0$	$\frac{H}{R_{p,0}}$

 Tabla 2.1: Valores de A_λ y B de acuerdo con el proceso de extinción.

llevando a la siguiente expresión para la altura efectiva de la atmósfera

$$h_{eff}(\lambda) = R_{p,0}(-1 + \sqrt{1 + c}). \quad (2.35)$$

La información atmosférica embebida puede ser accesada de forma directa para los rangos de longitud de onda ópticamente más activos usando

$$h_{eff}(\lambda) = R_{p,0}B(\gamma_{EM} + \ln A_\lambda), \quad (2.36)$$

donde γ_{EM} es la constante de Euler-Mascheroni [Euler, 1740]:

$$\gamma_{EM} = \lim_{n \rightarrow +\infty} \sum_{k=1}^n \frac{1}{k} - \ln n \approx 0,57722$$

En la ecuación 2.35, B es un múltiplo del factor de escala adimensional y A_λ es un trayecto oblicuo extendido de la profundidad óptica en el radio de referencia. La formulación exacta de B y A_λ dependen de los procesos de extinción (dispersión) que afectan la transmisión en la longitud de onda λ [de Wit and Seager, 2013b]. Para dispersión Rayleigh,

$$B = \frac{H}{R_{p,0}} y \quad (2.37)$$

$$A_\lambda = \sqrt{2\pi R_{p,0} \bar{H}} n_{sc,0} \sigma_{sc}(\lambda), \quad (2.38)$$

donde $n_{sc,0}$ es la densidad de partículas en $R_{p,0}$ y $\sigma_{sc}(\lambda)$ el área eficaz de los dispersores. Otros valores para B y A_λ se encuentran en la tabla 2.1 [de Wit and Seager, 2013b], donde p es la presión atmosférica y a_κ y b_κ son parámetros que se introducen para modelar los procesos.

Conceptualmente, la ecuación 2.36 es la altura a la cual la atmósfera se vuelve transparente para un trayecto oblicuo de profundidad óptica A_λ en un radio de referencia. Por ejemplo, si A_λ es 10^4 , entonces la atmósfera se vuelve transparente a ~ 9 factores de escala por encima del radio de referencia.

De acuerdo con [de Wit and Seager, 2013b]:

Lo más importante es que la ecuación 2.36 revela que un espectro de transmisión depende de

sus parámetros clave: en particular, A_λ únicamente depende del factor de escala, la presión de referencia, la temperatura y la densidad de partículas de los principales absorbentes atmosféricos, lo cual lleva a la masa molecular media μ . La unicidad de esta dependencia permite recuperar de manera independiente cada uno de estos parámetros clave. Por tanto, la masa de un planeta se puede restringir de manera única usando espectroscopia de transmisión.

Capítulo 3

Síntesis de poblaciones planetarias

Dentro del estudio de las ciencias planetarias uno de los tópicos de mayor relevancia en la actualidad son los exoplanetas. Desde hace casi tres décadas el descubrimiento de otros sistemas planetarios en la galaxia ha planteado nuevos retos en lo referente a la comprensión de la dinámica y el origen de los planetas.

El análisis de los sistemas exoplanetarios tiene dos componentes que son su fundamento y razón de ser: las observaciones de exoplanetas y los modelos de formación y evolución planetaria. Estas dos componentes deben ser complementarias y, además, retroalimentarse mutuamente. Por un lado, las observaciones imponen restricciones a lo que los modelos de formación y evolución planetaria deben predecir; a su vez, los modelos permiten comprender los procesos físicos que dieron como resultado las observaciones realizadas. Las observaciones sirven para hacer estadísticas con el objetivo de comprender aspectos generales acerca de los planetas y los modelos deben tener la capacidad de replicar los resultados de estas estadísticas. La síntesis de poblaciones planetarias está en el medio de estos dos componentes.

Por síntesis de poblaciones planetarias nos referimos a un conjunto de métodos y cálculos basados en los modelos de formación y evolución planetaria cuyo objetivo es reproducir, mediante simulaciones computacionales, sistemas planetarios consistentes con las observaciones. La síntesis de formación de planetas es un conjunto de cálculos computacionales que permiten evaluar los modelos para mejorarlos.

Los trabajos de síntesis planetaria deben tener dos componentes: (a) modelos globales de formación y evolución planetaria de principio a fin a partir de las propiedades de discos protoplanetarios y las distribuciones de probabilidad de las condiciones iniciales [Mordasini, 2018]; (b) submodelos que dan cuenta de los procesos físicos considerados en los modelos globales tales como la evolución del disco protoplanetario, la acreción de sólidos por parte de los planetecimales, la migración orbital y las colisiones e interacciones de N cuerpos entre los planetas que crecen de forma concurrente.

Actualmente existen varios modelos de síntesis planetaria y todos dan cuenta, más o menos, de las mismas observaciones y restricciones que estas observaciones imponen a los modelos teóricos de formación y evolución planetaria. Si bien reproducen resultados similares con mayor o menor precisión respecto a las observaciones, la principal diferencia entre estos modelos radica en el énfasis detallado de algunas etapas del proceso de formación de sistemas planetarios, el tiempo de ejecución, optimización y variaciones en las condiciones iniciales de las propiedades del disco protoplanetario.

A continuación se presenta una breve revisión de los modelos de síntesis planetaria más conocidos y luego se explicará el funcionamiento del modelo empleado en el código de síntesis desarrollado durante esta investigación.

3.1. Modelo de Berna de formación y evolución planetaria

El modelo de Berna¹ [Emsenhuber et al., 2021b] es un modelo global que calcula de forma auto consistente la evolución del disco de planetecimales y gas, la acreción de sólidos y gas por parte de los protoplanetas, su estructura interna y atmosférica. Este modelo también da cuenta de las interacciones entre protoplanetas y de éstos con el disco de gas.

En la serie de artículos [Emsenhuber et al., 2021b,a, Schlecker et al., 2021a, Burn et al., 2021, Schlecker et al., 2021b, Mishra et al., 2021] se hace una clasificación de los modelos de síntesis de poblaciones de planetas de acuerdo con las consideraciones de cada uno y, eventualmente, su aparición cronológica en la historia de los modelos de formación y evolución planetaria. De acuerdo con estos trabajos, ahora estamos en la Generación III de los modelos Berna de formación y evolución planetaria. La clasificación de los modelos se describe a continuación.

La **Generación I** comenzó con el modelo de [Alibert et al., 2004, 2005] para planetas individuales. Este modelo después fue usado por [Mordasini et al., 2009a,c] para poblaciones planetarias enteras. Inicialmente, este modelo calculaba la formación de un único planeta hasta que se dispersara el gas del disco y luego se dividió en dos ramas: una tenía el objetivo de seguir la evolución a largo plazo del planeta formado (**Generación Ib**, [Mordasini et al., 2012b,a]); la otra, fue capaz de formar sistemas multiplanetarios con una descripción mejorada del disco de planetecimales (**Generación II**, [Alibert et al., 2013, Fortier et al., 2013]).

Resumiendo cada una de las generaciones de los modelos de síntesis de sistemas planetarios, las características de cada una se puede esquematizar en los siguientes 10 puntos [Emsenhuber et al., 2021b]:

1. Número inicial de embriones en el disco, tipo de integrador de N cuerpos, masa de los embriones iniciales.
2. Fases simuladas.
3. Modo de acreción de planetecimales y tamaño de los mismos.
4. Fases con cálculos de estructura interna de los planetas.
5. Características del modelo del disco.
6. Migración orbital tipo I, tipo II y criterios de transición del tipo I al tipo II. En los modelos de migración, el carácter "térmico" se refiere a un criterio donde únicamente cuenta la razón entre el radio de Hill y el factor de escala vertical del disco; "térmico y viscoso" se refiere el criterio completo de [Crida et al., 2006] en el que, aparte de los torques gravitacional y por viscosidad, se debe tener en cuenta la fracción de torque gravitacional que es evacuado

¹La mayoría de los autores del artículo donde se expone este modelo están afiliados a la universidad de Berna, Suiza.

por ondas de soporte de presión. El flujo de momento angular transportado por estas ondas genera un torque de presión.

7. Tasa de acreción de gas limitada por el disco.
8. Adiciones posteriores y mejoras respecto a la generación anterior.
9. Resultados adicionales relativos a los modelos de generaciones anteriores.
10. Publicaciones de síntesis de poblaciones usando los criterios de la generación.

3.1.1. Generación I: basada en el modelo de [Alibert et al., 2005]

1. Un solo embrión por disco (no N cuerpos) de $0.6 M_{\oplus}$.
2. Solamente formación (hasta t_{disco}).
3. Acreción de planetecimales en régimen desbocado, 100 km.
4. Solamente fase ligada ².
5. Estructura vertical del disco, sin radiación ni evolución estelar.
6. Migración. Únicamente transiciones con criterio térmico: isotérmico para el tipo I y equilibrio para el tipo II.
7. Flujo de gas en equilibrio en el disco.
8. Irradiación estelar del disco [Fouchet et al., 2012].
9. Masas, distancias orbitales, composición volumétrica (*bulk*, en inglés).
10. [Mordasini et al., 2009a,c, 2012b, Alibert et al., 2011].

3.1.2. Generación Ib [Mordasini et al., 2012a]: inclusión de evolución a largo plazo

1. Un embrión por disco (no N cuerpos) de $0.6 M_{\oplus}$.
2. Formación (hasta t_{disco}) y evolución (termodinámica) (hasta 10 Gyr).
3. Acreción de planetecimales en régimen desbocado, 100 km.
4. Fase ligada, desligada ³ y evolucionaria ⁴, con estructura del núcleo.
5. Estructura vertical del disco, con radiación estelar, sin evolución estelar.
6. De acuerdo con [Dittkrist et al., 2014], migración tipo I no isotérmica, tipo II en no equilibrio, con criterio de transición viscosa y térmica.

²Se da este nombre a la fase en la que la envoltura de gas del protoplaneta está en equilibrio con el gas del disco, de tal forma que no se distingue un radio exterior del planeta.

³Se da este nombre a la fase en la la tasa de acreción de gas excede el máximo suministro que puede dar el disco, lo que incluye que el planeta ya no puede estar en una región donde el gas esté presente.

⁴Después de que se ha dispersado el gas del disco

7. Flujo de gas en no equilibrio en el disco.
8. Fusión de Deuterio, escape atmosférico.
9. Radios, luminosidades, tasa de evaporación de envolturas.
10. [Mordasini, 2014, Mordasini et al., 2017, Jin and Mordasini, 2018].

3.1.3. Generación II, [Alibert et al., 2013]: inclusión de interacciones de N cuerpos

1. Varios embriones por disco (integrador de N cuerpos EMPS), $0.01 M_{\oplus}$.
2. Solamente formación (hasta t_{disco}).
3. Acreción de planetecimales en régimen oligárquico, 300 m [Fortier et al., 2013].
4. Solamente fase ligada.
5. Estructura vertical del disco, sin radiación ni evolución estelar.
6. De acuerdo con [Dittkrist et al., 2014], migración tipo I no isotérmica, tipo II en no equilibrio, con criterio de transición viscosa y térmica.
7. Flujo de gas en no equilibrio en el disco.
8. Seguimiento (tracking) de la composición [Thiabaud et al., 2015].
9. Multiplicidad, excentricidades, MMR (mean-motion resonances).
10. [Pfyffer et al., 2015, Alibert and Benz, 2017].

3.1.4. Generación III: evolución a largo plazo y evolución de N cuerpos

1. Muchos embriones por disco (integrador de N cuerpos tipo Mercurio), $0.01 M_{\oplus}$.
2. Formación (hasta 20 Myr) y evolución termodinámica (hasta los 10 Gyr).
3. Acreción de planetecimales en régimen oligárquico, 300 m.
4. Fase ligada, desligada y evolucionaria (con fusión de Deuterio, escape atmosférico, estructura de núcleo).
5. Estructura (integrada) vertical del disco, con radiación y evolución estelar.
6. De acuerdo con migración tipo I no isotérmica, tipo II en no equilibrio, con criterio de transición viscosa y térmica.
7. Acreción de gas tipo Bondi [Bondi, 1952].
8. No hay adiciones.
9. Combina los resultados de las generaciones Ib y II.

10. La serie NGPPS (*The New Generation Planetary Population Synthesis*) [Emsenhuber et al., 2021b,a, Schlecker et al., 2021a, Burn et al., 2021, Schlecker et al., 2021b, Mishra et al., 2021].

3.2. PSyCo: Planetary Synthesis Code

Durante el desarrollo de esta investigación se escribió un código de síntesis de poblaciones planetarias que se ha sido llamado PSyCo (Planetary Synthesis Code) y que está basado, principalmente, en los trabajos de [Ida and Lin, 2008] y [Benz et al., 2014]. Por sus características, este código se puede clasificar en una etapa intermedia entre las generaciones Ib y II puesto que establece varios embriones planetarios en el disco y, aunque también se desarrolló un integrador de N cuerpos en esta investigación, éste no está integrado (aún) con este modelo de tal forma que los embriones planetarios no describen colisiones concurrentes con los procesos de acreción de material del disco. Estas colisiones se estudian en una etapa posterior a la de formación de masas planetarias. Las etapas y submodelos del código se describen a continuación.

3.2.1. Estructura y evolución del disco protoplanetario

Este código de síntesis de poblaciones planetarias parte del modelo de Nebulosa Solar de Masa Mínima (minimum mass solar nebula, MMSN [Hayashi, 1981]). El principal objetivo de este modelo es examinar las propiedades estadísticas de los planetas sintetizados, más que el estudio de los procesos individuales que regulan la estructura del disco. Por conveniencia computacional se usó de modelo de [Ida and Lin, 2008], que introduce los factores f_d y f_g con los que se escalan las densidades superficiales de gas Σ_g y polvo Σ_d en términos de los valores de referencia para la MMSN,

$$\begin{aligned}\Sigma_d &= \Sigma_{d,10} \eta_{ice} f_d (r/10 \text{ UA})^{-q_d}, \\ \Sigma_g &= \Sigma_{g,10} f_g (r/10 \text{ UA})^{-q_g},\end{aligned}\tag{3.1}$$

donde $\Sigma_{d,10} = 0,32 \text{ g cm}^{-2}$ y $\Sigma_{g,10} = 75 \text{ g cm}^{-2}$ son factores de normalización correspondientes a 1,4 veces los valores de Σ_d y Σ_g a 10 UA en el modelo MMSN y el factor que representa la razón hielo/polvo η_{ice} , cuyo valor es 1.0 dentro de la línea de hielos a_{ice} y 4.2 para $r > a_{ice}$.

Para un tratamiento autoconsistente de la acreción, la temperatura del disco está determinada por un equilibrio entre la disipación viscosa y transporte de calor [Shakura and Sunyaev, 1973]. En este modelo se ha despreciado el balance detallado de energía y se adoptó una distribución de temperatura de equilibrio en la prescripción de discos ópticamente delgados de [Hayashi, 1981] tal que,

$$T = 280 \left(\frac{r}{1 \text{ UA}} \right)^{-1/2} \left(\frac{L_*}{L_\odot} \right)^{1/4} \text{ K}.\tag{3.2}$$

La línea de hielos se estableció como aquella determinada por una temperatura de equilibrio en regiones del disco ópticamente delgadas,

$$a_{ice} = 2,7 \left(\frac{L_*}{L_\odot} \right)^{1/2} \text{ UA},\tag{3.3}$$

donde L_* y L_\odot son la luminosidad estelar y solar respectivamente. La magnitud de a_{ice} puede ser afectada por disipación viscosa y radiación estelar pero tales efectos no cambian significativamente la estructura del disco en etapas tardías de la evolución y por eso no son incluidos en este modelo.

