

Parcial Física Computacional

Óscar Andrés Calvo Rebellón¹, Julian Andres Barahona Giraldo²

¹oscara.calvor@uqvirtual.edu.co

²jabarahonag@uqvirtual.edu.co

Universidad del Quindío, Departamento de Física

La base de datos analizada es una base de datos en formato tabular 2D con extensión csv, la cual son datos del telescopio espectroscópico **LAMOST**¹ de baja resolución publicado por Weitao Zhang y demás autores [1], los datos son a espectros de Ca II (Calcio una vez ionizado), para observar el comportamiento de actividades cromosfericas de estrellas de tipo solar o mas específico, estrellas de tipo **G**² con luminosidad **V**³ que son estrellas con una masa alrededor de $[0.8M_{\odot}, 1.2M_{\odot}]$, dado que esa base de datos es de alrededor de 1.330.654 espectros⁴; los autores analizaron dichos datos en un clúster de servidores por lo cual nuestra potencia computacional no es suficiente para abarcar toda la base datos, por lo que se corta a 5.000 filas la base de datos teniendo la suficiente información para hacer un análisis. Para la limpieza de datos se observa que hay 5 filas donde hay valores NAN, la razón es porque esas estrellas no están registradas en el catalogo Gaia DR2, el cual soluciona 5 parámetros astronómicos en las estrellas del catalogo: posición del cielo (esfera celeste) dada coordenadas celestes (α, δ), paralaje, movimiento propio, brillo, magnitud absoluta.

Para mostrar las primeras relaciones de la tabla, se observa la relación señal- ruido en dos diferentes bandas en la verde(snr_g) y en la roja(snr_r), teniendo la gráfica:

¹telescopio espectroscópico de fibra multi-objeto de área de cielo grande

²Estrella en la secuencia principal

³Lo que condiciona que sean de tipo enanas

⁴Alrededor de 6 millones de filas

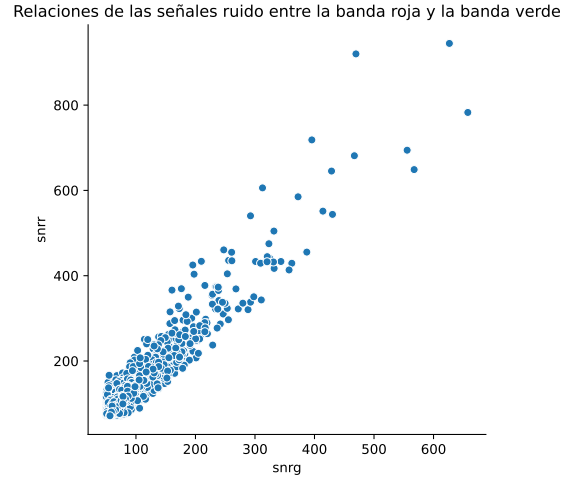


Fig. 1: Relación entre la señal-ruido en la banda verde y la señal-ruido en la banda roja, se observa cierto comportamiento lineal entre las dos, en el arreglo de los puntos o parámetros físicos.

Se observa que los puntos tienen un comportamiento lineal⁵, por lo que hacemos una regresión lineal en los datos y se obtiene una relación entre en snrr y snrg:

