Métodos Numéricos para la Ciencia e Ingeniería:

Informe Tarea n° 8

María Constanza Flores V.

25 de noviembre de 2015

1. Pregunta 1 y 2

1.1. Introducción

En esta pregunta se busca poder encontrar la constante de Huble H_0, dada por el modelo propuesto por Edwin Hubble, en donde se comparan las velocidades de recesión de las galaxias lejanas con las distancias entre dichas galaxias y la Tierra. El modelo está dado por:

$$v = H_0 * D \tag{1}$$

Donde H_0 es la *constante de Hubble* y se expresa generalmente en unidades de km/s/Mpc, v es la velocidad de recesión en km/s y D la distancia en kilómetros.

Se entregan dos archivos con datos, por un lado el archivo `hubble_original.dat`, que contiene las mediciones originales que Hubble utilizó en este trabajo; y por otro lado el archivo `SNIa.dat`, el cual contiene aquellos datos con los que se pueden obtener una estimación más reciente de la constante de Hubble (Freedman et al. 2000).

1.2. Procedimiento

Para poder estimar la constante de Hubble se utilizan los siguientes modelos:

- Modelo 1: $v = H_0 * D$
- Modelo 2: $D = v/H_0$

Debido a que no existe justificación alguna para elegir un modelo por sobre el otro (no se sabe si en realidad la velocidad está en función de la distancia o viceversa), se decide utilizar la bisección entre ambos como modelo final. Para dos rectas ($y = a_1 + b_1 * x$), ($y = a_2 + b_2 * x$), se tiene que la pendiente de la bisección está dada por:

$$b_{biseccion} = \frac{b_1 b_2 - 1 + \sqrt{(1 + b_1^2)(1 + b_2^2)}}{b_1 + b_2} \tag{2}$$

Se utiliza la funcion `curve_fit` de la librería `scipy.optimize` para poder minimizar el valor de χ^2, y así poder obtener el valor de H_0 óptimo para los modelos 1 y 2, y luego se calcula el H_0 dado por la bisección y descrito en la fórmula (2). También se utiliza una simulación de Bootstrap para calcular un intervalo de confianza al 95 %, del parámetro de H_0 calculado por la bisección.
Finalmente se procede a graficar tanto los datos, los ajustes lineales de cada uno de los modelos y un histograma en que se muestra lo realizado con la simulación de Bootstrap.

1.3. Resultados

1.3.1. Pregunta 1

Los valores de H_0 calculados para cada modelo, y el intervalo de confianza calculado con la simulación de Bootstrap están dados por:

- H_0 modelo 1 = 423.937322595 [km/s/Mpc]
- H_0 modelo 2 = 520.343909896 [km/s/Mpc]
- H_0 modelo bisección = 467.219313342 [km/s/Mpc]
- El intervalo de confianza al 95 % es: [385.805772258, 548.228046036]

Los gráficos obtenidos se muestran en la figura a continuación:

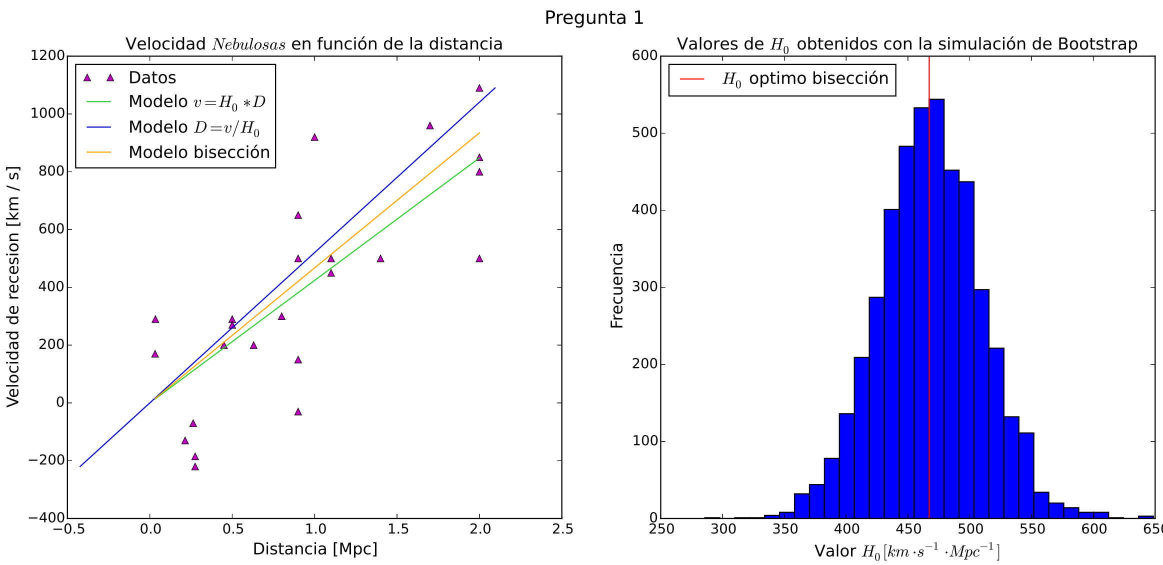

Figura 1: Resultados obtenidos de la pregunta 1

1.3.2. Pregunta 2

Los valores de H_0 calculados para cada modelo, y el intervalo de confianza calculado con la simulación de Bootstrap están dados por:

- H_0 modelo 1 = 70.6672159855 [km/s/Mpc]
- H_0 modelo 2 = 71.0154082682 [km/s/Mpc]
- H_0 modelo bisección = 70.8408843619 [km/s/Mpc]
- El intervalo de confianza al 95 % es: [68.6812517539, 73.562729413]

Los gráficos obtenidos se muestran en la figura a continuación:

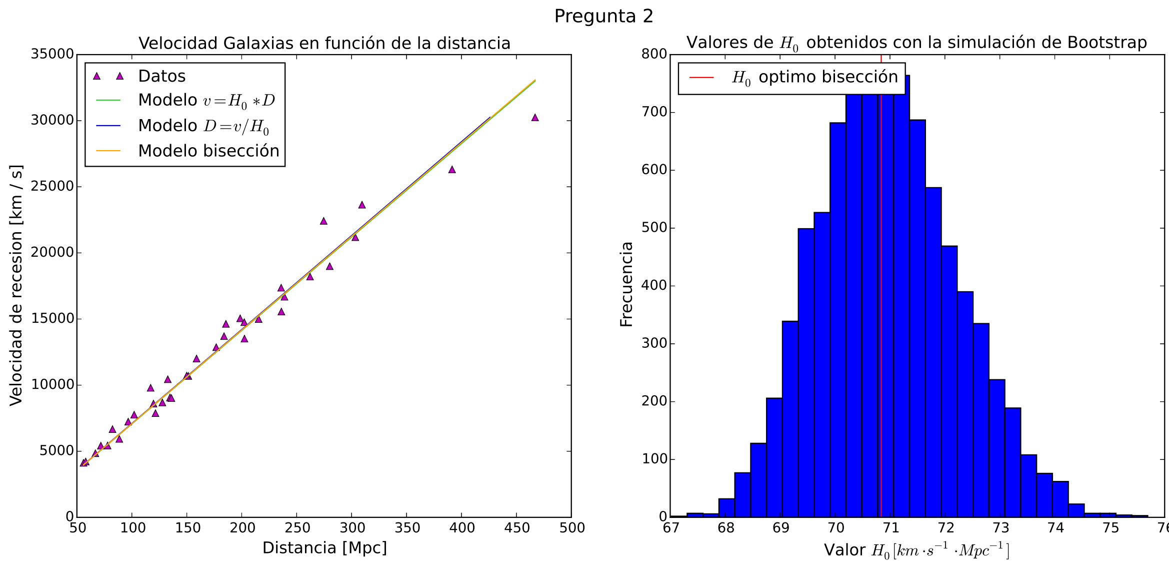

Figura 2: Histograma

1.4. Conclusiones

Los valores de H_0 obtenidos en la primera y en la segunda pregunta fueron valores bastante distantes. Es posible notar que para la pregunta 1 el modelo no se ajusta considerablemente a los datos (Figura 1), y eso puede deberse a lo dispersos que éstos se encuentran, pues no existe una recta clara que los pueda unir, y a la baja cantidad de datos que se tienen.

Por otro lado para la segunda pregunta, no se puede notar la diferencia entre el modelo 1 y 2, y el modelo de la bisección (Figura 2).

Otra forma de evidenciar la mala calidad de los datos de la pregunta 1, es al ver su intervalo de confianza al 95 %, donde dicho intervalo tiene una diferencia de 190 [km/s/Mpc], en cambio el intervalo de confianza de la pregunta dos tiene un ancho de 5 [km/s/Mpc].

Por otra parte se puede notar que los histogramas obtenidos al graficar el set de datos obtenido por la simulación de Bootstrap son similares a una una distribución gaussiana: esto es debido a que, al no conocer los valores de los errores de los datos utilizados en esta pregunta, se asume que ellos tienen una distribución gaussiana.