La densidad superficial del disco se puede determinar a partir de la ecuación de difusión dada por [Lin and Papaloizou, 1979],

$$\frac{\partial \Sigma_g}{\partial t} - \frac{1}{r} \frac{\partial}{\partial r} \left[3r^{1/2} \frac{\partial}{\partial r} (\Sigma_g \nu r^{1/2}) \right] = 0, \quad (3.4)$$

que corresponde a la ecuación 2.3 donde se han omitido los términos de reducción del gas debidos a la foto evaporación y a la acreción por núcleos de gigantes. Por conveniencia computacional (reducción de los tiempos de cómputo), en ocasiones se adopta la prescripción de α constante. [Shakura and Sunyaev, 1973]. En otras ocasiones, como en el modelo implementado en el código de síntesis planetaria desarrollado durante esta investigación, α adopta valores en el rango entre $10^{-4} - 10^{-2}$ puesto que se asume un disco visco-turbulento y α constituye un parámetro que se ajusta para que los tiempos de evolución del disco correspondan con aquellos deducidos de las observaciones de discos protoplanetarios debido a que su viscosidad no puede definirse clásicamente. De esta manera,

$$\nu = \alpha c_s H, \quad (3.5)$$

donde c_s , $H = \sqrt{2}c_s/\Omega_K$ y Ω_K son, respectivamente, la rapidez del sonido, escala de altura del disco y la frecuencia angular kepleriana.

En este modelo no se resuelve directamente la ecuación diferencial parcial 3.4 sino que se proponen soluciones en leyes de potencias dadas por las ecuaciones 3.1. Para estas ecuaciones,

$$f_g = f_{g,0} \exp(-t/\tau_{dep}), \quad (3.6)$$

donde τ_{dep} es el tiempo de vida del disco y se ha tomado como un parámetro entre $10^6 - 10^7$ yr. En particular, nuestros cálculos asumen $\tau_{dep} = 10^7$ yr. De forma análoga a la evolución del gas, nuestro modelo asume que la fracción de material sólido del disco, f_d , evoluciona como

$$f_d = f_{d,0} \exp(-t/\tau_{dep}), \quad (3.7)$$

con $f_{d,0} = f_{g,0}(Z_*/Z_\odot)$, donde Z_* representa la metalicidad de la estrella central y $Z_\odot = 0,0196$ [von Steiger and Zurbuchen, 2015]. En el trabajo original de [Ida and Lin, 2004a] no se tuvo en cuenta la evolución temporal de f_d como consecuencia de la acreción de planetecimales por parte de núcleos en su zona de alimentación y también de núcleos en procesos de migración; sin embargo, estos últimos no hacen decrecer tanto f_d como los primeros.

3.2.2. PSyCo: crecimiento de núcleos y migración tipo I

Para un disco de densidad de gas Σ_g y polvo Σ_d alrededor de una estrella de masa M_* el modelo de la masa de los embriones en una posición a en el tiempo t que se ha tomado es, de acuerdo con [Ida and Lin, 2005],

$$M_c(t) \simeq \left(\frac{t}{0,48 \text{ Myr}} \right)^3 \left(\frac{\Sigma_d}{10 \text{ g cm}^{-2}} \right)^3 \left(\frac{\Sigma_g}{2,4 \times 10^3 \text{ g cm}^{-2}} \right)^{6/5} \\ \times \left(\frac{m}{10^{22} \text{ g}} \right)^{-2/5} \left(\frac{a}{1 \text{ UA}} \right)^{-9/5} \left(\frac{M_*}{M_\odot} \right)^{1/2} M_\oplus, \quad (3.8)$$

donde m es la masa típica de los planetecimales acretados por los embriones. Con esto, la tasa de acreción de masa de los núcleos es

$$\dot{M}_c = \frac{M_c}{\tau_{acc}}, \quad (3.9)$$

donde, de acuerdo con [Ida and Lin, 2008],

$$\tau_{c,acc} = 2,2 \times 10^5 \eta_{ice}^{-1} f_d^{-1} f_g^{-2/5} 10^{(2/5)(3/2-q_g)} \\ \times \left(\frac{a}{1 \text{ UA}} \right)^{27/10+(q_d-3/2)+(2/5)(q_g-3/2)} \left(\frac{M_*}{M_\odot} \right)^{-1/6} \text{ yr}. \quad (3.10)$$

f_d y f_g cambian con el tiempo y $M_c(0) = m = 10^{20}$ g. Cuando $f_g < 10^{-3}$ se usa una tasa de acreción en la que el tiempo de acreción cambia [Ida and Lin, 2004b],

$$\tau_{c,acc} = 2 \times 10^7 \eta_{ice}^{-1} f_d^{-1} \left(\frac{a}{1 \text{ UA}} \right)^{3+(q_d-3/2)} \left(\frac{M_*}{M_\odot} \right)^{-1/2} \text{ yr}. \quad (3.11)$$

La acreción se detiene de forma artificial cuando el núcleo o planetoide alcanza la *masa de aislamiento* (*isolation mass* en inglés) $M_c = M_{c,iso}$ [Ida and Lin, 2004b,a, 2005] dada por

$$M_{c,iso} \simeq 0,16 \eta_{ice}^{3/2} f_d^{3/2} \left(\frac{a}{1 \text{ UA}} \right)^{3/4-(3/2)(q_d-3/2)} \left(\frac{M_*}{M_\odot} \right)^{-1/2} M_\oplus. \quad (3.12)$$

Por otro lado, el desbalance de los torques de marea entre las partes interior y exterior del disco, respecto de la órbita del núcleo, genera lo que conocemos como la migración tipo I. La escala de tiempo de esta migración está dada por [Tanaka et al., 2002],

$$\tau_{mig} \simeq (5 \times 10^4) 10^{(3/2-q_g)} \frac{1}{C_1 f_g} \left(\frac{M_c}{M_\oplus} \right)^{-1} \left(\frac{a}{1 \text{ UA}} \right)^{q_g} \left(\frac{M_*}{M_\odot} \right)^{3/2} \text{ yr}, \quad (3.13)$$

de donde calculamos la tasa de migración tipo I como

$$\frac{da}{dt} = \frac{a}{\tau_{mig}} \simeq 1,6 \times 10^5 \frac{1}{C_1 f_g} \left(\frac{M_c}{M_\oplus} \right)^{-1} \left(\frac{a}{1 \text{ UA}} \right) \left(\frac{M_*}{M_\odot} \right)^{3/2} \text{ yr}, \quad (3.14)$$

para $q_g = 1,5$ y $q_d = 1,5$.

3.2.3. PSyCo: formación de planetas gigantes

En principio, núcleos con masas mucho menores que la masa de la Tierra pueden acretar gas. Sin embargo, si el calor liberado durante la acreción de gas y planetecimales no se distribuye ni se irradia hacia afuera, se establece un equilibrio hidrodinámico y cuasi-térmico que detiene el flujo

de gas hacia el núcleo. Para masas bajas la temperatura y la densidad son bajas de tal forma que el calor no alcanza a distribuirse a través de la envoltura de gas pero, cuando el bombardeo de planetecimales hace crecer la masa más allá de un valor crítico

$$M_{c,\text{hydro}} \simeq 10 \left(\frac{\dot{M}_c}{10^{-6} M_{\oplus} \text{ yr}^{-1}} \right)^{0,25} M_{\oplus}, \quad (3.15)$$

tanto el transporte de energía por radiación y convección se vuelven lo suficientemente eficientes para que la envoltura se contraiga dinámicamente. En regiones donde los núcleos han alcanzado el aislamiento se reduce la tasa de acreción \dot{M}_c y $M_{c,\text{hydro}}$ se vuelve comparable a $\sim 1.0 M_{\oplus}$. La acreción de gas libera energía y depende de la eficiencia de la transferencia radiativa de la envoltura, tal que

$$\frac{dM_p}{dt} \simeq \frac{M_p}{\tau_{KH}}, \quad (3.16)$$

donde $M_p = M_c + M_g$ es la masa planetaria e incluye la masa de la envoltura. La escala temporal de la contracción de Kelvin-Helmholtz τ_{KH} está dada por [Ida and Lin, 2004a]

$$\tau_{KH} \simeq 10^{k_1} \left(\frac{M_p}{M_{\oplus}} \right)^{-k_2} \text{ yr}, \quad (3.17)$$

donde $8 \leq k_1 \leq 10$ y $3 \leq k_2 \leq 4$ para tener en cuenta incertidumbres asociadas con el bombardeo de planetecimales y sedimentación de polvo. La tasa de acreción de gas está dada por

$$\dot{M}_g \simeq \frac{M_g}{\tau_{KH}}. \quad (3.18)$$

La acreción de gas termina y se forma una brecha cuando el radio de Hill se vuelve mayor que la escala de altura del disco [Lin and Papaloizou, 1986], lo que se conoce como *la condición térmica*,

$$M_{g,\text{iso}} \simeq 0,95 \times 10^3 \left(\frac{a}{1 \text{ UA}} \right)^{3/4} \left(\frac{L_*}{L_{\odot}} \right)^{3/8} \left(\frac{M_*}{M_{\odot}} \right)^{-1/2} M_{\oplus}. \quad (3.19)$$

3.2.4. PSyCo: migración tipo II

La migración tipo II puede darse bajo dos condiciones: (i) antes de que la masa del gas decaiga a valores comparables con la masa del planeta, el planeta migra con acreción del disco; (ii) cuando la masa de gas del disco ya es comparable con la masa del planeta, una fracción del flujo de momento angular del disco es utilizada por el planeta en su evolución orbital. Para el caso (i) la tasa de acreción está dada por

$$\tau_{\text{mig2,ii}} \simeq 5 \times 10^5 f_g^{-1} \left(\frac{C_2 \alpha}{10^{-4}} \right)^{-1} \frac{M_p}{M_J} \left(\frac{a}{1 \text{ UA}} \right)^{1/2} \left(\frac{M_*}{M_{\odot}} \right)^{-1/2} \text{ yr}. \quad (3.20)$$

Cuando $\tau_{\text{mig2,ii}}$ es más corta que la escala temporal de la correspondiente a la migración en el caso (i), se usa $\tau_{\text{mig2,i}}$, dada por

$$\tau_{\text{mig2,i}} \simeq 0,7 \times 10^5 \left(\frac{\alpha}{10^{-3}} \right)^{-1} \frac{a}{1 \text{ UA}} \left(\frac{M_*}{M_{\odot}} \right)^{-1/2} \text{ yr}. \quad (3.21)$$

Nuestro modelo de síntesis de poblaciones de planetas es mucho más sencillo que aquellos modelos globales descritos en 2.1 puesto que no incluye cálculos de la estructura interna de los núcleos planetarios, no considera efectos de fotoevaporación del gas ni la interacción gravitacional entre planetecimales (integrador de N cuerpos). Aún así, es capaz de generar planetas de masas correspondientes a planetas rocosos y gigantes (gaseosos y helados) en diversas posiciones dentro del disco y las migraciones de cada tipo de planeta. La razón de ello es que el principal objetivo de este modelo es generar datos para hacer estadísticas sobre la distribución de masas de las poblaciones planetarias, más que comprender de forma detallada los procesos físicos implicados en la formación y evolución de los planetas.

Capítulo 4

Métodos estadísticos para el ajuste de masas planetarias a una FDP

La cantidad de exoplanetas confirmados por diferentes métodos de observación nos permite llegar a la conclusión de que la formación de planetas en el universo es un fenómeno bastante común y se han desarrollado diferentes modelos capaces de describir las condiciones bajo las cuales ocurre este fenómeno. Sin embargo, la observación de exoplanetas no es tarea fácil puesto que las distancias a la cuales se encuentran (hasta ahora solamente se han detectado y confirmado exoplanetas en nuestra galaxia), el brillo y variabilidad de sus estrellas huéspedes dificultan la detección de los mismos. Como consecuencia de esto, podemos afirmar que cualquier investigación estadística y sus resultados acerca de la población de planetas en el universo no podrá ser tomada como definitiva e irá mejorando con el paso de los años a medida que dispongamos de más datos observacionales.

Las definiciones de los conceptos básicos de la Estadística necesarios para comprender los métodos utilizados en esta investigación pueden ser consultados en el apéndice A.

En esta investigación se han usado los datos disponibles de los exoplanetas confirmados hasta el 20 de junio de 2022 en los catálogos [Exoplanet Team, 2022] y [NASA, 2022].

La búsqueda de la FDP a la que mejor se ajustan los datos de las masas de los exoplanetas confirmados requiere determinar la naturaleza continua o discreta del conjunto de datos al que pertenece la muestra.

En el caso de las masas de exoplanetas confirmados por VR y teniendo en cuenta que no existe razón alguna para suponer que la inclinación i del plano orbital depende de la dirección de observación, la única suposición válida acerca de i es que es completamente aleatoria y continua. Adicionalmente, considerando los diferentes modelos de formación planetaria tampoco se encuentran restricciones a las masas de los planetas diferentes al hecho de que, para masas mayores a las $\sim 12 M_J$, se inician procesos nucleares propios de cuerpos estelares, aunque aún está abierta la discusión acerca de si cuerpos masas superiores a $12 M_J$ pueden o no ser considerados como planetas. [Schneider et al., 2011]. Haciendo a un lado esta condición se consideraron como planetas aquellos cuerpos con m_p en el intervalo $(0 M_J, 12 M_J)$.

De acuerdo con lo anterior podemos considerar que el conjunto de masas confirmadas por VR corresponde a una variable aleatoria continua. En el caso de las masas de exoplanetas confirmadas por el método del TP no existen restricciones al rango de masas, por lo que este conjunto también se puede considerar como una variable aleatoria continua. De esta forma, para el análisis estadístico

se usarán funciones de distribución continuas.

Determinar la función de distribución de un conjunto de datos implica usar un método de ajuste de datos y esto se logra mediante una **prueba de bondad de ajuste** 4.0.2. En este tipo de pruebas se comparan los datos con una FDP cuyos parámetros deben estar plenamente establecidos previamente o pueden ser estimados a partir de los datos de la muestra. De los diferentes métodos de estimación de parámetros disponibles en la estadística inferencial hemos usado el de **estimación de máxima verosimilitud** que se explica a continuación.

4.0.1. Estimadores de Máxima Verosimilitud

Para una muestra $\mathbf{x} = (x_1, \dots, x_n)$, la función definida como

$$L(\mathbf{x}|\theta) = L(x_1, \dots, x_n|\theta_1, \dots, \theta_k) = \prod_{i=1}^n f(x_i|\theta_1, \dots, \theta_k) \quad (4.1)$$

se conoce como *función de verosimilitud*. Los estimadores de máxima verosimilitud son los valores $\hat{\theta}_1, \dots, \hat{\theta}_k$ que maximizan la función $L(\mathbf{x}|\theta)$.

Para estimar $\hat{\theta}_1, \dots, \hat{\theta}_k$, primero se calcula el logaritmo de $L(\mathbf{x}|\theta)$ para convertir la productoria en la sumatoria

$$\log L(x_1, \dots, x_n|\theta_1, \dots, \theta_k) = \sum_{i=1}^n f(x_i|\theta_1, \dots, \theta_k) \quad (4.2)$$

lo cual se justifica en el hecho de que los valores que maximizan $L(\mathbf{x}|\theta)$ también maximizan a $\log L(\mathbf{x}|\theta)$ y esta maximización se obtiene, como de costumbre, derivando parcialmente la función con respecto al parámetro θ_i , igualando a cero cada una de las derivadas y resolviendo el sistema de ecuaciones resultante.

$$\frac{\partial \log L(\mathbf{x}|\theta)}{\partial \theta_i} = \frac{\partial}{\partial \theta_i} \sum_{i=1}^n f(x_i|\theta_1, \dots, \theta_i, \dots, \theta_k) = 0 \quad (4.3)$$

Cuando $N \rightarrow \infty$, donde N es el tamaño de la muestra (una muestra se considera grande cuando $n > 30$ [Lehmann, 1999]), los estimadores de máxima verosimilitud poseen propiedades deseables:[Rodó, 2019]

- **Suficiencia:** un estimador $\hat{\theta}$ del parámetro θ es suficiente si la estimación utiliza toda la información relevante de la muestra.
- **Insesgamiento:** un estimador $\hat{\theta}$ se dice que es insesgado cuando su valor esperado $E[\hat{\theta}]$ es igual al parámetro θ .
- **Eficiencia:** se dice que un estimador $\hat{\theta}$ es eficiente si es el que tiene menor varianza que cualquier otro estimador insesgado para todos los posibles valores del parámetro.
- **Consistencia:** se dice que un estimador $\hat{\theta}$ es consistente si, cuando el número de observaciones es grande, el estimador $\hat{\theta}$ converge en probabilidad a su valor verdadero.