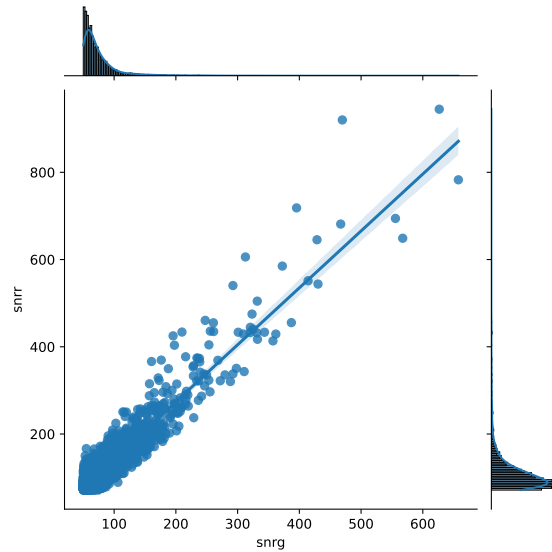


Fig. 2: Regresión lineal en la relación entre la señal-ruido en la banda verde y la señal-ruido en la banda roja, se observa sobre el eje x y sobre el eje y, el error de la interpolación, la relación entre las señales- ruido es $snrr = 1.3047869845619968 \cdot snrg + 13.136296670676828$ con un error estándar de ~ 0.007 , también toca mencionar que estas señales están en un rango de temperatura efectiva entre 4800K y 6800 K, además $snrg \geq 50$ y $snrr \geq 71.43$, siendo los puntos azules los espectros de snrg y snrr.

⁵Se puede ver desde la matriz de correlación hecha en el código pero por su extenso tamaño no se anexo

Dando la relación $snrr = 1.3047869845619968 \cdot snrg + 13.136296670676828$ con un error estándar de ~ 0.007

Se gráfica la gravedad superficial en escala logarítmica en función de la temperatura efectiva, para observar cuales son las estrellas analizadas en un diagrama similar al diagrama de Hertzsprung-Russell:

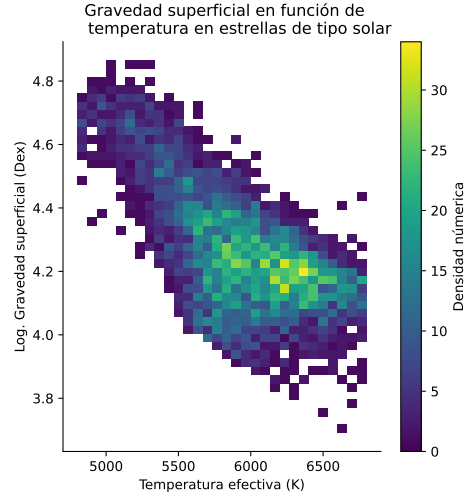


Fig. 3: Se observa la gravedad superficial en escala logarítmica en función de la temperatura efectiva, cada punto tiene un valor de densidad numérica de las estrellas en un sector del diagrama, dado los límites de la temperatura y en comparación de diagrama del diagrama Hertzsprung-Russell se observa que este fragmento es de estrellas en la secuencia principal, en un futuro serán gigantes rojas y mas adelante se podrían convertir en una enana blanca o una nebulosa planetaria

De manera ilustrativa se muestra también en la coordenadas celestes cada estrella observada.

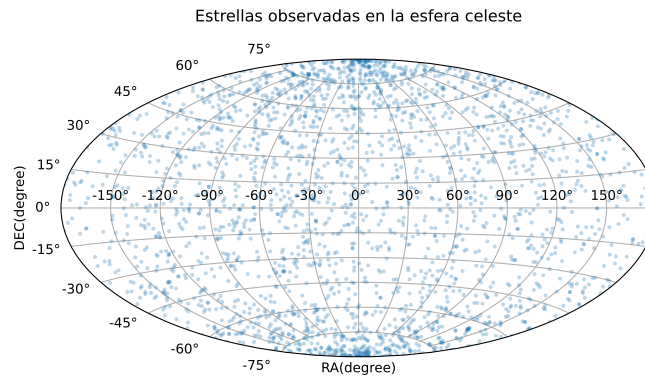


Fig. 4: Estrellas en la proyección de la esfera celeste, los ejes son en declinación (DEC) o latitud y Ascensión Recta (RA) o longitud.