2. Pregunta 3

2.1. Introducción

Se desea encontrar una línea recta que mejor modele la relación entre el flujo de banda i y el flujo de banda z, ambos datos dados en el archivo `DR9Q.dat`, el cual es una sección recortada del catálogo de cuasares del *Data Release 9* del *Sloan Digital Sky Survey.* También se pide encontrar los intervalos de confianza al 95 % para los parámetros de dicha línea recta, es decir, la pendiente y el coeficiente de posición.

2.2. Procedimiento

Para poder calcular los valores de la pendiente y el coeficiente de posición, se utiliza la rutina `numpy.polyfit`, la cual entrega un ajuste polinomial (de primer grado para este caso) a un par de set de datos.

Por otro lado para calcular el intervalo de confianza para cada parámetro (pendiente y coeficiente de posición), se realiza una simulación de Montecarlo, en la cual, a grandes razgos, se generan N_{mc} sets de datos sintéticos, que se estiman a partir de los datos guardados y sus errores, de la forma:

$$muestra_i = flujo\,banda_i + error_i * r$$

$$muestra_z = flujo\,banda_z + error_z * r$$

Que corresponde a una distribución normal con los valores de los flujos de bandas y sus errores asociados. Y a partir de cada set se calculan valores de la pendiente y coeficiente de posición con la rutina de polifyt mencionada anteriormente. Los N_{mc} valores de la pendiente y de los coeficientes de posición se ordenan y se calculan sus percentiles de 2.5 % y 9.75 % .

Finalmente se procede a graficar los datos y el ajuste lineal obtenido.

2.3. Resultados

El intervalo de confianza al 95 % para la pendiente es: [0.861140759214, 1.04129702615].

El intervalo de confianza al 95 % para el coeficiente de posicion es: [-0.917891360923, 4.19039496042].

El valor obtenido para la pendiente al hacer un ajuste lineal corresponde a: 1.10255388035, y el valor del coeficiente de posición es: 3.14925730948

El gráfico obtenido se muestra en la figura a continuación:

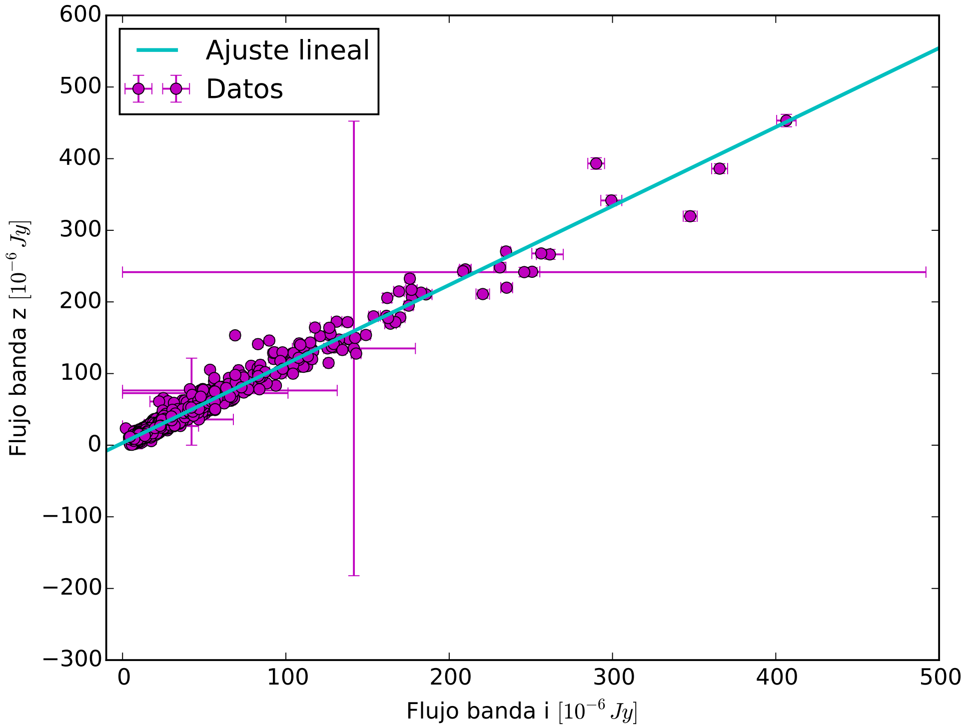

Figura 3: Ajuste lineal para los flujos de banda i, z

2.4. Conclusiones

El ajuste lineal obtenido satisface lo pedido, lo que es evidenciable en la Figura 3, donde se puede notar que la linea celeste, que representa la recta generada con los valores obtenidos con `polyfit`, se adecúa de buena forma a la tendencia de los datos. Y los intervalos de confianza obtenidos para la pendiente y el coeficiente de posición, tienen anchos lo suficientemente pequeños como para determinar que la muestra realizada es buena.