Considerando estas propiedades y que el tamaño de la muestra que hay a disposición es grande, en esta investigación se han usado los estimadores de máxima verosimilitud para ajustar las distribuciones propuestas.

4.0.2. Prueba de bondad de ajuste

Luego del cálculo de los estimadores de los parámetros de la FDA a la que se desean ajustar los datos se procede con una **prueba de hipótesis**.

Una *hipótesis* estadística es una afirmación o conjetura que se hace acerca de un parámetro de una distribución o de la misma distribución de probabilidad de la población que se está estudiando.

En toda prueba de hipótesis se enuncian dos hipótesis acerca de la población y que son complementarias o contrarias. Estas son llamadas **hipótesis nula** e **hipótesis alternativa** y se denotan respectivamente como H_0 y H_1 . El objetivo de una prueba de hipótesis es decidir, con base en los datos de la muestra (es decir, la *evidencia*), si se acepta o se rechaza la hipótesis nula.

Vale aclarar que aceptar la hipótesis nula no significa que esta sea cierta sino que *no existe evidencia suficiente para rechazarla* y que, por tanto, se debe seguir trabajando con ella [Viedma, 2018]. El procedimiento de una prueba de hipótesis es una regla que especifica:

- los valores para los cuales la decisión es aceptar H_0 como cierta;
- los valores para los cuales la decisión es rechazar H_0 o, equivalentemente, aceptar H_1 como cierta.

El subconjunto del espacio muestral para el cual se rechaza H_0 se conoce como **región de rechazo** o **región crítica**. El complemento de la región de rechazo es la llamada **región de aceptación**. De los diferentes procedimientos de prueba de hipótesis disponibles en la estadística inferencial se describirá únicamente el conocido como **prueba de bondad de ajuste**.

Una prueba de bondad de ajuste se usa en situaciones en las cuales se desea determinar si un conjunto de datos puede ser considerado como proveniente de una muestra tomada al azar de una población con una distribución dada. Esto significa que se *supone* que la población tiene una distribución y se desea saber si los datos obtenidos corresponden a una muestra aleatoria de esa población. De esta manera, las hipótesis nula y alternativa son:

- H_0 : los datos provienen de una muestra cuya población tiene una FDP $f(\mathbf{x}|\theta)$.
- H_1 : los datos *no* provienen de una muestra cuya población tiene una FDP $f(\mathbf{x}|\theta)$.

La aceptación o rechazo de la hipótesis nula puede ser una decisión correcta o un error dependiendo de si H_0 es verdadera o falsa. De esta forma se definen los errores tipo I y II de la siguiente manera [Viedma, 2018]:

- **Error tipo I:** rechazar H_0 cuando ésta es verdadera.
- **Error tipo II:** aceptar H_0 cuando ésta es falsa.

La probabilidad de cometer un error tipo I se denomina **nivel de significación** y se denota por α . Después de plantear las hipótesis nula H_0 y alternativa H_1 se procede a establecer una **regla de decisión** que corresponde al nivel de significación α y determinar un **estadístico de prueba** que se usa como **criterio de decisión**. La principal ventaja de la prueba de bondad de ajuste radica en que se usan los histogramas de densidad o de frecuencia de los datos y se comparan con los valores de la FDP escogida. Si entre estos dos valores existen diferencias estadísticamente significativas se rechaza H_0 . Si estas diferencias no lo son, se acepta H_0 . El nivel de significación marca el umbral que diferencia entre diferencias estadísticamente significativas y no significativas.

4.0.3. Prueba Chi-Cuadrado

Esta es una prueba de bondad de ajuste basada en la diferencia entre la FDP estimada a partir de la muestra y el histograma de densidades. Es una prueba que depende del tamaño N de la muestra y el estadístico de prueba está dado por

$$\chi_0^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i}, \quad (4.4)$$

donde χ_0^2 es el valor correspondiente a una distribución ji-cuadrada con $k - 1$ grados de libertad, O_i son las frecuencias **observadas** en los datos y E_i son las frecuencias **esperadas** si los datos provinieran de una población cuya distribución es la escogida. Desde el punto de vista del análisis gráfico, O_i representa las áreas de las barras del histograma de densidades y E_i representa las áreas bajo la curva de la FDP propuesta entre los valores extremos de cada clase. De acuerdo con lo anterior, el valor de χ_0^2 nos dice qué tan cercanos son los valores de las áreas en los gráficos del histograma y la FDP.

El valor de χ_0^2 en la ecuación 4.4 se compara con el correspondiente valor de referencia $\chi_{1-\alpha, k-1}^2$, para $k - 1$ grados de libertad, según el nivel de significación α establecido previamente, que se obtiene de la distribución chi-cuadrada. Estos valores de referencia se encuentran tabulados en diferentes libros de estadística. Si $\chi_0^2 > \chi_{1-\alpha, k-1}^2$ se rechaza H_0 . Si bien el nivel de significación α es la regla de decisión hace falta un criterio de decisión. Lo usual es tomar el **valor p** como tal criterio.

El valor p (*p-value*, e inglés) se define como la probabilidad condicional de encontrar (en la población) un valor del estadístico de prueba tan extremo como el calculado a partir de la muestra, dado que H_0 sea cierta [Bernstein and Bernstein, 1999]. Se puede pensar el valor p como un estadístico descriptivo que muestra qué tanto pueden los datos soportar la aceptación de la hipótesis nula: entre más pequeño sea p menor es tal soporte. Dado que el valor p se define como una probabilidad, su valor está en el intervalo $0 < p < 1$. Así las cosas, ¿qué nivel de soporte es considerado demasiado pequeño como para que la hipótesis nula sea rechazada?

El umbral que determina el nivel de soporte se basa en la probabilidad de cometer un error tipo I, es decir, el nivel de significación α . De esta forma, si $p < \alpha$ se rechaza H_0 . Por otro lado, si $p > \alpha$ se acepta H_0 , lo que se interpreta como que no hay evidencias suficientes para rechazar la hipótesis alternativa.

Es usual interpretar el valor p como un indicador de qué tan probable es que las diferencias encontradas en la prueba de bondad de ajuste, es decir, el valor del estadístico de prueba, sean estadísticamente significativas. En otras palabras, el valor p indica que tan probable es que las diferencias se deban a efectos meramente aleatorios en la selección de la muestra. De acuerdo con lo anterior un valor de $p < \alpha$ significa que la probabilidad de que H_0 sea cierta es muy baja.

4.0.4. Prueba de Kolmogorov-Smirnov

La prueba de Kolmogorov-Smirnov (prueba K-S) es una prueba de bondad de ajuste no paramétrica que se basa exclusivamente en la comparación de la **función de densidad acumulada empírica** (FDAE) de los datos de la muestra con la función de densidad acumulada de la distri-

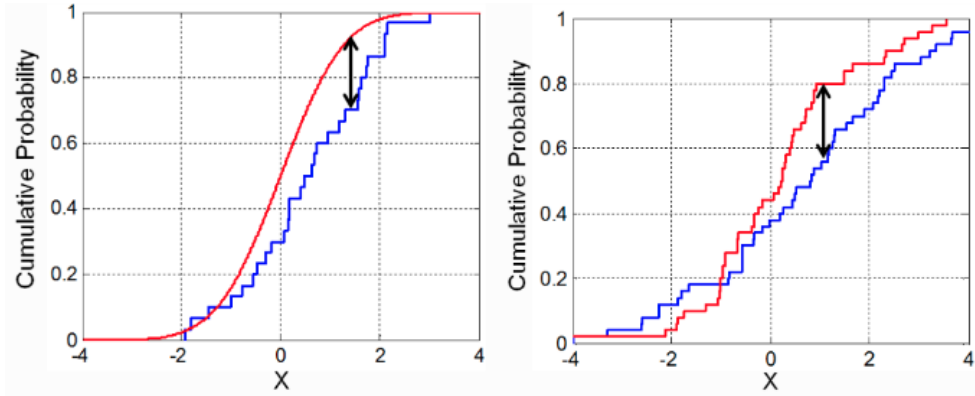


Figura 4.1: Izquierda: ilustración del estadístico de Kolmogorov–Smirnov. La línea roja muestra la función de distribución acumulada teórica, la azul la función de distribución acumulada empírica, y la flecha negra es el estadístico K-S. Derecha: ilustración del estadístico de Kolmogorov–Smirnov entre dos muestras. Las líneas roja y azul muestran la función de distribución acumulada empírica de dos muestras, y la flecha negra es el estadístico K-S.

bución teórica a la que se desean ajustar los datos [Chakravarti et al., 1967]. Esta prueba compara dos conjuntos de datos: los de la muestra y la población (distribución teórica) o dos muestras aleatorias. En nuestro caso usaremos la prueba K-S de una muestra. Como en la prueba χ^2 , en esta prueba se debe definir previamente el nivel de significación α y calcular un estadístico de prueba que servirá como criterio de decisión sobre la aceptación o rechazo de la hipótesis nula H_0 . Para realizar una prueba K-S se deben desarrollar los siguientes pasos:

- Ordenar los N datos $x_1, x_2, x_3, \dots, x_N$ de la muestra de forma ascendente.
- Calcular la FDAE, la cual está dada por

$$F_N(t) = \frac{1}{N} \sum_{i=1}^N 1_{\{x_i \leq t\}} \quad (4.5)$$

- Calcular el estadístico de prueba K-S que está dado por

$$D = \max_{1 \leq i \leq N} \left(F(x_i) - \frac{i-1}{N}, \frac{i}{N} - F(x_i) \right) \quad (4.6)$$

- Comparar el valor p de la prueba con el nivel de significación α y rechazar H_0 si $p < \alpha$.

El estadístico de esta prueba obtenido mediante la ecuación 4.6 calcula las diferencias entre la FDAE de la muestra y la FDA de la distribución de la que se espera proviene la población y se queda con la mayor de estas diferencias. De esta forma, se cuenta con la ventaja de que emplea los datos de la muestra de forma directa y sin recurrir a los parámetros de la distribución teórica y que no depende del tamaño N de la muestra. A pesar de esta ventaja, esta prueba cuenta con algunas desventajas [Stephens, 1974] como:

- es muy sensible a la mediana de la muestra,
- la FDA con la que se comparan los datos debe estar plenamente identificada (es decir, tener el conocimiento de todos sus parámetros) antes de iniciar la prueba y, por tanto, falla cuando

los parámetros se estiman a partir de la muestra [Keutelian, 1991, Durbin, 1976, Massey Jr, 1951].

Otra prueba de bondad de ajuste no paramétrica y que es una modificación de la prueba K-S es la prueba de Anderson-Darling, la cual ofrece una mejora respecto a aquella de la cual proviene en el sentido de que, a diferencia de la prueba K-S, esta también es sensible en los extremos de los datos de la muestra. Debido a que esta prueba es una modificación de la prueba K-S, los pasos que se deben seguir son los mismos que esta y solamente se diferencian en el estadístico de prueba. Para la prueba Anderson-Darling el estadístico A^2 está dado por,

$$A^2 = -N - S, \quad (4.7)$$

donde,

$$S = \sum_{i=1}^N \frac{2i-1}{N} [\ln(F(x_i)) + \ln(1 - F(x_{N+1-i}))].$$

Existen otras pruebas de bondad de ajuste con sus respectivos estadísticos de prueba, paramétricas y no paramétricas, que se emplean exclusivamente para muestras aleatorias discretas o que son específicas de ciertas distribuciones. Las aquí expuestas se pueden emplear con muestras discretas, continuas y sin importar la distribución a la que se desean ajustar los datos.

Capítulo 5

Resultados

En el capítulo 4 se explicaron los métodos estadísticos para analizar los datos de una muestra para lograr, a través de ese análisis, inferir características generales de la población de la cual provienen esos datos. Esos métodos fueron empleados para el análisis de los datos de masas de exoplanetas medidas por velocidad radial y tránsito planetario.

Los datos (muestra) empleados fueron tomados del catálogo de exoplanetas de la NASA¹ [NASA, 2022]. En la actualidad existen tres grandes catálogos de exoplanetas [Exoplanet Team, 2022], los cuales difieren en la cantidad de exoplanetas de la cual ofrecen información y datos debido a que tienen diferentes criterios para filtrar aquellos planetas que reportan. Estos criterios [Christiansen, 2018] se resumen en la tabla 5.1.

De acuerdo con estos criterios, el adoptado por el catálogo de exoplanetas de la NASA es más restrictivo puesto que solamente agregan al catálogo datos de exoplanetas cuya confirmación ha sido publicada en una revista indexada. Por esta razón el análisis estadístico se hizo con estos datos.

Otro asunto importante es el criterio de masa para definir qué es y qué no es un exoplaneta. El consenso general, de acuerdo con los modelos de formación estelar, es que la masa mínima para generar procesos de fusión nuclear es de unas $13 M_J$. Existe aún una discusión acerca de si esta masa necesariamente corresponde a un cuerpo estelar porque algunos autores [Schneider et al., 2011] argumentan que ello depende del proceso de crecimiento de estas masas: si esta masa es recolectada por acreción no se presentarían procesos de fusión nuclear y, por tanto, son posibles planetas con masas superiores a las $13 M_J$. Lo que argumentan es que a partir de este valor se generan procesos nucleares solamente si esta masa se ha formado en un proceso de colapso gravitacional en lugar de acreción de gas y sólidos [Faherty et al., 2021].

En las observaciones de cuerpos orbitando alrededor de estrellas es difícil determinar el proceso mediante el cual se han formado aquellos cuya masa supera las $13 M_J$ puesto que en los diferentes catálogos hay cuerpos de hasta $50 M_J$ y aun así han sido clasificados como planetas.

De acuerdo con lo anterior, en esta investigación se debía adoptar un criterio de decisión acerca de los cuerpos que se considerarían planetas y que, por tanto, serían objeto de análisis. Se adoptó el criterio más conservador apelando al consenso antes mencionado: solamente se consideraron aquellos cuerpos con masas inferiores a las $13 M_J$. Una razón adicional para este criterio, desde el punto de vista práctico, es que al considerar masas superiores a este valor se generan vacíos o *desiertos* planetarios, es decir, rangos de masas en las que no hay registros de exoplanetas y estos

¹<https://exoplanets.nasa.gov/discovery/exoplanet-catalog/>

Catálogo	Criterio de masa	Criterio de confianza	Planetas ²
Exoplanet Encyclopaedia	$M_p - 1\sigma < 60M_{\text{Jup}}$	Artículo sometido, conferencia	5106
NASA Exoplanet Archive	$M_p < 30M_{\text{Jup}}$	Artículo aceptado, referenciado	5044
Open Exoplanet Catalog	No especificado	Open-source	4743

Tabla 5.1: Contenidos de los más grandes catálogos de exoplanetas.

desiertos afectan los resultados de los métodos de inferencia estadística.

Teniendo en cuenta los criterios de selección del catálogo fuente y las masas sujetas a análisis, las características de los datos con los que se trabajó son los siguientes:

- a la fecha, el catálogo de NASA³ reporta 5044 planetas confirmados;
- de estos, 2781 planetas tienen reportada su masa total;
- 2012 planetas tienen reportada la masa mínima $m_p \sin(i)$;
- entre estos dos conjuntos 92 planetas tienen las dos masas reportadas y en todos ellos se cumple que $m_p \sin(i) \leq m_p$.

El histograma de frecuencias de cada uno de estos dos grupos de masas se muestra en la figura 5.1. Los histogramas de la figura 5.1 sugieren escoger una FDP cuya forma matemática tenga una curva decreciente. Teniendo esto en cuenta, con cada uno de los dos grupos de masas se hizo una prueba de bondad de ajuste χ^2 (ver sección 4.0.2), para las funciones de distribución:

- Beta

$$f(x, \alpha, \beta) = \frac{\Gamma(\alpha + \beta)x^{\alpha-1}(x-1)^{\beta-1}}{\Gamma(\alpha)\Gamma\beta} \quad \text{con} \quad 0 \leq x \leq 1 \text{ y } \alpha > 0, \beta > 0 \quad (5.1)$$

- Pareto

$$f(x, b) = \frac{b}{x^{b+1}} \quad \text{con} \quad x \geq 1 \text{ y } b > 0 \quad (5.2)$$

- Ley de potencias

$$f(x, a) = ax^{a-1} \quad \text{con} \quad 0 \leq x \leq 1 \text{ y } a > 0 \quad (5.3)$$

De cada una de estas funciones se hizo la estimación de parámetros usando el método de máxima verosimilitud (ver sección 4.0.1) y los procedimientos descritos en el capítulo 4. Para el caso de las pruebas χ^2 los parámetros estimados de cada FDP, tanto con el conjunto de masas m como $m_p \sin(i)$ están listados en la tabla 5.2 y, teniendo en cuenta los tamaños de cada muestra, estos estimadores cumplen las condiciones de ser **suficientes**, **insesgados**, **eficientes** y **consistentes**. Una vez estimados los parámetros de estas tres FDP se compararon sus gráficas con el histograma de frecuencias de las masas observadas (figura 5.2).