Aunque el telescopio ya da la información de la velocidad en la línea de la visual (LOS) de cada astro de manera ilustrativa se muestra como es el proceso un poco complejo de obtener estos datos, por lo que se introduce un cubo de datos a partir de una simulación MHD de una mancha solar[2], donde se sintetiza la línea espectral $\lambda = 6032.49\text{\AA}$, para observar movimiento del flujo se sintetizo con un ángulo heliocéntrico de 60° , con tamaño de píxel de 48 Km, ahora los datos sintetizados nos da un cubo de datos el cuál tiene parámetros físicos de (longitud de onda, posición del píxel en x, posición del píxel en y, polarización de la luz en parámetros de Stokes), el usado en este desarrollo es de luz no polarizada o parámetro de Stokes I.

dado este análisis se obtiene la imagen del continuo:

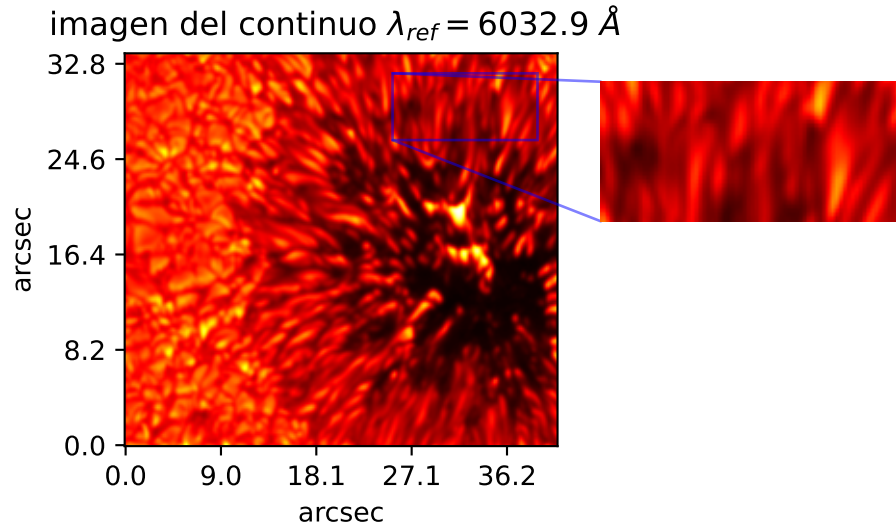


Fig. 5: Imagen de la simulación MHD tomada en el continuo alrededor de la línea sintética $\lambda_{ref} = 6032.49\text{\AA}$, se observa en el panel de la derecha un panel de la penumbra solar de la mancha solar.

Cada píxel presenta un perfil de absorción, para obtener la velocidad en la línea de la visual se usa el método de los bisectores[3], de modo sencillo se hace un corte en $\hat{I} = 0.5$, siendo \hat{I} la intensidad normalizada en el píxel, de manera mas didáctica, se muestra como se observa los parámetros físicos de un píxel situado en la penumbra.

