

Métodos numéricos para la ciencia e Ingeniería:

Informe Tarea 9

José Guillermo Araya

24 de noviembre de 2015

1. Pregunta 1 y 2

1.1. Introducción

El objetivo de esta pregunta es encontrar la constante de Hubble H_0 usando los datos de el archivo "data/hubble_original.dat" luego de "data/SNIa.dat" su intervalo de confianza para 95 %, donde:

$$v = H_0 * D$$

1.2. Procedimiento

En primer lugar para encontrar H_0 , como el problema no es simétrico debemos ajustar una recta para $v = H_0 * D$ y otra para $D = v/H_0$, y luego tomar un promedio de ambas para obtener una alternativa simétrica.

Para ajustar las rectas se ocupa la minimización de chi cuadrado de scipy, implementada en `curve_fit`, como se mencionó anteriormente se obtienen 2 valores para H , de modo que:

$$H_0 = \frac{H_1 + H_2}{2}$$

Para encontrar el intervalo de confianza, dado que no tenemos información sobre los errores de medición, utilizamos una simulación de bootstrap, donde generamos $N_{bootstrap}$ sets de mediciones sintéticas a partir de los datos, con cada una de estas muestras repetimos el proceso para estimar H_0 y guardamos ese valor. Finalmente ordenamos el arreglo que contiene todos los H_0 , y para tomar una muestra de 95 % simétrica, elegimos desde el índice $N_{boot} * 0,025$, hasta el índice $N_{boot} * 0,975$, donde los extremos son los límites del intervalo de confianza.

El procedimiento es análogo para la pregunta 2 solo cambiando el archivo de donde se importan los datos.

1.3. Resultados

Usando los datos de "data/hubble_original.dat" se obtiene:

$$H_0 = 472,1406173$$

y el siguiente intervalo de confianza:

Intervalo de confianza para 95 % = [364,92847445 : 600,40188199]

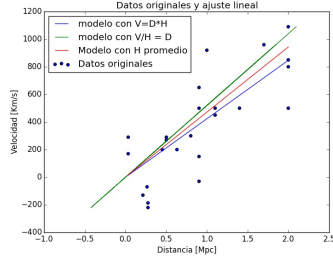


Figura 1: Ajuste lineal a los datos de "data/hubble_original.dat"

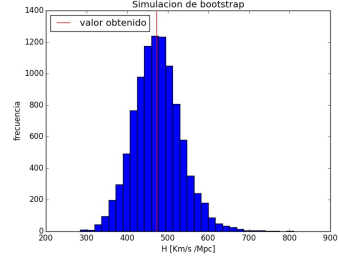


Figura 2: Histograma de los valores sintéticos de H_0

Para la pregunta 2, es decir, usando los datos de "data/SNIa.dat", se obtiene:

$$H_0 = 70,84131213$$

Intervalo de confianza para 95 % = [67,991450161 : 74,4967658056]

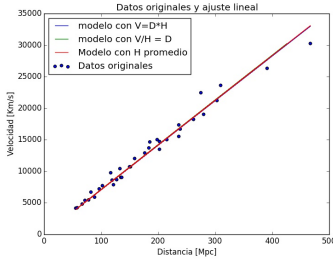


Figura 3: Ajuste lineal a los datos de "data/SNIa.dat"

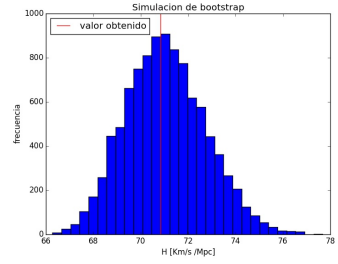


Figura 4: Histograma de los valores sintéticos de H_0

1.4. Conclusiones

Se ve claramente que el ajuste lineal con los datos originales dista mucho del ajuste con datos actualizados, aunque ambos predicen que las estrellas más lejanas se mueven con mayor velocidad, el primero está incorrecto por un factor 10 que es considerablemente grande.

Respecto a el ajuste lineal en sí, a simple vista concluimos que el algoritmo está funcionando correctamente pues la línea pareciera aproximar bien los datos en ambos casos, y el valor encontrado está dentro de los intervalos de confianza mencionados, cabe notar que el intervalo de confianza para el segundo ajuste es mucho menor pues es una mucho mejor aproximación.

2. Pregunta 3

2.1. Introducción

El objetivo de esta pregunta es encontrar la línea recta que mejor ajusta el flujo de la banda i con el flujo de la banda z, usando los datos del catálogo de

cuasares del Data Release 9 del Sloan Digital Sky Survey (SDSS). Además de indicar el intervalo de confianza para el 95 % de los parámetros de la recta

2.2. Procedimiento

Para encontrar la línea recta que mejor aproxima la relación entre ambos flujos se utilizará la función "polyfit" de numpy, que retorna el coeficiente de posición y la pendiente de la recta.