³<https://exoplanetarchive.ipac.caltech.edu/>

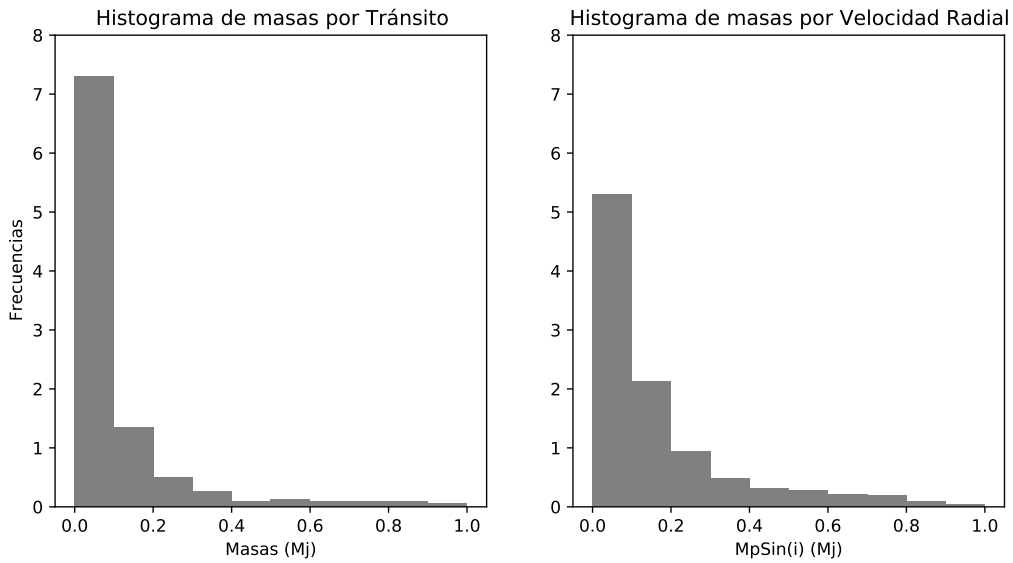


Figura 5.1: Histogramas de los dos grupos de masas observadas. *Izquierda*: masas medidas por tránsito planetario. *Derecha*: masas medidas por velocidad radial.

Si bien la comparación de los histogramas con las gráficas de las FDP dan indicios de a cuál o cuáles funciones de distribución se pueden ajustar bien los datos es necesario un análisis numérico que permita sacar conclusiones con base en evidencia objetiva, que no dependa de las apariencias. Para ello se procedió a realizar las pruebas de bondad de ajuste de para cada conjunto de datos.

En una prueba de bondad de ajuste se proponen dos hipótesis contrarias acerca de las características de la muestra análisis respecto a la población de la que proviene. Las hipótesis nula H_0 y alternativa H_1 en esta investigación fueron

- H_0 : las masas planetarias provienen de una muestra cuya población está descrita por la FDP $f(\mathbf{x}|\theta)$.
- H_1 : las masas planetarias *no* provienen de una muestra cuya población está descrita por la FDP $f(\mathbf{x}|\theta)$.

en donde las FDP son las descritas anteriormente y que están dadas por las ecuaciones 5.1, 5.2 y 5.3 con los parámetros listados en la tabla 5.2.

El objetivo de una prueba de bondad de ajuste es decidir si se tiene *suficiente* evidencia a favor de rechazar la hipótesis nula. Esto significa que no se pretende saber a cuál distribución se ajustan *mejor* los datos sino si la evidencia es suficiente para poder afirmar que la muestra no se ajusta a determinada función. De esta forma, el resultado de una prueba de bondad de ajuste

Distribución	Conjunto m_p	Conjunto $m_p \sin(i)$
Beta	$\alpha = 0,4579, \beta = 15,7518$	$\alpha = 0,5111, \beta = 2,9056$
Pareto	$b = 1,9655$	$b = 1,8288$
Ley de potencias	$a = 0,2535$	$a = 0,3252$

Tabla 5.2: Estimadores de los parámetros de las distribuciones utilizadas para el ajuste de los datos.

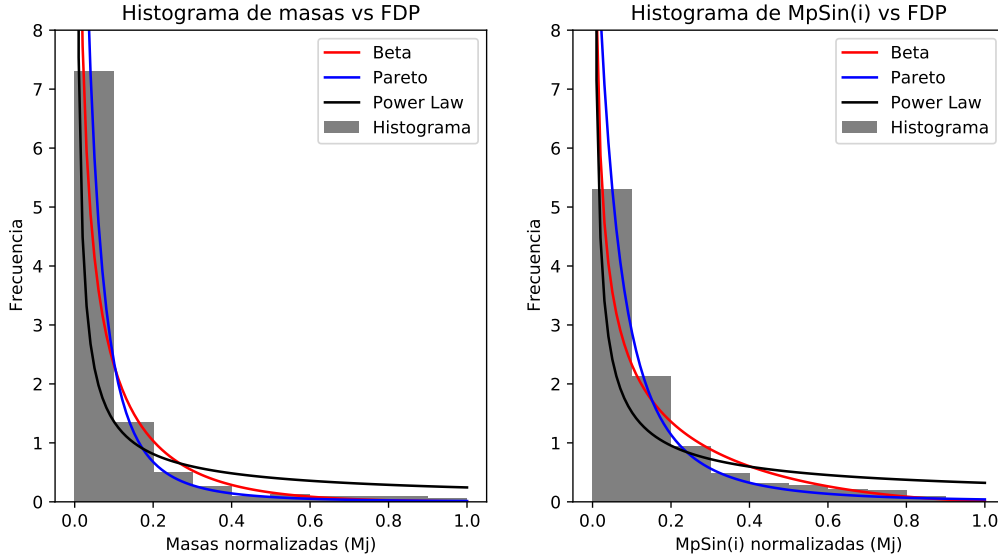


Figura 5.2: Histogramas de masas observadas comparadas con las gráficas de las FDP usadas en esta investigación. *Izquierda*: masas medidas por tránsito planetario. *Derecha*: masas medidas por velocidad radial.

puede ser que *se rechaza* o *no se rechaza* la hipótesis nula... pero esto no nos dice a cuál FDP se ajusta mejor la muestra.

El criterio de decisión acerca de rechazar o no la hipótesis nula viene dado por el tipo de prueba realizado. Las pruebas más comunes para ajustes de datos a distribuciones de probabilidad son la χ^2 (sección 4.0.3) y la prueba Kolmogorov-Smirnov (sección 4.0.4).

En la prueba χ^2 se comparan las áreas de las barras de los histogramas con las áreas bajo la FDP entre los valores correspondientes al ancho de estas barras. El cálculo de áreas bajo curvas corresponde al cálculo de las integrales de las FDP y estas integrales son las Funciones de Densidad Acumulada FDA descritas en la sección A.1. Estas áreas se muestran en gráficos de barras contra los histogramas en la figura 5.3

Esta forma de mostrar las áreas permite visualizar cuáles áreas, en total, se aproximan mejor a las áreas del histograma. Aun así, hace falta todavía un cálculo que nos permita decidir sin depender de lo que cada observador *vea* en las gráficas. Para esto se calculó el estadístico de prueba χ^2 mediante la ecuación 4.4 en donde O_i (los valores observados) corresponden a las áreas de las barras del histograma y E_i (los valores esperados) son las áreas bajo las FDP dentro del ancho de cada barra del histograma. Los estadísticos χ^2 calculados para todas las masas detectadas se muestran en la tabla 5.3.

Estos estadísticos deben ser ahora comparados con el valor crítico del estadístico de prueba χ_c^2

Datos	Beta	Pareto	Ley de potencias
Masas TP	0.862308	0.735821	1.98892
Masas VR	0.345541	0.498121	1.72849

Tabla 5.3: Valores del estadístico de prueba χ^2 calculado para los dos conjuntos de datos y las tres distribuciones propuestas para el ajuste.

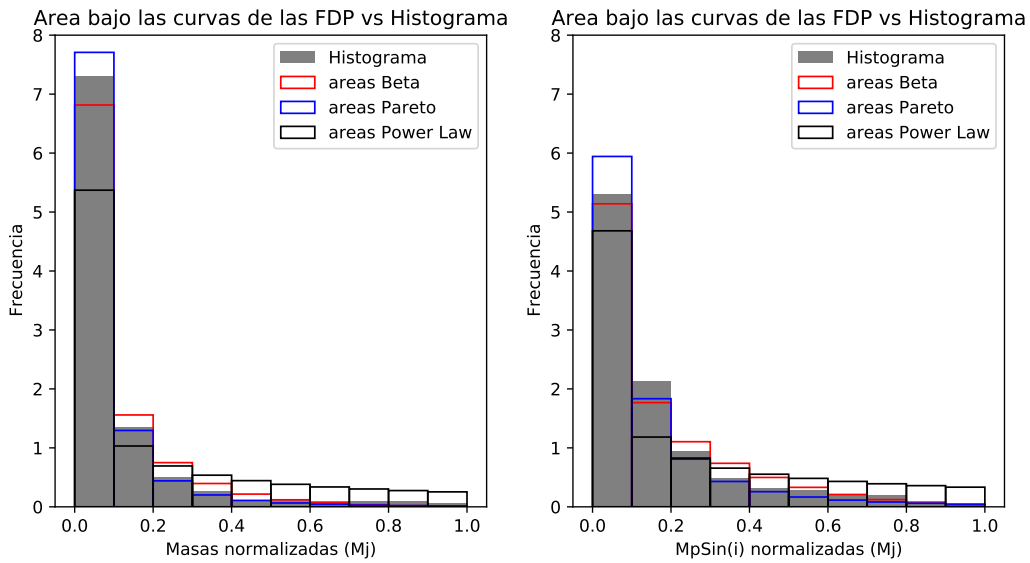


Figura 5.3: Áreas bajo las curvas de las diferentes FDP comparadas con los histogramas de las masas observadas. *Izquierda*: masas medidas por tránsito planetario. *Derecha*: masas medidas por velocidad radial.

y este valor crítico depende de los grados de libertad de la prueba. En una prueba χ^2 el número de grados de libertad gl (en inglés *degrees of freedom* o *df*) es igual al número N de categorías estudiadas menos 1, es decir, $gl = N - 1$. Para esta prueba $gl = 9$ puesto que los valores χ^2 fueron calculados a partir de 10 categorías o rangos de masa dados por el histograma de frecuencias. El criterio de decisión en una prueba χ^2 , basado en el estadístico de prueba es el siguiente:

Si $\chi^2 > \chi_c^2$, se rechaza H_0 . En caso contrario, se acepta.

El valor crítico χ_c^2 también depende del nivel de significancia estadística α escogido para la prueba. Nuevamente, lo tradicional ha sido tomar $\alpha = 0,05$ y ha sido el tomado en esta investigación. De esta forma, para $\alpha = 0,05$ y $df = 9$ tenemos que $\chi_c^2 = 16,919$ y, como ninguno de los estadísticos mostrados en la tabla 5.3 es mayor que χ_c^2 , se acepta la hipótesis nula, es decir, los datos se ajustan a las tres distribuciones propuestas: Beta, Pareto y Ley de Potencias.

Un criterio adicional que suele usarse con mayor confianza que el estadístico de prueba es el *valor-p* de la prueba y su interpretación es que:

si el *valor-p* $< \alpha$ se rechaza H_0 . En caso contrario, se acepta.

En esta investigación los valores p calculados para las tres distribuciones en los dos conjuntos de masa se muestran en la tabla 5.4. Teniendo en cuenta los valores de la tabla 5.4 y el criterio de decisión, se concluye que no hay evidencia suficiente para rechazar la hipótesis nula para ninguno de los dos conjuntos de masa observadas. De esta forma se confirma la conclusión basada en el estadístico de prueba χ^2 . La segunda prueba de hipótesis descrita en en el capítulo 4, es decir, la prueba de Kolmogorov-Smirnov, no se aplicó a los datos de las masas observadas debido a que esta prueba requiere que la FDP con la que se van a comparar los datos de la muestra esté completamente determinada previamente. Como no se conocen los parámetros reales de las FDP ensayadas y estos deben ser estimados a partir de la muestra, esta prueba no es aplicable al conjunto de masas observacionales.

Datos	Beta	Pareto	Ley de potencias
Masas TP	0.999695	0.999843	0.991642
Masas VR	0.999994	0.99997	0.99507

Tabla 5.4: Valores del *valor-p* calculado para los dos conjuntos de datos y las tres distribuciones propuestas para el ajuste.

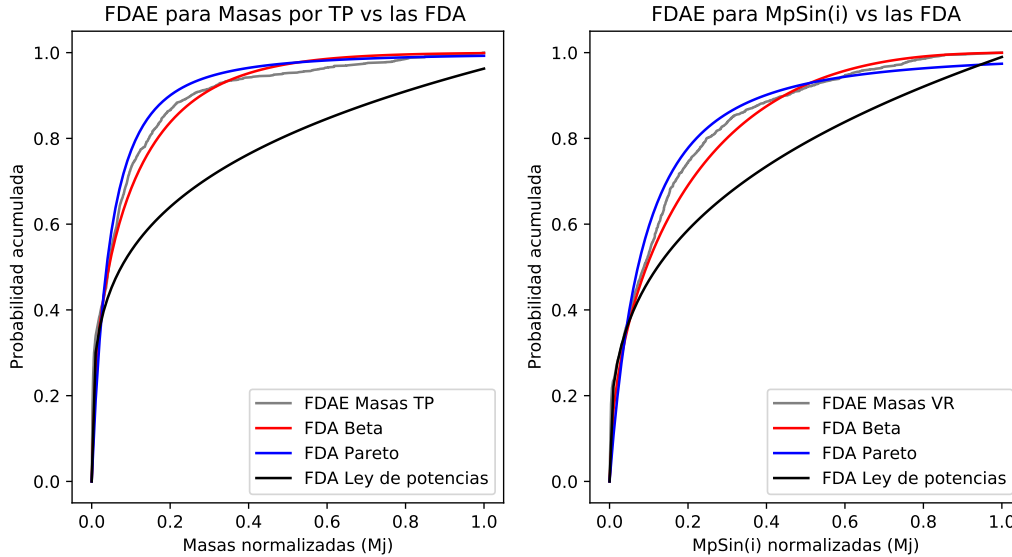


Figura 5.4: Comparación de la FDAE de los conjuntos de masas con las FDA de las distribuciones propuestas para el ajuste. *Izquierda*: masas medidas por TP. *Derecha*: masas medidas por VR

5.1. Resultados de PSyCo

Los datos de las masas de planetas observados tienen desde el principio un sesgo ligado a las limitaciones instrumentales y condiciones de observación de los métodos de detección. Estas limitaciones tienen que ver con la calibración y precisión de los instrumentos, distancias a las que se encuentran las estrellas, tiempos de observación, entre otras. Debido a esto se hace necesario recurrir a los programas de síntesis planetaria para comprender los procesos que llevan a la formación o no de planetas alrededor de una estrella y, a partir de allí, hacer estadísticas que revelen información importante acerca de la población de planetas en el universo. Teniendo en cuenta lo anterior se analizaron los resultados obtenidos con **PSyCo**, el código de síntesis planetaria desarrollado durante esta investigación. La cantidad de planetas sintetizados con los que se hizo el siguiente análisis fue $N = 6410$ con masas entre $4,015084 \times 10^{-10} M_{\oplus}$ y $3829,414133 M_{\oplus}$ o, equivalentemente, entre $1,262977 \times 10^{-12} M_J$ y $12,045729 M_J$. El histograma de las masas sintetizadas se muestra en la figura 5.5. Los parámetros estimados de las tres FDP con las que se

Datos	Beta	Pareto	Ley de potencias
Masas TP	0.0881727	0.18022	0.229841
Masas VR	0.0868903	0.1495	0.171926

Tabla 5.5: Valores del estadístico D de la prueba Kolmogorov-Smirnov para las tres distribuciones propuestas.