Perfil de la línea espectral $\lambda_{ref} = 6302.49\text{\AA}$

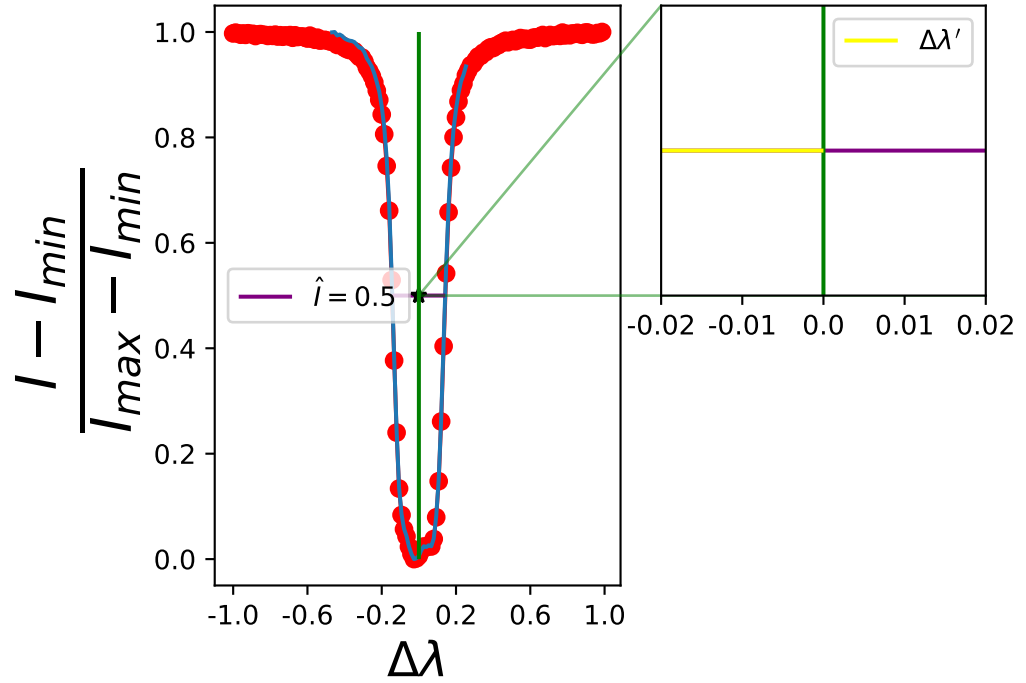


Fig. 6: perfil de absorción de Fe I normalizado. En el panel de la izquierda los puntos rojos muestran la línea de absorción tomada en un píxel de la penumbra. La línea azul es una interpolación por splines cuadrados. La línea verde representa la longitud de onda referencia. La línea morada presenta uno de los bisectores a una intensidad de $\hat{I} = 0.5$. El panel de la derecha es una ampliación en el sector de la distancia entre el punto medio del bisector y la longitud de onda de referencia, en donde la línea amarilla es el corrimiento Doppler en este píxel.

Usando la ecuación doppler no relativista, se encuentra la velocidad en ese píxel.

Si se hace para todos los píxeles, se puede encontrar la V_{LOS} de la fotosfera en un bisector $\hat{I} = 0.7$, y consiguiendo un dopplergrama, teniendo:

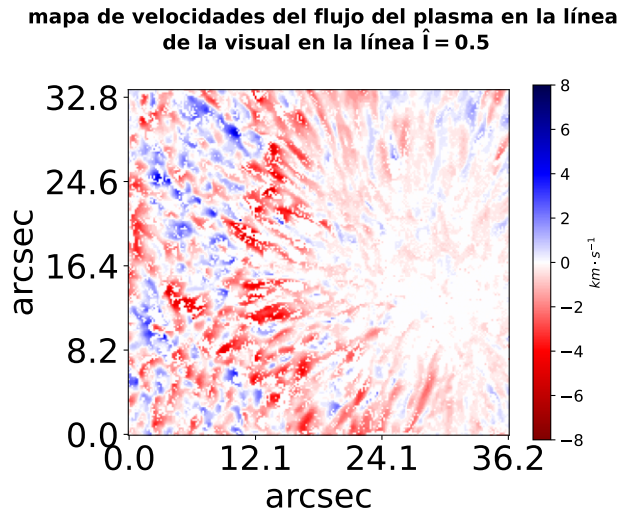


Fig. 7: Dopplergrama en la línea de la visual en el bisector o línea referencia $\hat{I} = 0.5$, se observa que el flujo penumbral tiene un corrimiento hacia el rojo, lo cual se puede concluir por el comportamiento del flujo Evershed, que la mancha esta al este del hemisferio solar [4, 5]

Ya en este punto se puede conseguir bastantes cosas, por ejemplo, suponiendo una simetría axial por medio de coordenadas esféricas se puede obtener la velocidad radial dado el plano tangente de la mancha y la velocidad azimuthal, sale del ejemplo que queremos mostrar.

Ya mostrado el ejemplo pequeño en comparación de los que hace el telescopio que es para BASTANTES estrellas, se continuara analizando el conjunto de datos cromosfericos.

Se hace un análisis de como es la metalicidad en función de la gravedad superficial para hallar algunas relaciones cualitativas.