En este caso tenemos información sobre los errores de medición asociados, por lo que podemos utilizar una simulación de montecarlo para obtener los intervalos de confianza, para esto creamos N_{mc} sets de datos a partir de los datos del catálogo y asumiendo que el error es gaussiano la muestra queda:

$$muestra_i = datos_i + error * r$$

donde *error*, es la información que tenemos sobre el error, y *r* es una variable aleatoria Gaussiana. Con estas muestras se realiza "polyfit" nuevamente obteniendo un arreglo de parámetros para la curva, y al igual que en las preguntas anteriores, estos arreglos se ordenan y se toman los extremos de índice $N_{mc} * 0,025$ y $N_{mc} * 0,975$ para encontrar el intervalo de confianza para el 95 %

2.3. Resultados

De el procedimiento mencionado anteriormente se obtiene:

recta : $1,10255388035 \times x + 3,14925730948$

El intervalo de confianza al 95 % para la pendiente es:

$$[0,94925942032 : 1,14279614126]$$

El intervalo de confianza al 95 % para el coef de posición es:

$$[2,25622648987 : 7,76884537235]$$

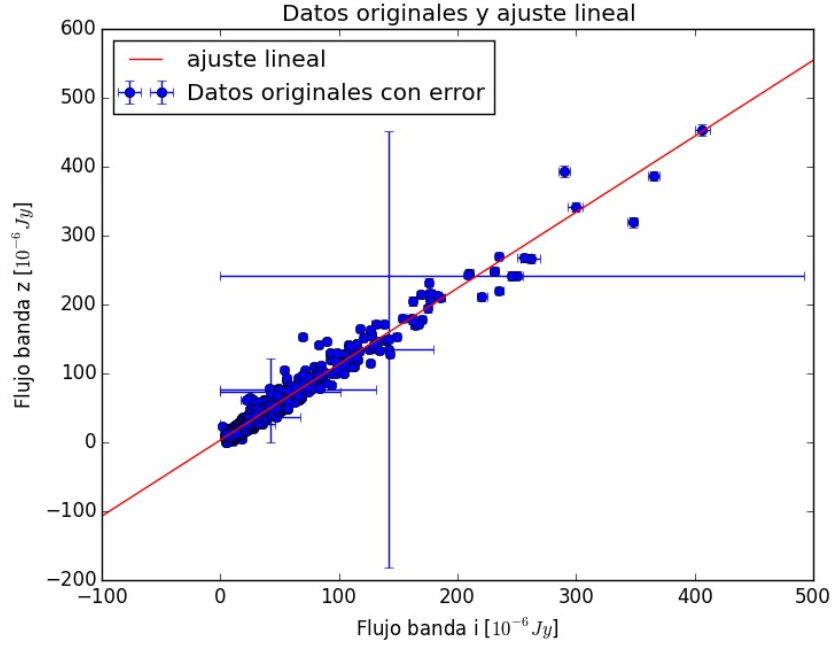


Figura 5: Ajuste lineal para los datos "de data/DR9Q.dat"

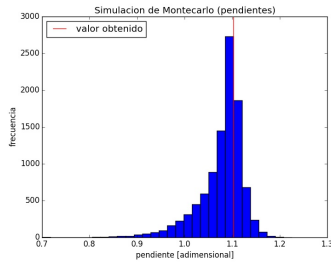


Figura 6: Simulación de montecarlo para el coeficiente lineal (pendiente)

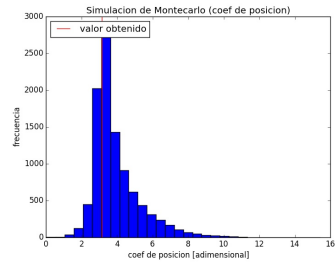


Figura 7: Simulación de monte-carlo para el coeficiente de posición

2.4. Conclusiones

En primer lugar, la recta obtenida tiene sentido pues aproxima los datos del catálogo correctamente. El hecho de poseer datos sobre los errores de medición nos permite realizar una simulación de montecarlo con nuevos datos, de manera que podemos crear sets de muestras muy distintas a diferencia de las preguntas anteriores donde no era posible hacer esto y se necesitaba hacer una simulación de bootstrap usando los mismos datos del archivo.

Ambos parámetros de la recta aproximada se encuentran dentro de el intervalo de confianza dado.

En otra oportunidad sería interesante analizar la relación entre los flujos de otras bandas y observar si todos siguen una relación similar a la encontrada para la banda i y la banda z.