Datos	Beta	Pareto	Ley de potencias
Masas TP	$9,3805 \times 10^{-19}$	$2,26152 \times 10^{-77}$	$3,79158 \times 10^{-126}$
Masas VR	$3,07447 \times 10^{-13}$	$1,88172 \times 10^{-38}$	$8,60966 \times 10^{-51}$

Tabla 5.6: Valores p de la prueba Kolmogorov-Smirnov para las tres distribuciones propuestas.

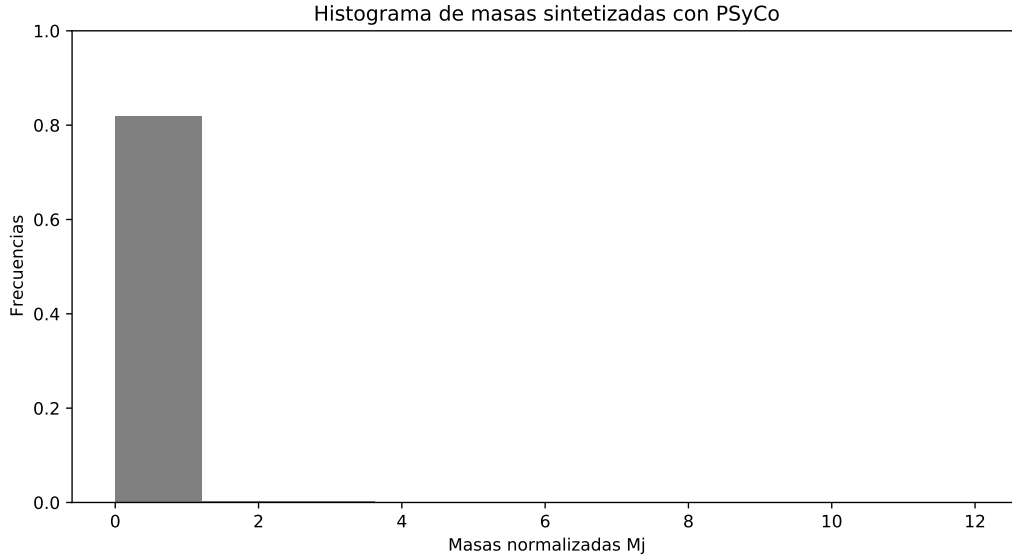


Figura 5.5: Histograma de las masas sintetizadas con PSyCo

hizo el ajuste de estas masas se encuentran en la tabla 5.7. Con los parámetros de la tabla 5.7 se graficaron las FDP contra el histograma de masas sintetizadas (figura 5.6). Con los parámetros estimados se calcularon las áreas bajo las curvas de las FDP entre los rangos de masas del histograma. Mostrando estas áreas en forma de gráfico de barras para comprender mejor las diferencias se obtuvieron los resultados de la figura 5.7. Los resultados de las pruebas χ^2 en las que O_i son las áreas de las barras del histograma y E_i las áreas bajo las FDP se muestran en la tabla 5.7

	Beta	Pareto	Ley de potencias
Estimadores	$\alpha = 0.125761$ $\beta = 37.557161$	3.92644	0.0745821
χ^2	0.0393522	9.3711×10^7	0.152874
Valores p	0.9999999996047415	0.0	0.9999998306124057

Tabla 5.7: Estimadores, valores del estadístico de prueba χ^2 y valores p de las pruebas χ^2 , para las distribuciones utilizadas en el ajuste de los datos de PSyCo.

Una vez más, adoptando un nivel de significancia $\alpha = 0,05$ y con $gl = 9$, el valor crítico de la prueba es $\chi^2 = 16,919$. De acuerdo con este criterio se rechazaría H_0 en el caso de la FDP Pareto y no se rechaza en las otras dos distribuciones.

Para tener otro criterio de decisión se calcularon los *valores-p* de esta prueba para las tres distribuciones y los resultados se muestran en la tabla 5.7. De acuerdo con los *valores-p* calculados y el nivel de significancia adoptado $\alpha = 0,05$ se confirma el rechazo de H_0 en el caso de la distribución de Pareto y el no rechazo de la hipótesis nula para las otras dos distribuciones.

Como en el caso de las masas observacionales, la prueba de Kolmogorov-Smirnov no puede

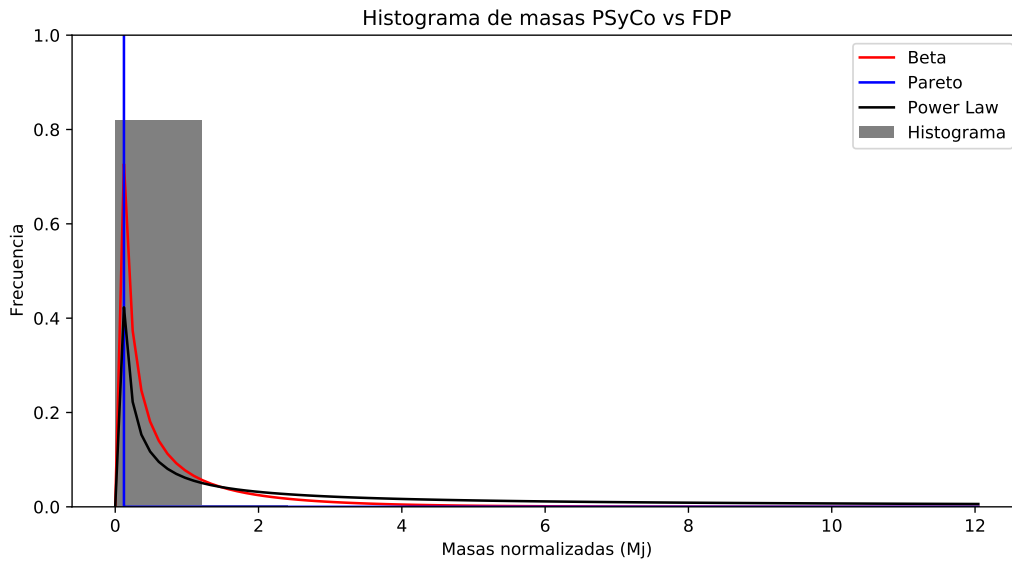


Figura 5.6: Histograma de las masas sintetizadas con PSyCo contra las FDP propuestas

ser aplicada para el conjunto de masas sintetizadas porque no se conocen los parámetros de las funciones de distribución que se proponen y tendrían que ser estimados a partir de estas masas. Al no cumplirse la condición del conocimiento completo previo de estas FDP, solamente podemos sacar conclusiones a partir de la prueba χ^2 . Finalmente, se hace una comparación de los resultados

Dato	Beta	Pareto	Ley de potencias
D	0,403052	0,339126	0,379853
$Valor-p$	0,0	0,0	0,0

Tabla 5.8: Estadístico D y $valor-p$ de la prueba de Kolmogorov-Smirnov para evaluar el ajuste de las masas sintetizadas a las distribuciones propuestas

de PSyCo con aquellos correspondientes a las masas medidas por TP, VR y del modelo en el cuál se basó el modelo de síntesis planetaria de esta investigación. Esta comparación arroja luces sobre la manera de interpretar los resultados aquí mostrados. Primero, el gráfico de dispersión de las masas observadas en función de los semiejes mayores se muestra en la figura 5.9.

En los gráficos de la figura 5.9 se observa que se han detectado muchos planetas muy masivos orbitando a una distancia menor a 1 UA de su estrella. Los resultados de las masas sintetizadas y sus radios orbitales para PSyCo se muestran en la figura 5.10

Teniendo en cuenta que los datos observacionales imponen restricciones a los resultados que se deben obtener de un modelo de síntesis planetaria para que estos últimos logren replicar las observaciones, los resultados de masas y distancias orbitales finales de los planetas generados con PSyCo no se acercan a las distribuciones correspondientes a los datos observacionales. Aún así, el código logra formar diferentes tipos de planetas y puede (y debe) ser mejorado para obtener resultados mas cercanos a los de las observaciones.

El modelo de síntesis PSyCo también logró reproducir las migraciones tipo I y II dentro del tiempo de evolución de cada sistema planetario. Algunos resultados de evolución de la masa de planetas junto con sus migraciones se muestran en las figuras 5.11

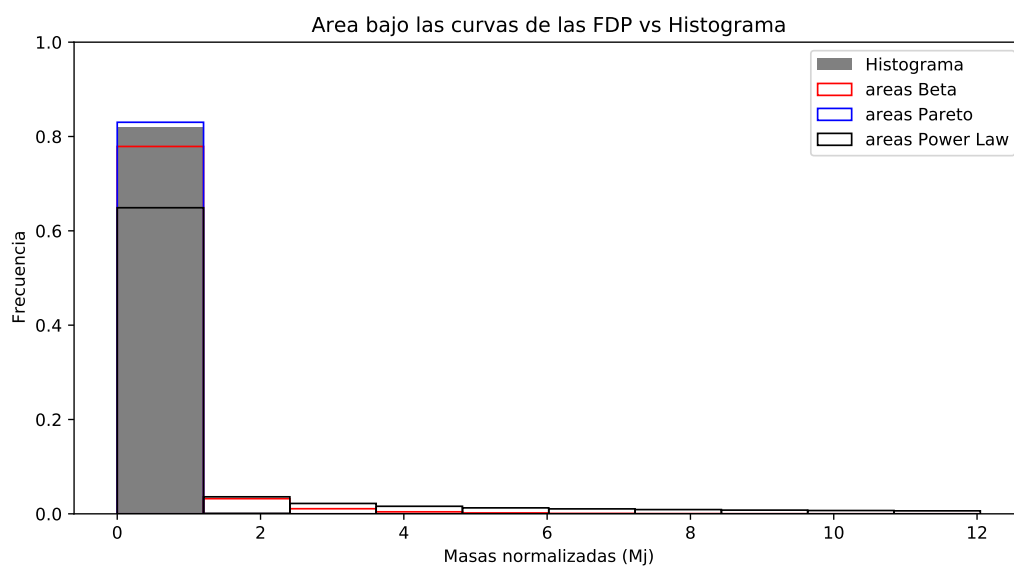


Figura 5.7: Comparación de las áreas bajo las curvas de las FDP propuestas vs el histograma de masas sintetizadas

En la figura 5.11 se observa que no en todas las masas solares se logran formar planetas gigantes.

Todos estos resultados y sus interpretaciones serán discutidos en las **Conclusiones**.

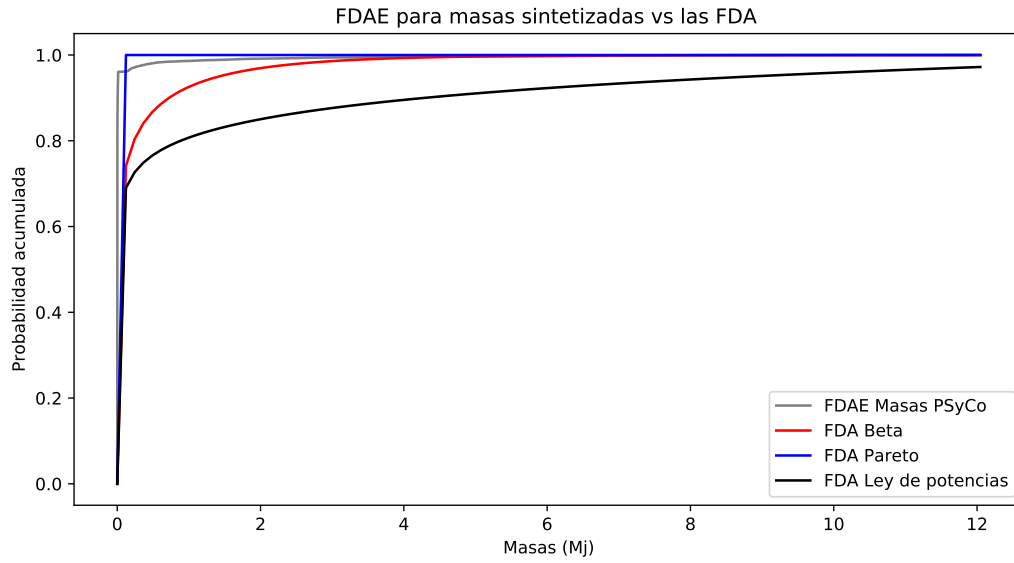


Figura 5.8: Comparación de la FDEA de las masas sintetizadas con las FDA de las distribuciones propuestas.

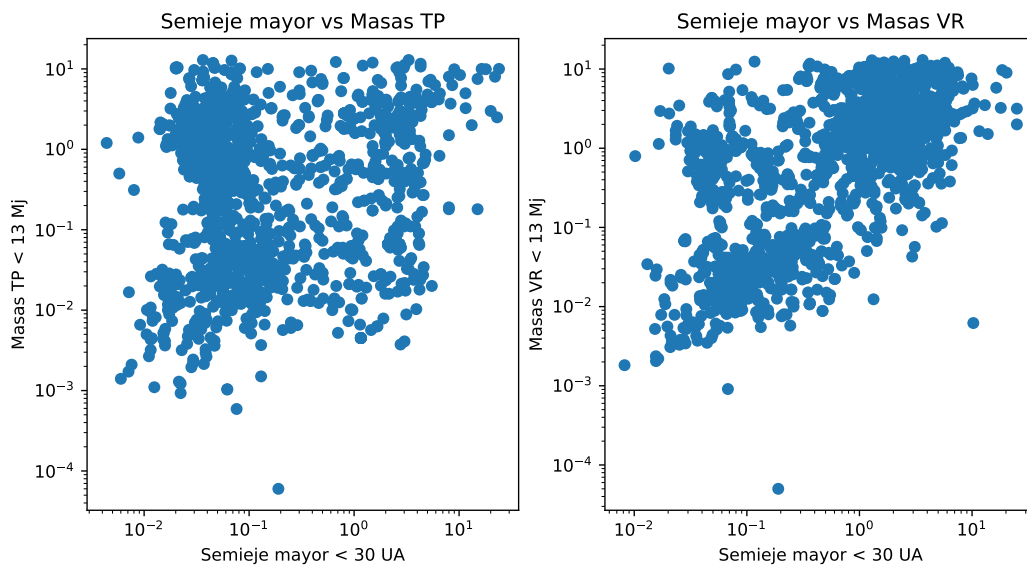


Figura 5.9: Gráficos de dispersión de las masas observadas y sus radios orbitales. *Izquierda*: masa por tránsito planetario. *Derecha*: masas por velocidad radial.

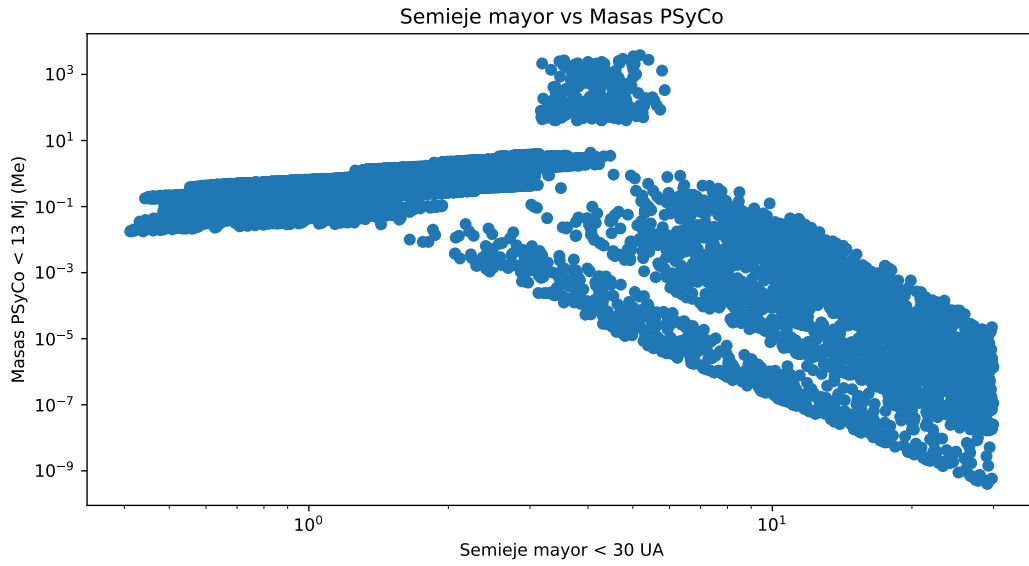


Figura 5.10: Síntesis planetaria con PSyCo. Masas vs semiejes mayores en una de las corridas del modelo.