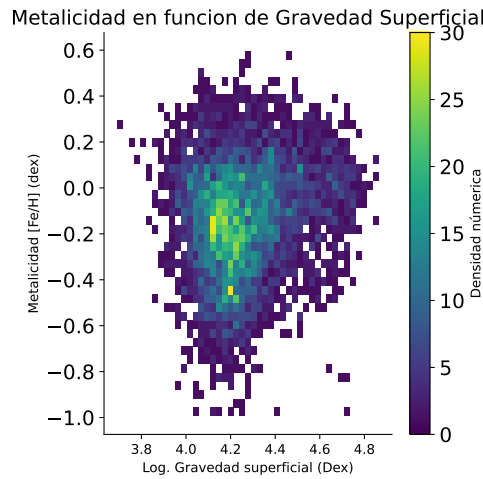


Fig. 8: Metalicidad en función de la gravedad superficial en escala logarítmica, cada punto tiene una distinción de la densidad numérica de las estrellas, la concentración para estrellas de tipo solar es que mayormente no hay tanto metales presentes, y es normal porque la temperatura y la edad de las estrellas no son suficientes para producir hierro o silicio, los pocos metales presentes deben ser a estrellas que estuvieron presentes en una nebulosa planetaria o de cualquier tipo, que las enriquecio en metales, las cuales son reducidas.

Dado que es un conjunto de datos reducido podemos hacer un análisis bivariado para estimar el comportamiento del resto de las estrellas, visualmente sería:

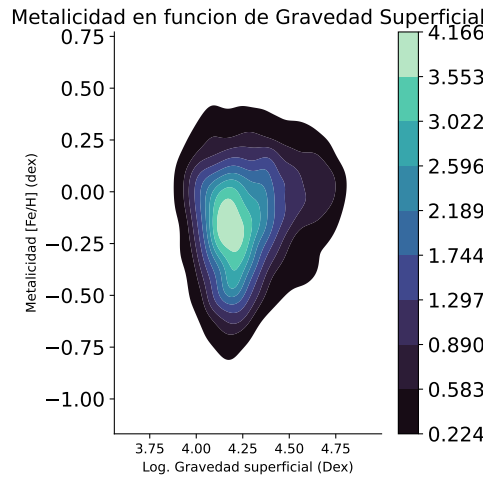


Fig. 9: Imagen de análisis bivariado de la metalicidad en función de la gravedad superficial en escala logarítmica, se observa como es el comportamiento entre las dos variables para intuir de alguna manera como sería el comportamiento de las demás estrellas tipo solar del resto del dataset, se observa una gran concentración de estrellas con metalicidad inexistente, dado que tiende a valores negativos, en proporción hay mucho mas hidrógeno que hierro en las estrellas de tipo solar.

Ahora mirando el asombroso trabajo que hace el telescopio, analizando cada velocidad radial en función de su gravedad superficial se obtiene:

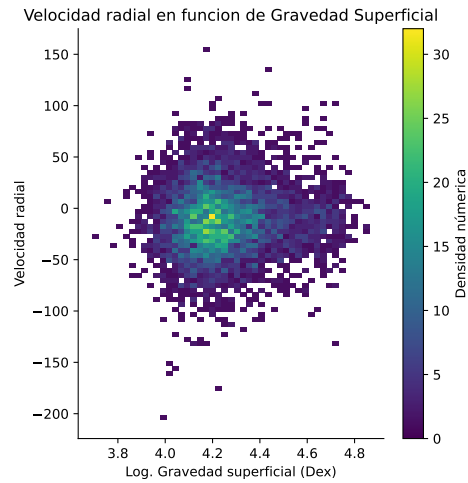


Fig. 10: Velocidad radial V_{LOS} en función de la gravedad superficial en escala logarítmica, cada punto tiene una distinción de la densidad numérica de las estrellas, la concentración para estrellas de tipo solar es que mayormente hay demasiado movimiento, y es normal porque la lejanía de las etrellas y posiblemente muchas sean estrellas binarias, por lo que su velocidad radial observada desde el telescopio es demasiado alta.

De igual forma para estimar las demas estrellas que no tomamos en cuenta por el tamaño del dataset, usamos una estimación bivariada.

