UNIVERSIDAD DE CHILE, FACULTAD DE CIENCIAS FISICAS Y MATEMATICAS
Metodos Numericos para la Ciencia y Ingenieria

Tarea 9: Constante de Hubble; estimación e intervalos de confianza

Maria Jose Hernández Pozo

November 25, 2015

1 Pregunta 1

1.1 Introducción

Se proporciona el archivo ubicado en el repositorio: data/hubble_original.dat que contiene las mediciones originales que Edwin Hubble utilizó en 1929, comparando la velocidad de recesión de galaxias lejanas con las distancias entre estas y la Tierra. Su modelo es:

$$v = H_0 * D$$

Donde H_0 es la constante de Hubble y generalmente se expresa en unidades de km / s / Mpc (Mpc: Mega parsec, 3.086e24 cm). Utilice los datos originales para derivar la constante de Hubble incluyendo su intervalo de confianza al 95%.

1.2 Procedimiento

Lo primero a observar es que el resultado de modelar $v = HD$ (modelo 1); ó $D = v/H$ (modelo 2) es distinto y dado que no hay motivo para preferir uno por sobre el otro, se busca una alternativa simétrica. Para realizar esto, se modelarán ambas ecuaciones por separado y su respectivo resultado se promediará como solución simétrica. Es decir:

$$H_{0prom} = \frac{H_0^1 + H_0^2}{2}$$

Donde H_0^1 y H_0^2 representan el resultado de la constante buscada para el modelo 1 y 2 respectivamente.
Para determinar cada constante a partir de los datos se utilizó la minimización χ^2 presente en curve_fit de la librería scipy.optimize, la cual al entregarle la función a minimizar, las variables dependientes e independientes de la muestra, más una adivinanza nos retorna un ajuste de la constante.

Sin embargo no se poseen los errores ni tampcoo se posee información acerca de la distribución de estos, se opta entonces por una simulación de Bootstrap, la cuál desde los mismos datos, utilizando una selección aleatoria de estos, obtiene un promedio de aquella con un error asociado.
En este caso iteraremos 5000, y en cada iteración calcularemos mediante bootstrap una secuencia aleatoria de los datos muestreados, luego se hará uso de curve_fit para obtener una aproximación de H_0. Entonces con las distintas constantes se calcula un promedio y un intervalo de confianza en torno a este.

1.3 Resultados

Para los datos del repositorio data/hubble_original.dat, la constante de Hubble es:

$$H_{0promedio} = 474.440919491 \frac{km}{s\ Mpc}$$

Si reemplazamos este resultado en el primer modelo:

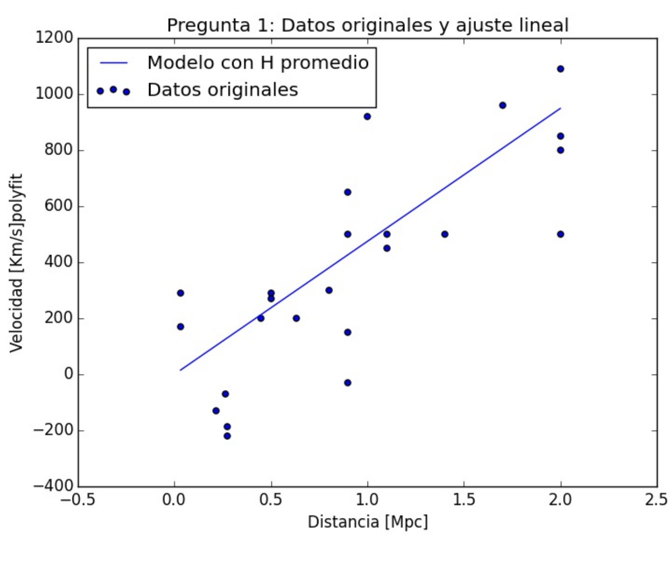

Figure 1: Ajuste usando la constante de Hubble promedio en comparación con los datos dispuestos de data/hubble_original.dat según el modelo 1

Se presenta también el intervalo de confianza de 95%: [397.044263136-558.804914415]
El cuál se representa en el siguiente histograma:

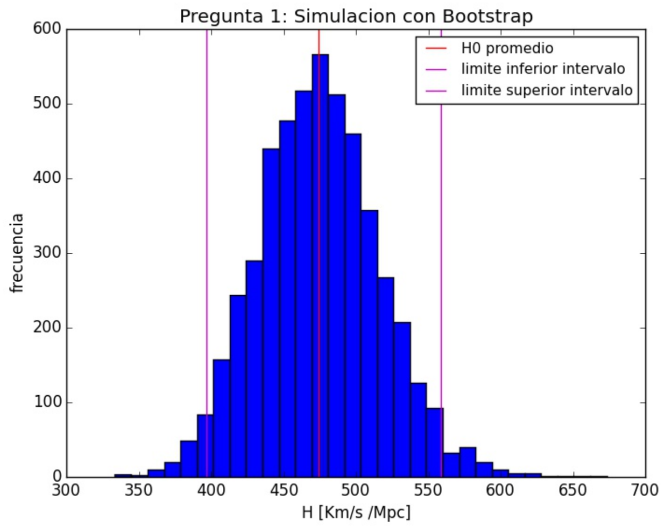

Figure 2: Histograma para los H0 promedio, su intervalo de confianza y el mejor valor

1.4 Conclusiones

Podemos observar que el ajuste lineal si bien se muestra equidistante de los puntos, presenta una gran dispersión. Cabe destacar que dado los distintos modelos trabajados como un promedio, quizás hubiese tenido un resultado más simétrico una relación logarítmica entre estos datos.
Respecto al histograma, se muestra simétrico y el valor encontrado esta dentro del intervalo de confianza, que es bastante amplio.

2 Pregunta 2

2.1 Introducción

Hubble, sin embargo, cometió un gran error en su estimación de H_0. Este deriva de utilizar una calibración equivocada de la relación período–luminosidad (entre otras cosas).
Una estimación más reciente de la constante de Hubble se obtiene con los datos ubicados en el archivo: data/SNIa.dat (Freedman et al. 2000) que utiliza Super Novas tipo I para estimar distancias para una muestra de galaxias. Entre otras

ventajas, el método permite estimar distancias muy superiores a las que se pueden medir con el método de las Cefeidas.
Vuelva a estimar la constante de Hubble para estos datos incluyendo su intervalo de confianza del 95%. Comente.

2.2 Procedimiento

Dado que esta pregunta es bastante similar a la primera solo diferenciandose en los datos entregados, se replican los mismos procedimientos que en la pregunta anterior.

2.3 Resultados

Para los datos del repositorio data/SNIa.dat, la constante de Hubble es:

$$H_{0promedio} = 0.0140992388498 \frac{km}{s\ Mpc}$$

Si reemplazamos este resultado en el primer modelo:

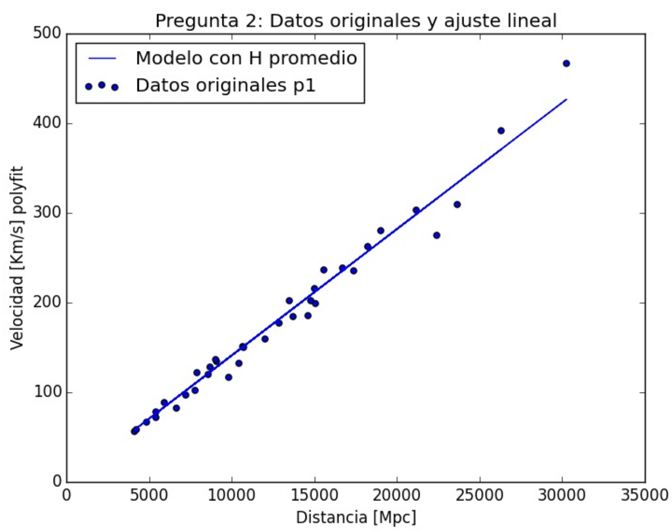

Figure 3: Ajuste usando la constante de Hubble promedio en comparación con los datos dispuestos de SNIa.dat según el modelo 1

Se presenta también el intervalo de confianza de 95%: [0.0135869453555-0.0145575178335]
El cuál se representa en el siguiente histograma:

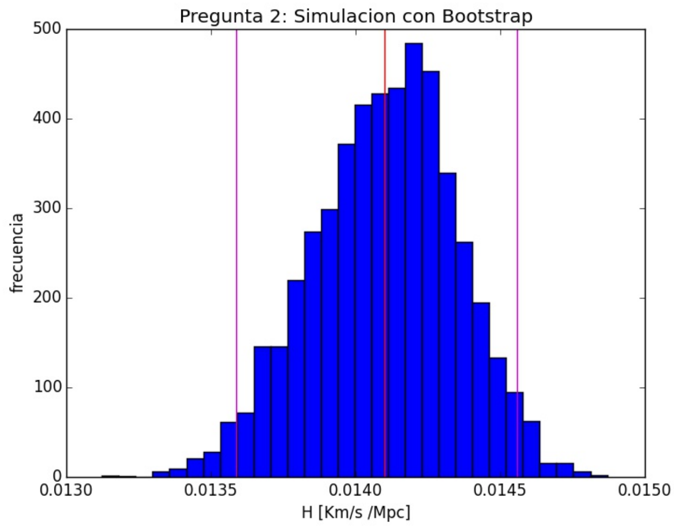

Figure 4: Histograma para los H0 promedio, su intervalo de confianza y el mejor valor

2.4 Conclusiones

El ajuste lineal presenta poca dispersión y buena aproximación. Cabe destacar, tal como en la primera pregunta, que dado los distintos modelos que se trabajaron como promedio, quizás hubiese tenido un resultado más simétrico una relación logarítmica entre estos datos.
Respecto al valor de la constante, dista mucho de la calculada anterior, si bien ambas son positivas, la primera en 10^4 la primera, lo que representa un error grave de estimación, el que deriva de utilizar una calibración equivocada de la relación período–luminosidad.
Respecto al histograma, se muestra menos simétrico que en la pregunta 1, el valor encontrado esta dentro del intervalo de confianza, que es menor respecto al encontrado anteriormente.

3 Pregunta 3

3.1 Introducción

El archivo data/DR9Q.dat es una sección recortada del catálogo de cuasares del Data Release 9 del Sloan Digital Sky Survey (SDSS). Encuentre la línea recta que mejor modela la relación entre el flujo en la banda i y la banda z, incluyendo los intervalos de confianza al 95% para los parámetros de la línea recta.

3.2 Procedimiento

Para determinar la pendiente y la constante a partir de los datos se utilizó la función np.polyfit de la librería numpy, a la cual basta entregarle las variables dependientes e independiente y grado del polinomio que nos retornará.

Esta vez tenemos errores asociados a ambas variables, se opta entonces por una simulación de MonteCarlo, la cuál supone una distribución normal de los errores, de esta forma itera 10000 y en cada iteración elabora una lista de las variables con su error ponderado por la distribución normal (datos + error * r, r número de una distribución normal con mu=0 y varianza=1), usando polyfit, obtenemos entonces una pendiente y una constante por cada iteración y podemos construir un histograma, ordenando los datos con np.sort(), podemos establecer el intervalo de confianza.

3.3 Resultados

La recta queda de la siguiente forma:

$$1.10255388035x + 3.14925730948$$

Donde el número que acompaña a x corresponde a la pendiente y el otro al coeficiente de posición.

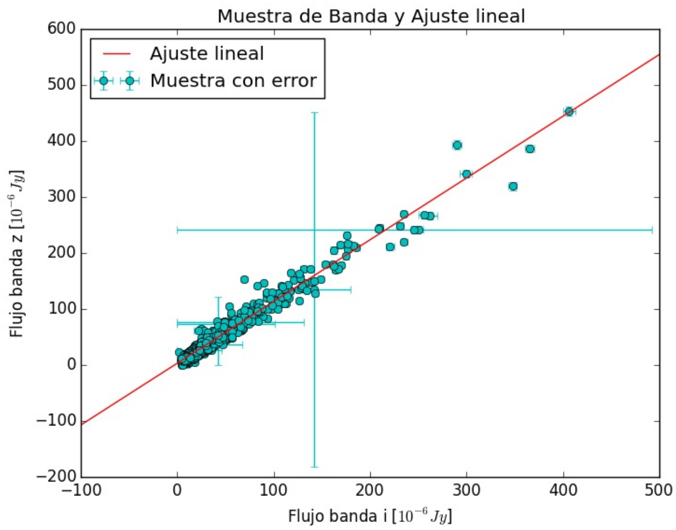

Figure 5: Datos de bandas z e i, con su aproximación lineal

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

$$[0.945737796117 - 1.14281661307]$$

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

$$[2.25491572198 - 7.929453357]$$

3.4 Conclusiones

El ajuste con polyfit muestra coherencia con los datos, tanto como la pendiente como el coeficiente de posición tienen intervalos pequeños lo que nos transmite que la aproximación es buena y que los datos tienen poca dispersión y error.