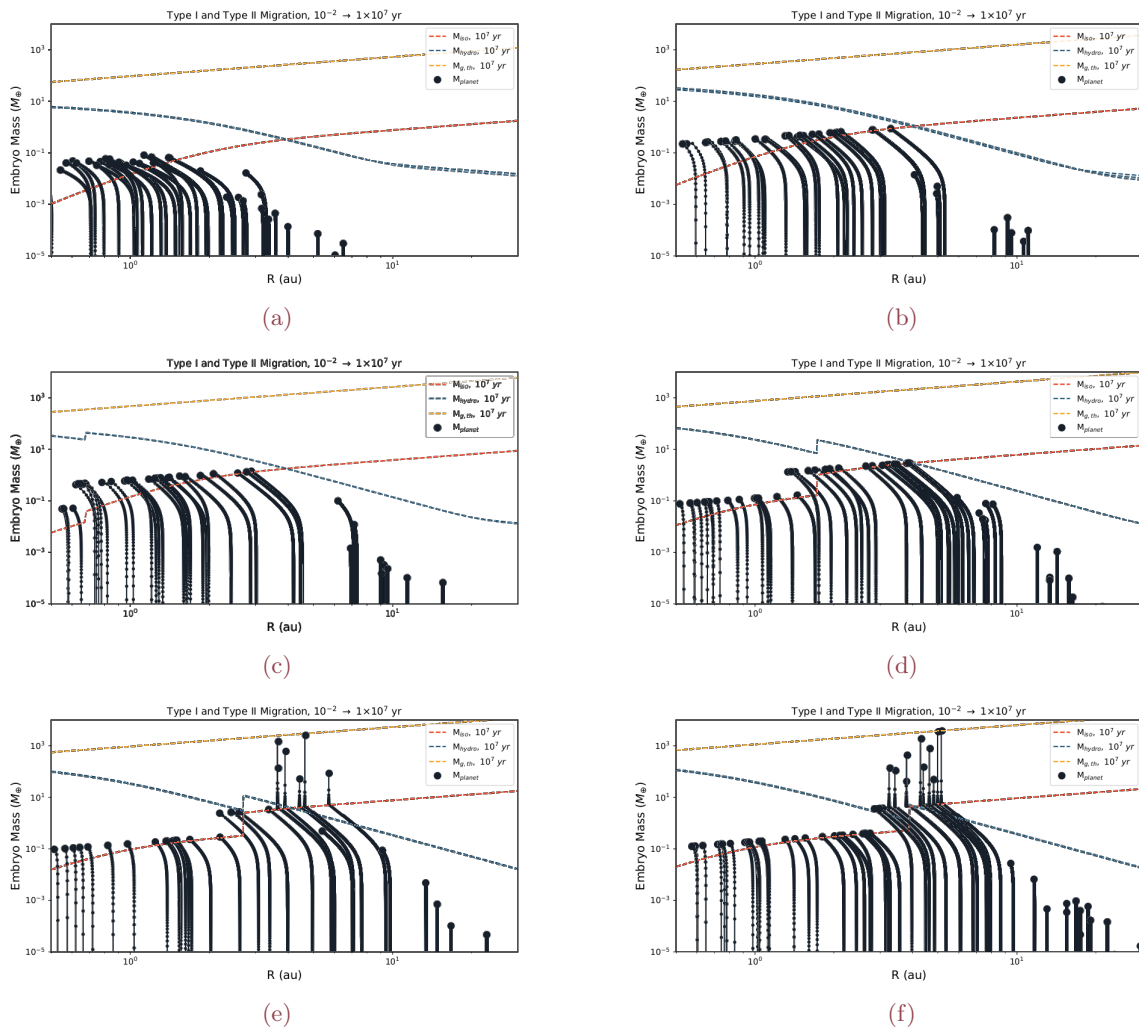


Figura 5.11: Evolución de la masa y la posición de planetas en el modelo PSyCo para diferentes masas estelares: (a). $0,1M_{\odot}$, (b). $0,3M_{\odot}$, (c). $0,5M_{\odot}$, (d). $0,8M_{\odot}$, (e). $1,0M_{\odot}$ y (f). $1,2M_{\odot}$

Conclusiones

Los resultados arrojados por los cálculos y simulaciones computacionales desarrollados durante esta investigación permiten llegar a las siguientes conclusiones respecto a los objetivos planteados:

- Los métodos de la estadística inferencial clásica, es decir, no Bayesiana, nos permiten inferir información sobre la población de planetas en el universo a partir de la muestra correspondiente a los exoplanetas confirmados. Aunque estas muestras poseen sesgos relacionados con los métodos e instrumentos empleados en la detección y caracterización de los exoplanetas y las condiciones físicas de la observación, la gran cantidad de planetas confirmados permiten tener confianza en las características de suficiencia, eficiencia, consistencia e insegmentación de los estimadores de los parámetros de las FDP propuestas.
- Los histogramas de las masas observadas tanto por velocidad radial como por tránsito planetario, dependiendo de la cantidad de clases en las que se desean clasificar estas masas, muestran “huecos” o rangos de masas que aún no han sido detectados. Estos “huecos” (llamados también “desiertos”), de acuerdo con [Ida and Lin \[2004b\]](#), pueden ser debidos a restricciones físicas durante la formación de planetas dentro de estos rangos de masas y tales restricciones están relacionadas con la tasa de acreción de planetesimales en determinados rangos de semiejes mayores. Exoplanetas con masas dentro del rango de estos “desiertos” quizás sí existan y puedan ser detectados en el futuro cuando se tengan mejores instrumentos. Estos rangos de masas no detectados pueden haber incidido en algunos de los resultados, lo que podría explicar la variabilidad de los mismos.
- En principio, los datos observacionales deben imponer ciertas restricciones a los resultados que deben producir los modelos de síntesis planetarias. En nuestro caso, la comparación entre la distribución de masas y semiejes mayores de los planetas confirmados en los catálogos y los generados por PSyCo muestran semejanzas en la diversidad de tipos de planetas y en los rangos de masas. Las diferencias en la distribución espacial final de los planetas alrededor de la estrella huésped pueden ser explicadas teniendo en cuenta que los planetas observados son aquellos más cercanos a su estrella y PSyCo produce planetas con una distribución de semiejes mayores aleatoria y en un rango amplio. De todas formas, PSyCo es un código factible de mejoras en muchos aspectos y la mejora en la capacidad de detección permitirá en un futuro confirmar la presencia de planetas a distancias mayores, de tal forma que su distribución sea comparable con la síntesis de PSyCo.
- PSyCo es un código que sigue en desarrollo y no es un trabajo terminado. Existen subrutinas del modelo que aún no se han incluido. Entre estos procesos faltantes está el que considera los efectos de la fotoevaporación y la interacción magnética entre el disco protoplanetario

y la estrella. Esto se hace evidente en el hecho de que PSyCo no resuelve numéricamente la ecuación de la evolución del disco (ecuación 2.3) sino que adopta la solución en ley de potencias empleada por [Ida and Lin, 2008] (ecuaciones 3.1) y esta solo contempla efectos dinámicos. Adicionalmente, aunque en esta investigación también se desarrolló un código de integración de N cuerpos que analiza la dinámica final de los planetoides y núcleos sintetizados, éste no se ha integrado a PSyCo de forma que se ejecute de manera concurrente dentro de todo el proceso de formación planetaria. Por ahora se ha usado luego de la síntesis de planetas para analizar la distribución final de objetos del sistema. Este integrador de N cuerpos es importante y necesario dentro del proceso de síntesis de sistemas planetarios porque las colisiones entre cuerpos sólidos dentro del disco protoplanetario afectan las distribuciones finales de semiejes mayores, excentricidades e inclinaciones de los planos orbitales.

- PSyCo logra sintetizar los rangos de masas planetarias que se esperarían en la formación de sistemas planetarios y, además, produce muchos objetos de masas pequeñas a distancias medianas que corresponderían a cinturones de planetoides.
- Los resultados de las pruebas de hipótesis basadas en la prueba χ^2 arrojaron que los datos se ajustan a las distribuciones Beta, Pareto y Ley de potencias. Este resultado no es sorprendente puesto que la interpretación que se hace en este tipo de pruebas es acerca de que la evidencia sea suficiente o no para rechazar la hipótesis nula. En otras palabras, los resultados de las pruebas χ^2 , tanto con el criterio del estadístico de prueba como el criterio del *valor-p* lo que muestran es que no hay evidencia suficiente para rechazar ninguna de las distribuciones propuestas pero no significa que la muestra se ajuste a las tres funciones. Podrían proponerse otras distribuciones cuya forma funcional sea semejante a las utilizadas en esta investigación y las pruebas de bondad de ajuste basadas en la prueba χ^2 quizás arrojen el mismo resultado: que no se pueden descartar. Ante esta situación se proponen tres posibles soluciones: 1) esperar que con el tiempo y el incremento acelerado de la cantidad de exoplanetas confirmados cada año se logre llegar a un número de exoplanetas en rangos de masa tales que algunas de las FDP propuestas ya no logre ajustarse y, por tanto, se descarte la hipótesis nula correspondiente y se reduzcan las opciones sobre las posibles distribuciones que describen correctamente la población de planetas en el universo. 2). Que en la actualidad exista (y no es conocida aún por este investigador) o en el futuro se desarrolle (si es que en la actualidad no existe) un nuevo método estadístico frecuentista que permita discernir, entre todas las posibles distribuciones para las cuales la prueba χ^2 no resulte en el rechazo de H_0 , la distribución a la que "mejor" se ajusten los datos observacionales. 3) Que existan métodos propios de la estadística Bayesiana que permitan discernir a cuál distribución se ajustan mejor los datos o al menos descartar opciones que no se logran descartar con métodos de la estadística frecuentista.
- El análisis de inferencia estadística hecho con las masas planetarias producidas por PSyCo dió como resultado el rechazo de la hipótesis nula para la distribución de Pareto. Esto es así por la cantidad de datos disponibles para la evaluación tanto del estadístico de prueba como del *valor-p* y está de acuerdo con la primera de las soluciones planteadas para la resolución del problema de que estas pruebas sirven para rechazar o descartar distribuciones y no para

confirmar alguna.

- Se sugiere continuar el desarrollo del código de PSyCo para incorporarle nuevas rutinas que permitan incluir más procesos físicos implicados en la formación de sistemas planetarios, más que planetas individuales. Esto será un trabajo del nivel de doctorado puesto que requiere dedicación exclusiva y mucha investigación acerca de los diversos modelos existentes, sus fortalezas, sus debilidades y muchas pruebas de ensayo y error para corregir parámetros de las ecuaciones que deben ser resueltas en cada etapa del proceso de formación y evolución de los planetas.
- Otras acciones que pueden explorarse en futuras investigaciones sobre la distribución de las masas planetarias podrían ser: *(a)* la transformación de los conjuntos de datos de masas sintetizadas para que, empleando criterios adecuados, se asemejen a las condiciones de los sistemas planetarios que son susceptibles de ser observados y *(b)* hacer la separación analítica de m_p y $\sin i$ a partir de la FDP de las masas medidas por velocidad radial y comparar el resultado con la FDP a la que se ajusten las masas medidas por tránsito planetario.

Elementos de Estadística

Los conceptos, definiciones y métodos estadísticos utilizados en esta investigación fueron elaborados a partir de los trabajos de [Blanco Castañeda \[2004\]](#), [Casella and Berger \[2002\]](#) y [Velasco Sotomayor and Wisniewski \[2001\]](#).

A.1. Definición de conceptos

La *estadística* es una ciencia que usa y analiza un conjunto de datos con el objetivo de hacer inferencias acerca de la población de la cual provienen estos datos.

- *Experimento aleatorio*: un experimento se dice que es aleatorio si su resultado no puede ser determinado de antemano.
- *Espacio muestral*: es el conjunto S de todos los resultados posibles de un experimento aleatorio.
- *Evento*: es cualquier subconjunto del espacio muestral S , incluyendo el propio S .
- *Probabilidad de un evento*: cuando el espacio muestral S es finito y está formado por eventos elementales que tienen la misma probabilidad de ocurrencia, la probabilidad de un evento A es un número real $P(A)$ que se obtiene calculando el cociente entre el número de resultados del evento A y el número de resultados del espacio muestral S , es decir,

$$P(A) = \frac{\text{Número de resultados del evento } A}{\text{Número de resultados del espacio muestral } S} \quad (\text{A.1})$$

- *Variable aleatoria*: es una función X cuyo dominio es el espacio muestral S y cuyo rango es el conjunto de los números reales \mathbb{R}
- *Variable aleatoria discreta*: se dice que una variable aleatoria X es discreta cuando el conjunto de todos los posibles valores de ésta son contables.
- *Función de masa de probabilidad (FMP)*: para una variable aleatoria discreta X , la f.m.p. es una función que asigna a cada elemento x del espacio muestral S la probabilidad p_x de que ocurra el evento $X = x$. Matemáticamente se define la FMP como

$$f(x) = P(X = x) = p_x \quad (\text{A.2})$$

- *Función de distribución acumulada (FDA)*: si X es una variable aleatoria discreta, la función dada por

$$F_X(x) = P(X \leq x) = \sum_{i=-\infty}^x P(X = i) \quad (\text{A.3})$$

para toda $-\infty < x < \infty$.

Una FDA tiene las siguientes propiedades:

- $\lim_{x \rightarrow -\infty} F_X(x) = 0$ y $\lim_{x \rightarrow \infty} F_X(x) = 1$
 - Si $a < b$, entonces $F(a) \leq F(b)$ para $a, b \in \mathbb{R}$.
- *Variable aleatoria continua*: se dice que una variable aleatoria X es continua cuando existe una función $f(x)$ continua tal que

$$P(a \leq X \leq b) = \int_a^b f(x)dx \quad (\text{A.4})$$

- *Función de densidad de probabilidad (FDP)*: es la función $f(x)$ que aparece en la ecuación A.4 y satisface las siguientes condiciones
- $f(x) \geq 0$ para $-\infty < x < \infty$.
 - $\int_{-\infty}^{\infty} f(x)dx = 1$.

En el caso de una variable aleatoria continua X , la FDA está dada por

$$F_X(x) = P(X \leq x) = \int_{-\infty}^x f(t)dt, \text{ para todo } x \quad (\text{A.5})$$

Adicionalmente, si $f(x)$ y $F(x)$ son las FDP y FDA respectivamente de una variable aleatoria continua X y $a, b \in X$, entonces

$$P(a \leq X \leq b) = F(b) - F(a), \text{ para } a \leq b \quad (\text{A.6})$$

y

$$f(x) = \frac{dF_X(X)}{dx} \quad (\text{A.7})$$

A.2. Pruebas de hipótesis

De acuerdo con el análisis que se hace con los datos de una muestra de una población y las conclusiones que se pueden obtener de ésta, se clasifica la Estadística en *descriptiva* e *inferencial*.

La **estadística descriptiva** es una parte de la estadística que busca obtener información de una población con base en el análisis de los datos de una muestra proveniente de ella. Este análisis se puede hacer a partir de gráficos y medidas de tendencia central. En esta sección se hablará únicamente del análisis gráfico.