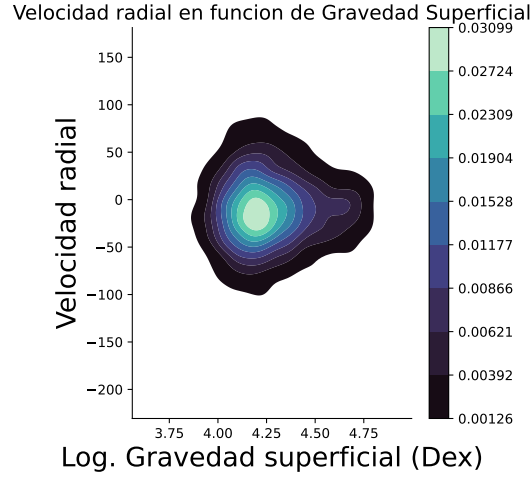


Fig. 11: Imagen de análisis bivariado de la Velocidad radial V_{LOS} en función de la gravedad superficial en escala logarítmica, se observa como es el comportamiento entre las dos variables para intuir de alguna manera como sería el comportamiento de las demás estrellas tipo solar del resto del dataset, se observa una gran concentración de estrellas con una velocidad relativamente alta, posiblemente por la gran cantidad de estrellas binarias en el universo, además el movimiento se incrementa dado que las estrellas están vibrando (nodos p), por lo que el perfil radial dado el espectro, se desplaza de la longitud de onda de emisión en este caso el del calcio ionizado

1 Testeo de hipótesis

1. Como las líneas de flujo de Ca II tienen relación directa con la actividad cromosferica, y este flujo fue medido para dos "bandpass" 1.09A FWHM triangular y 1A rectangular para la obtención de los índices de actividad cromosferica S_{tri} y S_{rec} respectivamente. Nos lleva a pensar, ¿Sera que hay alguna diferencia entre los índices de actividad cromosfericas?

Como $p_{value} = 0.0002716358536594022$ es menor que $\alpha = 0.05$ Se rechaza la hipótesis nula, por lo que las medidas de actividad cromosferica si difieren de la banda ancha de emisión utilizada. Este resultado es un poco controversial, dado que en la gráfica de las densidades de estos índices, pareciera que no difieren mucho.

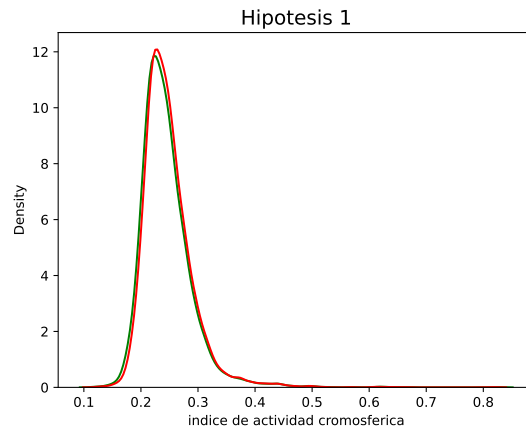


Fig. 12: Densidades de probabilidad de los índices de actividad cromosferica.

2. Con respecto al resultado controversial anterior, nos lleva a pensar si los valores de los flujos obtenidos para los distintos "bandpass" de 1.09Å FWHM triangular y 1Å rectangular ($H_{mean_{tri}}$ y $H_{mean_{rec}}$) difieren.

Al hacer el t-test, se evidencia que el $p_{value} = 0.7656830657043792$ es mayor que $\alpha = 0.05$. No se puede rechazar la hipótesis nula, esto es, las medias de ambos datos son las mismas, además que son el apoyo del Levene test, se concluye que sus varianzas también lo son, por lo que los datos de estos flujos no varían como se puede apreciar en las siguientes figuras. La primera es de los datos ajustados a la distribución normal mediante la función fit().