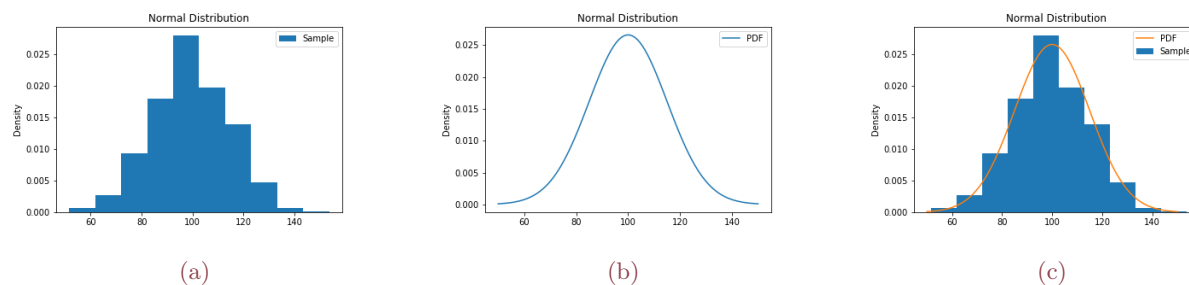


Figura A.1: (a). Histograma de densidades de 1000 números generados aleatoriamente a partir de una distribución normal. (b) Gráfica de una FDP normal. (c) Histograma de densidades de (a) y la FDP de (b) superpuestos [Dekermenjian, 2020].

A.2.1. Histograma

Un *histograma de frecuencias* o, simplemente, *histograma*, es un gráfico de barras que se obtiene a partir de los datos de la muestra. Para hacerlo, se ordenan los datos en forma creciente o decreciente y se organizan en *clases* o *categorías*. El número de datos que entra en cada clase se conoce con el nombre de *frecuencia de clase*. Para calcular el número k de clases en el que se dividirán los n datos de la muestra se suele emplear como regla

$$k = \sqrt{n}$$

Cada clase k es un intervalo de valores cuya cota inferior es cerrada. Gráficamente, el ancho de cada barra corresponde al valor de la clase y la altura de esta muestra el número de datos que pertenecen a cada clase. De esta manera, el histograma representa la distribución de frecuencias de la muestra y este gráfico se compara con el de la FDP a la que se desean ajustar los datos para inferir la distribución de la población. En ocasiones se usa un *histograma de densidades* que se obtiene a partir de la tabla de frecuencias dividiendo la cantidad de datos de cada clase por el ancho de la misma. Este tipo de histogramas conserva la misma forma del histograma de frecuencias pero la altura de cada barra tiene un valor entre 0 y 1. Este proceso de análisis gráfico de muestra en la figura A.1. Este primer análisis gráfico no es determinante sino exploratorio puesto que sirve de insumo para ver qué tipo de FDP podrían tener una forma que asemeje a la del histograma. En otras palabras, permite validar o no la intuición inicial del investigador.

La **estadística inferencial** es una parte de la estadística que busca obtener conclusiones basadas en la muestra y hacerlos válidos para toda la población de la cual proviene. Para ello, la estadística inferencial emplea las técnicas de *estimación de parámetros* y *pruebas de hipótesis*.

La estimación de parámetros es un proceso mediante el cual se desea saber cuáles son, para una determinada muestra, los valores más cercanos a los parámetros de la distribución que describe a la población de la cual proviene la muestra. Los métodos empleados por la estadística inferencial para la estimación de parámetros son la *Estimación de Máxima Verosimilitud* o *EMV* (en inglés, *Maximum Likelihood Estimation* o MLE) y el *Método de los Momentos*. En este trabajo se describirá el MLE, el usado en la investigación.

Sean X_1, X_2, \dots, X_n los datos de una muestra aleatoria obtenida a partir de una población cuya FDP está dada por $f(x|\theta_1, \theta_2, \dots, \theta_n)$ donde $\theta_1, \theta_2, \dots, \theta_n$ son parámetros desconocidos que van a ser estimados.

Agradecimientos

Quiero agradecer a todos los elementos humanos, circunstanciales y (¿por qué no?) sobrenaturales que hicieron posible la culminación exitosa de este trabajo.

En primer lugar, quiero agradecer a aquello, etéreo para algunos y concreto para otros, que me permitió seguir adelante cuando sentía que había perdido el sentido de lo que estaba haciendo, porque estaba afectando mi salud, mis relaciones familiares, personales, laborales y de otra índole. Aquello que hace que en ocasiones, para un científico como yo, las cosas se den en contra de toda lógica puesto que su resolución no estaba dentro de las posibilidades. Ese aquello a lo que me refiero es para mí algo concreto y que solamente es posible accederlo a partir de una experiencia de vida. Quiero agradecer a Dios, motor, impulso y esperanza cuando la batalla está perdida y siento que he caído vencido.

En segundo lugar quiero agradecer a mi familia y mi hogar. Si bien la familia (padre, madre y hermanos) son para siempre y nunca podré ser un ex-hijo o ex-hermano, siempre me apoyaron, me alentaron y me perdonaron las veces que decidí no estar con ellos porque tenía compromisos académicos de los cuales ocuparme. A mi hogar, que son mi hijo y mi esposa, sobre todo a ella, porque son los que tuvieron que vivir mi ausencia durante mi presencia con ellos: ignorarlos tantas veces, desatender sus reclamos justos, optar por hacer algo no tan fácil ni agradable en lugar de pasar con ellos sus mejores momentos. Si bien nunca podré ser ex-papá de mi hijo, en varias ocasiones sentí que podría llegar a ser ex-esposo de la mujer que decidió dedicarme su vida a cambio de que yo hiciera lo mismo por ella. Si ustedes no hubiera sido posible levantarme de nuevo en mis caídas.

En tercer lugar quiero agradecer a mi asesor de trabajo de grado por su paciencia: le causé desilusiones y rabias y él solamente me dió ánimos, consejo, sabiduría y todo lo mejor de él como ser humano, como profesional y como científico. Aquí también agradezco a mi co-asesora porque sin su experiencia en la estadística y su instinto, su visión, no habríamos encontrado el camino correcto hacia lo que queríamos: pensábamos investigar sobre algo que ella misma nos mostró que ya estaba resuelto y nos iluminó con su intuición hacia dónde debíamos seguir. Sus aportes, dedicación y paciencia nos mostraron un mundo desconocido para nosotros y que nos permitió ver el universo de otra manera.

Lo que soy en este momento y seré en adelante se lo debo a todos ustedes.

Medellín, 2022

Fabio Onier Osorio Peláez

Bibliografía

- R. Alexander, I. Pascucci, S. Andrews, P. Armitage, L. Cieza, H. Beuther, R. Klessen, C. Dullemond, and T. Henning. Protostars and planets vi, 2014.
- R. D. Alexander and P. J. Armitage. Giant planet migration, disk evolution, and the origin of transitional disks. *The Astrophysical Journal*, 704(2):989, 2009.
- Y. Alibert and W. Benz. Formation and composition of planets around very low mass stars. *Astronomy & Astrophysics*, 598:L5, 2017.
- Y. Alibert, C. Mordasini, and W. Benz. Migration and giant planet formation. *Astronomy & Astrophysics*, 417(1):L25–L28, 2004.
- Y. Alibert, C. Mordasini, W. Benz, and C. Winisdoerffer. Models of giant planet formation with migration and disc evolution. *Astronomy & Astrophysics*, 434(1):343–353, 2005.
- Y. Alibert, C. Mordasini, and W. Benz. Extrasolar planet population synthesis-iii. formation of planets around stars of different masses. *Astronomy & Astrophysics*, 526:A63, 2011.
- Y. Alibert, F. Carron, A. Fortier, S. Pfyffer, W. Benz, C. Mordasini, and D. Swoboda. Theoretical models of planetary system formation: mass vs. semi-major axis. *Astronomy & Astrophysics*, 558:A109, 2013.
- American Meteorological Society. Scale height, 2022.
- V. I. Ananyeva, A. E. Ivanova, A. A. Venkstern, I. A. Shashkova, A. V. Yudaev, A. V. Tavrov, O. I. Korablev, and J.-L. Bertaux. Mass distribution of exoplanets considering some observation selection effects in the transit detection technique. *Icarus*, 346:113773, 2020.
- P. André, J. Di Francesco, D. Ward-Thompson, S. Inutsuka, R. Pudritz, J. Pineda, H. Beuther, R. S. Klessen, C. P. Dullemond, and T. Henning. Protostars and planets vi, 2014.
- S. M. Andrews, D. Wilner, A. Hughes, C. Qi, and C. Dullemond. Protoplanetary disk structures in ophiuchus. ii. extension to fainter sources. *The Astrophysical Journal*, 723(2):1241, 2010.
- P. J. Armitage. *Astrophysics of planet formation*. Cambridge University Press, 2020.
- L. Arnold, F. Bouchy, and C. Moutou. Tenth anniversary of 51 peg-b: Status of and prospects for hot jupiter studies. In *Tenth Anniversary of 51 Peg-b: Status of and prospects for hot Jupiter studies*, 2006.

- A. Baranne, D. Queloz, M. Mayor, G. Adrianzyk, G. Knispel, D. Kohler, D. Lacroix, J.-P. Meunier, G. Rimbaud, and A. Vin. Elodie: A spectrograph for accurate radial velocity measurements. *Astronomy and Astrophysics Supplement Series*, 119(2):373–390, 1996.
- W. Benz, S. Ida, Y. Alibert, D. Lin, and C. Mordasini. Planet population synthesis. *arXiv preprint arXiv:1402.7086*, 2014.
- R. Bernstein and S. Bernstein. *Schaum's outline of elements of statistics II: Inferential statistics*. McGraw Hill Professional, 1999.
- T. Birnstiel, H. Klahr, and B. Ercolano. A simple model for the evolution of the dust population in protoplanetary disks. *Astronomy & Astrophysics*, 539:A148, 2012.
- L. Blanco Castañeda. Probabilidad. *Universidad Nacional de Colombia, Bogotá*, 2004.
- P. Bodenheimer and J. B. Pollack. Calculations of the accretion and evolution of giant planets: The effects of solid cores. *Icarus*, 67(3):391–408, 1986.
- P. Bodenheimer, O. Hubickyj, and J. J. Lissauer. Models of the in situ formation of detected extrasolar giant planets. *Icarus*, 143(1):2–14, 2000.
- H. Bondi. On spherically symmetrical accretion. *Monthly Notices of the Royal Astronomical Society*, 112(2):195–204, 1952.
- F. Bouchy and S. Team. Sophie: the successor of the spectrograph elodie for extrasolar planet search and characterization. In *Tenth Anniversary of 51 Peg-b: Status of and prospects for hot Jupiter studies*, pages 319–325, 2006.
- R. Burn, M. Schlecker, C. Mordasini, A. Emsenhuber, Y. Alibert, T. Henning, H. Klahr, and W. Benz. The new generation planetary population synthesis (ngpps)-iv. planetary systems around low-mass stars. *Astronomy & Astrophysics*, 656:A72, 2021.
- G. Casella and R. L. Berger. *Statistical inference*, volume 2. Duxbury Pacific Grove, CA, 2002.
- I. M. Chakravarti, R. G. Laha, and J. Roy. Handbook of methods of applied statistics. *Wiley Series in Probability and Mathematical Statistics (USA) eng*, 1967.
- J. Chambers. An analytic model for the evolution of a viscous, irradiated disk. *The Astrophysical Journal*, 705(2):1206, 2009.
- J. L. Christiansen. Exoplanet catalogues. *arXiv preprint arXiv:1803.11158*, 2018.
- P. Cresswell and R. P. Nelson. Three-dimensional simulations of multiple protoplanets embedded in a protostellar disc. *Astronomy & Astrophysics*, 482(2):677–690, 2008.
- A. Crida, A. Morbidelli, and F. Masset. On the width and shape of gaps in protoplanetary disks. *Icarus*, 181(2):587–604, 2006.
- J. N. Cuzzi, R. C. Hogan, and K. Shariff. Toward planetesimals: Dense chondrule clumps in the protoplanetary nebula. *The Astrophysical Journal*, 687(2):1432, 2008.

- J. de Wit and S. Seager. Constraining exoplanet mass from transmission spectroscopy. *Science*, 342(6165):1473–1477, 2013a.
- J. de Wit and S. Seager. Constraining exoplanet mass from transmission spectroscopy. *Science*, 342(6165):1473–1477, dec 2013b. doi: 10.1126/science.1245450. URL <https://doi.org/10.1126%2Fscience.1245450>.
- G. Dekermenjian. Probability distributions using scipy, 2020.
- K.-M. Dittkrist, C. Mordasini, H. Klahr, Y. Alibert, and T. Henning. Impacts of planet migration models on planetary populations-effects of saturation, cooling and stellar irradiation. *Astronomy & Astrophysics*, 567:A121, 2014.
- J. Durbin. Kolmogorov-smirnov tests when parameters are estimated. In *Empirical Distributions and Processes*, pages 33–44. Springer, 1976.
- G. D’Angelo and S. H. Lubow. Evolution of migrating planets undergoing gas accretion. *The Astrophysical Journal*, 685(1):560, 2008.
- A. Emsenhuber, C. Mordasini, R. Burn, Y. Alibert, W. Benz, and E. Asphaug. The new generation planetary population synthesis (ngpps)-ii. planetary population of solar-like stars and overview of statistical results. *Astronomy & Astrophysics*, 656:A70, 2021a.
- A. Emsenhuber, C. Mordasini, R. Burn, Y. Alibert, W. Benz, and E. Asphaug. The new generation planetary population synthesis (ngpps)-i. bern global model of planet formation and evolution, model tests, and emerging planetary systems. *Astronomy & Astrophysics*, 656:A69, 2021b.
- L. Euler. De progressionibus harmonicis observationes. *Commentarii academiae scientiarum Petropolitanae*, 7:150–161, 1740.
- Exoplanet Team. Catalog, 2022.
- J. Faherty, J. Gagne, A. Weinberger, E. Gonzales, R. Kiman, and M. Popinchalk. Brown dwarfs, exoplanets, what’s the difference? In *American Astronomical Society Meeting Abstracts*, volume 53, pages 314–02, 2021.
- A. Fortier, Y. Alibert, F. Carron, W. Benz, and K.-M. Dittkrist. Planet formation models: the interplay with the planetesimal disc. *Astronomy & Astrophysics*, 549:A44, 2013.
- L. Fouchet, Y. Alibert, C. Mordasini, and W. Benz. Effects of disk irradiation on planet population synthesis. *Astronomy & Astrophysics*, 540:A107, 2012.
- P. Goldreich and S. Tremaine. The excitation of density waves at the lindblad and corotation resonances by an external potential. *Astrophysical Journal*, 233(3):857–871, 1979.
- P. Goldreich and S. Tremaine. Disk-satellite interactions. *Astrophysical Journal*, 241(1):425–441, 1980.
- R. Greenberg, J. F. Wacker, W. K. Hartmann, and C. R. Chapman. Planetesimals to planets: Numerical simulation of collisional evolution. *Icarus*, 35(1):1–26, 1978.

- Y. Greenzweig and J. J. Lissauer. Accretion rates of protoplanets: ii. gaussian distributions of planetesimal velocities. *Icarus*, 100(2):440–463, 1992.
- Y. Hasegawa and R. E. Pudritz. Evolutionary tracks of trapped, accreting protoplanets: The origin of the observed mass–period relation. *The Astrophysical Journal*, 760(2):117, 2012.
- C. Hayashi. Structure of the solar nebula, growth and decay of magnetic fields and effects of magnetic and turbulent viscosities on the nebula. *Progress of Theoretical Physics Supplement*, 70:35–53, 1981.
- S. Ida and D. Lin. Toward a deterministic model of planetary formation. ii. the formation and retention of gas giant planets around stars with a range of metallicities. *The Astrophysical Journal*, 616(1):567, 2004a.
- S. Ida and D. Lin. Toward a deterministic model of planetary formation. iii. mass distribution of short-period planets around stars of various masses. *The Astrophysical Journal*, 626(2):1045, 2005.
- S. Ida and D. Lin. Toward a deterministic model of planetary formation. iv. effects of type i migration. *The Astrophysical Journal*, 673(1):487, 2008.
- S. Ida and D. Lin. Toward a deterministic model of planetary formation. vi. dynamical interaction and coagulation of multiple rocky embryos and super-earth systems around solar-type stars. *The Astrophysical Journal*, 719(1):810, 2010.
- S. Ida and D. N. Lin. Toward a deterministic model of planetary formation. i. a desert in the mass and semimajor axis distributions of extrasolar planets. *The Astrophysical Journal*, 604(1):388, 2004b.
- S. Ida and J. Makino. Scattering of planetesimals by a protoplanet: Slowing down of runaway growth. *Icarus*, 106(1):210–227, 1993.
- S. Ida, D. Lin, and M. Nagasawa. Toward a deterministic model of planetary formation. vii. eccentricity distribution of gas giants. *The Astrophysical Journal*, 775(1):42, 2013.
- M. Ikoma, K. Nakazawa, and H. Emori. Formation of giant planets: dependences on core accretion rate and grain opacity. *The Astrophysical Journal*, 537(2):1013, 2000.
- A. Ivanova, V. Ananyeva, A. Venkstern, I. Shashkova, A. Yudaev, A. Tavrov, O. Korablev, and J.-L. Bertaux. The mass distribution of transiting exoplanets corrected for observational selection effects. *Astronomy Letters*, 45(10):687–694, 2019.
- S. Jin and C. Mordasini. Compositional imprints in density–distance–time: a rocky composition for close-in low-mass exoplanets from the location of the valley of evaporation. *The Astrophysical Journal*, 853(2):163, 2018.
- S. Jin, C. Mordasini, V. Parmentier, R. Van Boekel, T. Henning, and J. Ji. Planetary population synthesis coupled with atmospheric escape: a statistical view of evaporation. *The Astrophysical Journal*, 795(1):65, 2014.