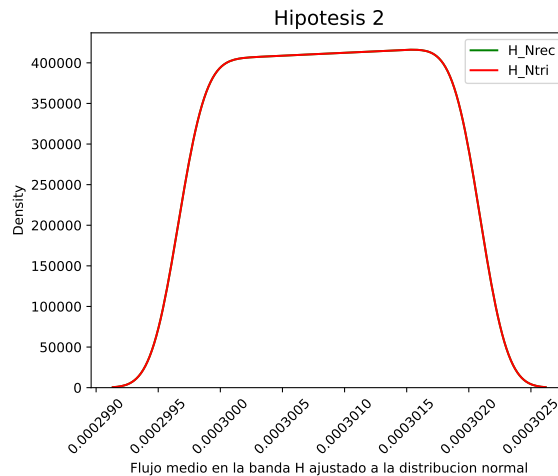


Fig. 13: Datos ajustados a la distribución normal en la banda H

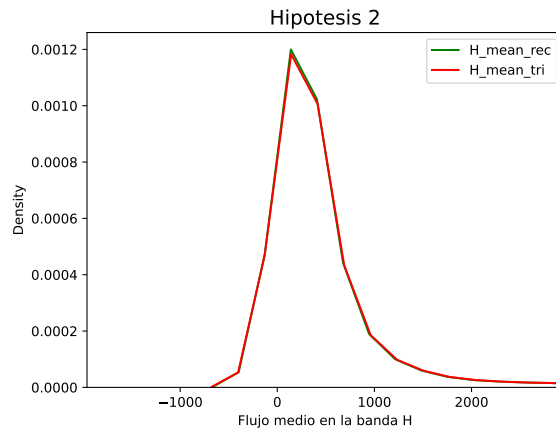


Fig. 14: Distribución de los flujos en la banda H

3. Como los valores del flujo en la banda H no varían, solo queda verificar si los flujos en la banda K varían, para tratar de dar respuesta al resultado anómalo de la hipótesis 1.

como $p_{value} = 0.7952538933888289$ es mayor que $\alpha = 0.05$ No se puede rechazar la hipótesis nula, esto es, las medias de ambos datos son las mismas. Este resultado nos deja con el sin sabor de no haber encontrado por que el t-test en la hipótesis aparentemente fallo; Se reafirman los resultados con las siguientes imágenes.

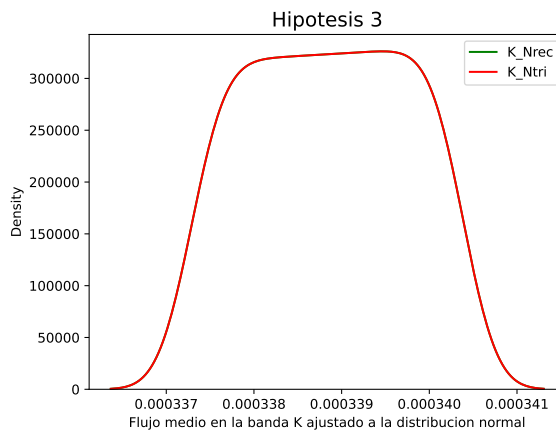


Fig. 15: Datos ajustados a la distribución normal en la banda K

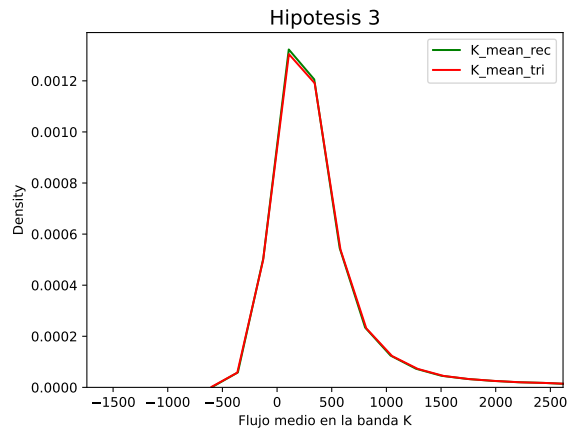


Fig. 16: Distribución de los flujos en la banda K

4. Podemos también analizar que tan influida se ve la velocidad radial con respecto a la temperatura efectiva de la estrella. Para lo cual se dividió el intervalo de temperatura efectiva en dos intervalos $I_1 = 4802.73 < teff < 5939.97$ e $I_2 = 5940.44 < teff < 6799.83$. A partir de las velocidades radiales asociadas para estos valores, se aplico el t-test para ver si diferían entre ellas.