- A. Jorissen, M. Mayor, and S. Udry. The distribution of exoplanet masses. *Astronomy & Astrophysics*, 379(3):992–998, 2001.
- H. Keutelian. The kolmogorov-smirnov test when parameters are estimated from data. *CDF Notes*, 1991.
- E. Kokubo and S. Ida. On runaway growth of planetesimals. *Icarus*, 123(1):180–191, 1996.
- E. Kokubo and S. Ida. Oligarchic growth of protoplanets. *Icarus*, 131(1):171–178, 1998.
- E. Kokubo and S. Ida. Formation of protoplanets from planetesimals in the solar nebula. *Icarus*, 143(1):15–27, 2000.
- E. Kokubo and S. Ida. Formation of protoplanet systems and diversity of planetary systems. *The Astrophysical Journal*, 581(1):666, 2002.
- M. Lambrechts and A. Johansen. Rapid growth of gas-giant cores by pebble accretion. *Astronomy & Astrophysics*, 544:A32, 2012.
- M. Lambrechts and A. Johansen. Forming the cores of giant planets from the radial pebble flux in protoplanetary discs. *Astronomy & Astrophysics*, 572:A107, 2014.
- E. L. Lehmann. *Elements of large-sample theory*. Springer, 1999.
- D. Lin and J. Papaloizou. On the structure of circumbinary accretion disks and the tidal evolution of commensurable satellites. *Monthly Notices of the Royal Astronomical Society*, 188(2):191–201, 1979.
- D. N. Lin and J. Papaloizou. On the tidal interaction between protoplanets and the protoplanetary disk. iii-orbital migration of protoplanets. *The Astrophysical Journal*, 309:846–857, 1986.
- C. Lovis, D. Fischer, et al. Radial velocity techniques for exoplanets. *Exoplanets*, pages 27–53, 2010.
- R. Lüst. Die entwicklung einer um einen zentralkörper rotierenden gasmasse. i. lösungen der hydrodynamischen gleichungen mit turbulenter reibung. *Zeitschrift Naturforschung Teil A*, 7(1):87–98, 1952.
- D. Lynden-Bell and J. E. Pringle. The evolution of viscous discs and the origin of the nebular variables. *Monthly Notices of the Royal Astronomical Society*, 168(3):603–637, 1974.
- G. Marcy, R. P. Butler, D. Fischer, S. Vogt, J. T. Wright, C. G. Tinney, and H. R. Jones. Observed properties of exoplanets: masses, orbits, and metallicities. *Progress of Theoretical Physics Supplement*, 158:24–42, 2005.
- G. W. Marcy, R. P. Butler, D. A. Fischer, and S. S. Vogt. Properties of extrasolar planets. In *Scientific frontiers in research on extrasolar planets*, volume 294, pages 1–16, 2003.
- G. W. Marcy, A. W. Howard, et al. The occurrence and the distribution of masses and radii of exoplanets. *Proceedings of the International Astronomical Union*, 6(S276):3–12, 2010.

- F. Masset, A. Morbidelli, A. Crida, and J. Ferreira. Disk surface density transitions as protoplanet traps. *The Astrophysical Journal*, 642(1):478, 2006.
- F. J. Massey Jr. The kolmogorov-smirnov test for goodness of fit. *Journal of the American statistical Association*, 46(253):68–78, 1951.
- M. Mayor and D. Queloz. A jupiter-mass companion to a solar-type star. *Nature*, 378(6555):355–359, 1995.
- M. Mayor, M. Marmier, C. Lovis, S. Udry, D. Ségransan, F. Pepe, W. Benz, J.-L. Bertaux, F. Bouchy, X. Dumusque, et al. The harps search for southern extra-solar planets xxxiv. occurrence, mass distribution and orbital properties of super-earths and neptune-mass planets. *arXiv preprint arXiv:1109.2497*, 2011.
- L. Mishra, Y. Alibert, A. Leleu, A. Emsenhuber, C. Mordasini, R. Burn, S. Udry, and W. Benz. The new generation planetary population synthesis (ngpps) vi. introducing kobe: Kepler observes bern exoplanets-theoretical perspectives on the architecture of planetary systems: Peas in a pod. *Astronomy & Astrophysics*, 656:A74, 2021.
- A. Morbidelli and D. Nesvorny. Dynamics of pebbles in the vicinity of a growing planetary embryo: hydro-dynamical simulations. *Astronomy & Astrophysics*, 546:A18, 2012.
- C. Mordasini. Grain opacity and the bulk composition of extrasolar planets-ii. an analytical model for grain opacity in protoplanetary atmospheres. *Astronomy & Astrophysics*, 572:A118, 2014.
- C. Mordasini. Planetary population synthesis. *arXiv preprint arXiv:1804.01532*, 2018.
- C. Mordasini, Y. Alibert, and W. Benz. Extrasolar planet population synthesis-i. method, formation tracks, and mass-distance distribution. *Astronomy & Astrophysics*, 501(3):1139–1160, 2009a.
- C. Mordasini, Y. Alibert, and W. Benz. Extrasolar planet population synthesis-i. method, formation tracks, and mass-distance distribution. *Astronomy & Astrophysics*, 501(3):1139–1160, 2009b.
- C. Mordasini, Y. Alibert, W. Benz, and D. Naef. Extrasolar planet population synthesis-ii. statistical comparison with observations. *Astronomy & Astrophysics*, 501(3):1161–1184, 2009c.
- C. Mordasini, Y. Alibert, C. Georgy, K.-M. Dittkrist, H. Klahr, and T. Henning. Characterization of exoplanets from their formation-ii. the planetary mass-radius relationship. *Astronomy & Astrophysics*, 547:A112, 2012a.
- C. Mordasini, Y. Alibert, H. Klahr, and T. Henning. Characterization of exoplanets from their formation-i. models of combined planet formation and evolution. *Astronomy & Astrophysics*, 547:A111, 2012b.
- C. Mordasini, P. Mollière, K.-M. Dittkrist, S. Jin, and Y. Alibert. Global models of planet formation and evolution. *International Journal of Astrobiology*, 14(2):201–232, 2015.

- C. Mordasini, G.-D. Marleau, and P. Mollière. Characterization of exoplanets from their formation-iii. the statistics of planetary luminosities. *Astronomy & Astrophysics*, 608:A72, 2017.
- K. Nakazawa, S. Ida, and Y. Nakagawa. Collisional probability of planetesimals revolving in the solar gravitational field. i-basic formulation. *Astronomy and Astrophysics*, 220:293–300, 1989.
- NASA. Nasa exoplanet archive: A service of nasa exoplanet science institute. <https://exoplanetarchive.ipac.caltech.edu/>, 2022. Accessed: 2022-06-20.
- N. Ndugu, B. Bitsch, and E. Jurua. Planet population synthesis driven by pebble accretion in cluster environments. *Monthly Notices of the Royal Astronomical Society*, 474(1):886–897, 2018.
- A. R. Neil and L. A. Rogers. A joint mass–radius–period distribution of exoplanets. *The Astrophysical Journal*, 891(1):12, 2020.
- C. Ormel and H. Klahr. The effect of gas drag on the growth of protoplanets-analytical expressions for the accretion of small bodies in laminar disks. *Astronomy & Astrophysics*, 520:A43, 2010.
- C. Ormel, C. Dullemond, and M. Spaans. Accretion among preplanetary bodies: the many faces of runaway growth. *Icarus*, 210(1):507–538, 2010.
- S.-J. Paardekooper, C. Baruteau, A. Crida, and W. Kley. A torque formula for non-isothermal type I planetary migration – I. Unsaturated horseshoe drag. *Monthly Notices of the Royal Astronomical Society*, 401(3):1950–1964, 01 2010a. ISSN 0035-8711. doi: 10.1111/j.1365-2966.2009.15782.x. URL <https://doi.org/10.1111/j.1365-2966.2009.15782.x>.
- S.-J. Paardekooper, C. Baruteau, A. Crida, and W. Kley. A torque formula for non-isothermal type i planetary migration–i. unsaturated horseshoe drag. *Monthly Notices of the Royal Astronomical Society*, 401(3):1950–1964, 2010b.
- O. Panić, M. R. Hogerheijde, D. Wilner, and C. Qi. A break in the gas and dust surface density of the disc around the t tauri star im lupi. *Astronomy & Astrophysics*, 501(1):269–278, 2009.
- L. Pasquini, A. Manescau, G. Avila, B. Delabre, H. Dekker, J. Liske, S. D’Odorico, F. Pepe, M. Dessauges, C. Lovis, et al. Espresso: a high resolution spectrograph for the combined coude focus of the vlt. In *Science with the VLT in the ELT Era*, pages 395–399. Springer, 2009.
- F. Pepe, M. Mayor, B. Delabre, D. Kohler, D. Lacroix, D. Queloz, S. Udry, W. Benz, J.-L. Bertaux, and J.-P. Sivan. Harps: a new high-resolution spectrograph for the search of extrasolar planets. In *Optical and IR telescope instrumentation and detectors*, volume 4008, pages 582–592. SPIE, 2000.
- S. Pfyffer, Y. Alibert, W. Benz, and D. Swoboda. Theoretical models of planetary system formation-ii. post-formation evolution. *Astronomy & Astrophysics*, 579:A37, 2015.
- M. Podolak, J. B. Pollack, and R. T. Reynolds. Interactions of planetesimals with protoplanetary atmospheres. *Icarus*, 73(1):163–179, 1988.

- D. Queloz, M. Mayor, S. Udry, M. Burnet, F. Carrier, A. Eggenberger, D. Naef, N. Santos, F. Pepe, G. Rupprecht, et al. From coralie to harps. the way towards 1 m s-1 precision doppler measurements. *The Messenger*, 105:1–7, 2001.
- S. N. Raymond and A. Morbidelli. Planet formation: key mechanisms and global models. In *Demographics of Exoplanetary Systems*, pages 3–82. Springer, 2022.
- P. Rodó. Propiedades de los estimadores. <https://economipedia.com/definiciones/propiedades-de-los-estimadores.html>, 2019. Accessed: 2022-09-5.
- V. Safronov. Evolution of the protoplanetary cloud and formation of the earth and planets, trans. *NASA TTF-677 (1972)(Nauka, Moscow)(Russian)*, 1969.
- M. Schlecker, C. Mordasini, A. Emsenhuber, H. Klahr, T. Henning, R. Burn, Y. Alibert, and W. Benz. The new generation planetary population synthesis (ngpps)-iii. warm super-earths and cold jupiters: a weak occurrence correlation, but with a strong architecture-composition link. *Astronomy & Astrophysics*, 656:A71, 2021a.
- M. Schlecker, D. Pham, R. Burn, Y. Alibert, C. Mordasini, A. Emsenhuber, H. Klahr, T. Henning, and L. Mishra. The new generation planetary population synthesis (ngpps)-v. predetermination of planet types in global core accretion models. *Astronomy & Astrophysics*, 656:A73, 2021b.
- J. Schneider, C. Dedieu, P. Le Sidaner, R. Savalle, and I. Zolotukhin. Defining and cataloging exoplanets: the exoplanet. eu database. *Astronomy & Astrophysics*, 532:A79, 2011.
- S. Seager. Exoplanet atmospheres. In *Exoplanet Atmospheres*. Princeton University Press, 2010.
- N. I. Shakura and R. A. Sunyaev. Black holes in binary systems. observational appearance. *Astronomy and Astrophysics*, 24:337–355, 1973.
- M. A. Stephens. EDF statistics for goodness of fit and some comparisons. *Journal of the American statistical Association*, 69(347):730–737, 1974.
- H. Tanaka, T. Takeuchi, and W. R. Ward. Three-dimensional interaction between a planet and an isothermal gaseous disk. i. corotation and lindblad torques and planet migration. *The Astrophysical Journal*, 565(2):1257, 2002.
- The Planetary Society. The radial velocity graph of 51 pegasi. <https://www.planetary.org/space-images/\radial-velocity-graph-51-pegasi>, 2022. Accessed: 2022-04-17.
- A. Thiabaud, U. Marboeuf, Y. Alibert, I. Leya, and K. Mezger. Gas composition of the main volatile elements in protoplanetary discs and its implication for planet formation. *Astronomy & Astrophysics*, 574:A138, 2015.
- E. W. Thommes, M. J. Duncan, and H. F. Levison. Oligarchic growth of giant planets. *Icarus*, 161(2):431–455, 2003.
- A. S. M. Uzsoy, L. A. Rogers, and E. M. Price. Radius and mass distribution of ultra-short-period planets. *The Astrophysical Journal*, 919(1):26, 2021.

- G. Velasco Sotomayor and P. M. Wisniewski. Probabilidad y estadística para ingeniería y ciencias. *México: Thomson Learning*, 2001.
- C. Viedma. *Estadística descriptiva e inferencial*. Madrid: Ediciones IDT, 2018.
- R. von Steiger and T. Zurbuchen. Solar metallicity derived from in situ solar wind composition. *The Astrophysical Journal*, 816(1):13, 2015.
- J. Wang, J. Ge, X. Wan, N. De Lee, and B. Lee. Accurate group delay measurement for radial velocity instruments using the dispersed fixed delay interferometer method. ii. application of heterodyne combs using an external interferometer filter. *Publications of the Astronomical Society of the Pacific*, 124(921):1159, 2012.
- W. R. Ward. Density waves in the solar nebula: Differential lindblad torque. *icarus*, 67(1):164–180, 1986.
- S. Weidenschilling. The distribution of mass in the planetary system and solar nebula. *Astrophysics and Space Science*, 51(1):153–158, 1977.
- G. Wetherill and G. Stewart. Formation of planetary embryos: Effects of fragmentation, low relative velocity, and independent variation of eccentricity and inclination. *Icarus*, 106(1):190–209, 1993.

Abreviaciones

VR	Velocidad Radial
TP	Tránsito Planetario
FDP	Función de Densidad de Probabilidad
FDA	Función de Densidad Acumulada
FDAE	Función de Densidad Acumulada Empírica
FMP	Función de masa de probabilidad
CM	Centro de masa
m_*	Masa de la estrella huésped
m_p	Masa del exoplaneta
M_{Jup}	Masa de Júpiter
M_\odot	Masa del Sol
HARPS	High Accuracy Radial velocity Planet Searcher
MARVELS	Multi-object Apache Point Observatory Radial Velocity Exoplanet Large-area Survey
SOPHIE	Spectrographe pour l'Observation des Phénomènes des Intérieurs stellaires et des Exoplanètes
ESPRESSO	Echelle SPectrograph for Rocky Exoplanets and Stable Spectroscopic Observations
EMV	Estimación de Máxima Verosimilitud
MLE	Maximun Likelihood Estimation
MMSN	Minimun Mass Solar Nebula
MMR	Mean-Motion Resonance
PSYCO	Population Synthesis Code