Tanto el levene test, como el t-test rechazan la hipótesis nula, lo que nos indica que las velocidades radiales en estos rangos difieren. Aunque con ayuda de las imagenes de sus distribuciones, se puede apreciar que la diferencia no es "tanta".

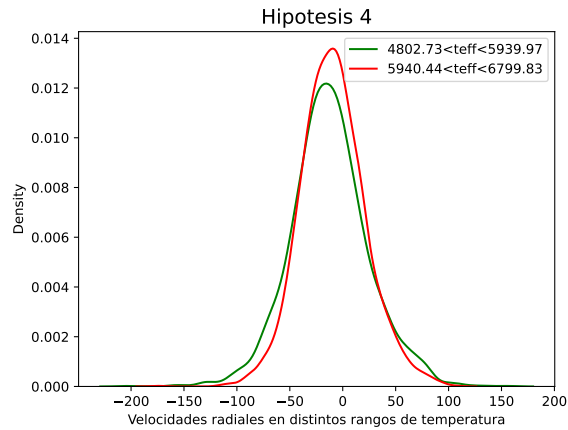


Fig. 17: Velocidades radiales para dos rangos de temperaturas

5. El uso apropiado de ANOVA en el análisis de los datos de la encuesta requiere que se cumplan algunos supuestos, incluyendo la distribución normal de la información, la independencia de los casos y la igualdad de varianza. Este método se utiliza para comparar dos medias de dos o mas grupos independientes (no relacionados) utilizando

la distribución F. La hipótesis nula para la prueba es que las dos medias sean iguales. Por lo tanto, un resultado significativo es que las dos medias sean desiguales. En este orden de ideas, era necesario disponer de grupos independientes con un significado semejante, para poder llevar a cabo la comparación, por lo que se escogieron los flujos de Ca II en las bandas H,K(previamente evaluados),R y V. Sin embargo, estos grupos difieren mucho entre si, aseveración que es respalda con el levene test, al proporcionar que alguna de las desviaciones estándar del conjunto de 4 grupos, es distinta. De esta manera, con el resultado anterior y el hecho que las distribuciones de los grupos no son normales, solo se cumpliría el supuesto de la independencia entre lo grupos, por lo que se es imposible realizar el test ANOVA. A pesar de todo, se realizo el test con el estadístico f y se pudo corroborar mediante el valor tan desfasado que dio($p_{value} = 4.8135672022947705 * 10^{-178}$, lo previamente dicho.

Referencias

- [1] W. Zhang, J. Zhang, H. He, Z. Song, A. Luo, H. Zhang, Stellar Chromospheric Activity Database of Solar-like Stars Based on the LAMOST Low-Resolution Spectroscopic Survey. arXiv e-prints arXiv:2209.15255 (2022). arXiv:2209.15255 [astro-ph.SR]
- [2] M. Rempel, Numerical Simulations of Sunspot Decay: On the Penumbra-Evershed Flow-Moat Flow Connection. **814**(2), 125 (2015). <https://doi.org/10.1088/0004-637X/814/2/125>
- [3] R.T. Ishikawa, Y. Katsukawa, T. Oba, M. Nakata, K. Nagaoka, T. Kobayashi, Study of the Dynamics of Convective Turbulence in the Solar Granulation by Spectral Line Broadening and Asymmetry. **890**(2), 138 (2020). <https://doi.org/10.3847/1538-4357/ab6bce>
- [4] J.S. Castellanos Durán, Strong magnetic fields and unusual flows in sunspots. Ph.D. thesis, University of Göttingen (2022)
- [5] J.S. Castellanos Durán, A. Lagg, S.K. Solanki, How rare are counter Evershed flows? **651**, L1 (2021). <https://doi.org/10.1051/0004-6361/202141159>