Métodos numéricos para la ciencia y la ingeniería
Tarea 9

Jean Paul Martel

24 de noviembre 2015

1 Determinación de la constante de Hubble

1.1 Introducción

- En 1929 Edwin Hubble estudio la relación existente entre la distancia de nebulosas y velocidad radial con respecto a la Tierra. En su modelo, la velocidad de recesión de una nebulosa (galaxia) distante es proporcional a su distancia de la Tierra. La relación es simple :

  $$v = H_0 * D$$

  Donde a H_0 se le llama "constante de Hubble" y representa la tasa de expansión del universo actualmente. Generalmente se escribe en unidades de km/s/Mpc.

  Para encontrar la expresión realizó un ajuste lineal a los datos obtenidos de la relación período–luminosidad [1] recientemente calibrada para estrellas de luminosidad variable tipo Cefeidas.

- Sin embargo, Hubble cometió errores considerables debido principalmente a una mala calibración de la relación período–luminosidad. Más recientemente se han hecho estimaciones mucho más precisas de la constante de Hubble. Un método más avanzado es usar Super Novas tipo I para estimar las distancias de las galaxias, pues permite estimar distancias muy superiores a las que se logran con el método de las Cefeidas. Por ejemplo, el año 2000 Freedman (et al.) utilizó este método logrando datos más precisos y en mayor cantidad que los obtenidos por Hubble.

- A continuación se busca:

  1. Estimar H_0 con intervalo de confianza al 95%, utilizando los datos originales de Hubble encontrados en el archivo *hubble_original.dat*.
  2. Estimar H_0 con intervalo de confianza al 95%, utilizando los datos Freedman, que se encuentran en el archivo *SNIa.dat*.

[1] Relación descubierta en 1912 por Henrietta Leavitt

1.2 Procedimiento

1.2.1 Ajuste lineal

- En general, un ajuste o regresión lineal es un método matemático que modela la relación entre una variable dependiente Y, las variables independientes X, suponiendo que esta relación sería aproximadamente lineal:

$$Y = a * X + b$$

  Dado los datos (x_i, y_i), se utiliza el criterio de los mínimos cuadrados para encontrar los mejores valores para a y b. Consiste en minimizar la suma de cuadrados:

$$S = \sum_{i=1}^{n} (y_i - ax_i - b)^2$$

- Para este problema particular quiere obtenerse la relación $v = H_0 * D$. Con esto se reconoce:
  $y_i = v_i$
  $x_i = D_i$
  $a = H_0$
  $b = 0$

  Es decir se minimiza: $S = \sum_{i=1}^{n} (v_i - H_0 D_i)^2$

  No es difícil encontrar que:

$$\frac{\partial S}{\partial H_0} = 0 \Rightarrow H_0 = \frac{\sum_{i=1}^{n} v_i D_i}{\sum_{i=1}^{n} D_i^2}$$

- Pero notar que el problema también puede modelarse como: $D = \frac{v}{H_0}$. La estimación para H_0 será distinta en este caso.

  Ahora debe minimizarse: $S = \sum_{i=1}^{n} (D_i - \frac{1}{H_0} v_i)^2$

  De donde se obtiene:—

$$\frac{\partial S}{\partial H_0} = 0 \Rightarrow H_0 = \frac{\sum_{i=1}^{n} v_i^2}{\sum_{i=1}^{n} v_i D_i}$$

- Por lo tanto, se utiliza una alternativa simétrica. Si se llama H_0^I a la constante obtenida con el primer modelo y H_0^{II} a la encontrada con el segundo modelo, se elije como estimación de la constante de Hubble el promedio:

$$H_0 = \frac{H_0^I + H_0^{II}}{2}$$

1.2.2 Intervalo de confianza para la estimación de H_0

- Debido a que no se disponen de errores de medición, se utiliza una simulación Bootstrap. A grandes rasgos este es uno de los llamados "métodos de remuestreo" que utilizan los datos observados para constituir un universo del cual extraer repetidas muestras. La idea es tomar una muestra tal que tenga el mismo tamaño de la original y en donde pueden repetirse elementos. Esto se repite N_{boot} veces, tal que cada vez se realiza la minimización de la suma de cuadrados y se encuentran nuevos valores para H_0. Estos valores son ordenados en un arreglo de menor a mayor. Al límite inferior del intervalo de confianza le corresponde 100(alpha), mientras que al límite superior 100(1-alpha/2). Para un intervalo de confianza del 95% se tiene $alpha = 0.05$

1.3 Resultados

- Utilizando los datos originales de Hubble se encuentra:

  $H_0 = 472.14$

  Intervalo de Confianza: $[394.39, 559.50]$

- Usando los datos de Freedman:

  $H_0 = 70.84$

  Intervalo de Confianza: $[68.66, 73.57]$

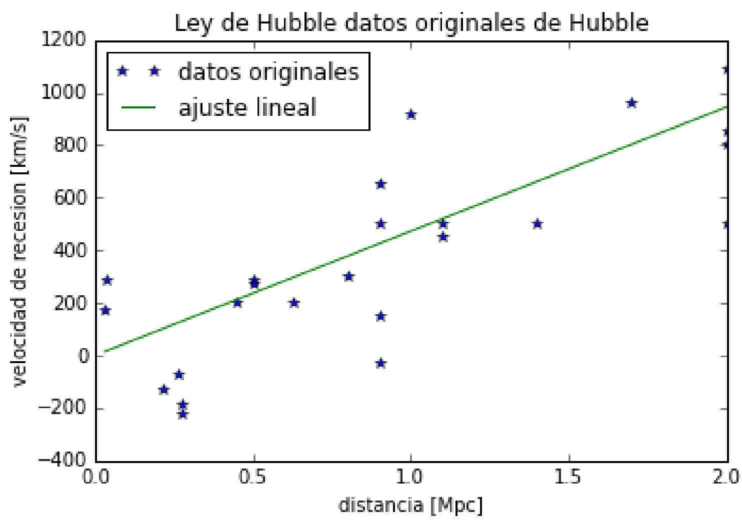

Figure 1: Gráfico velocidad de recesión versus distancia. Se muestran los datos originales de Hubble obtenidos de relación periodo-luminosidad, y la linea recta que se ajusta a estos datos

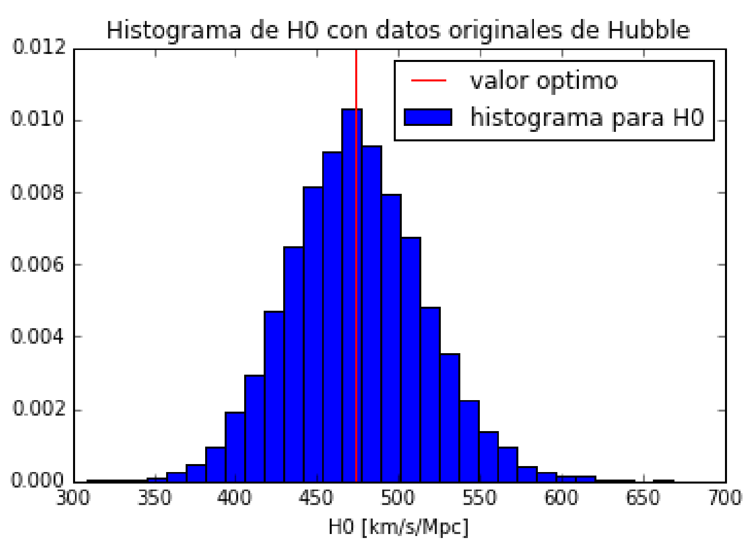

Figure 2:

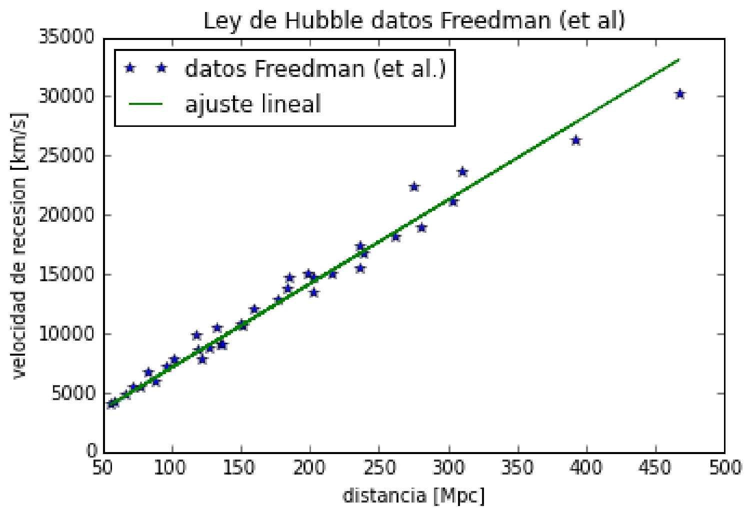

Figure 3: Gráfico velocidad de recesión versus distancia. Se muestran los datos de Freedman usando Super Novas tipo , y la linea recta que se ajusta a estos datos

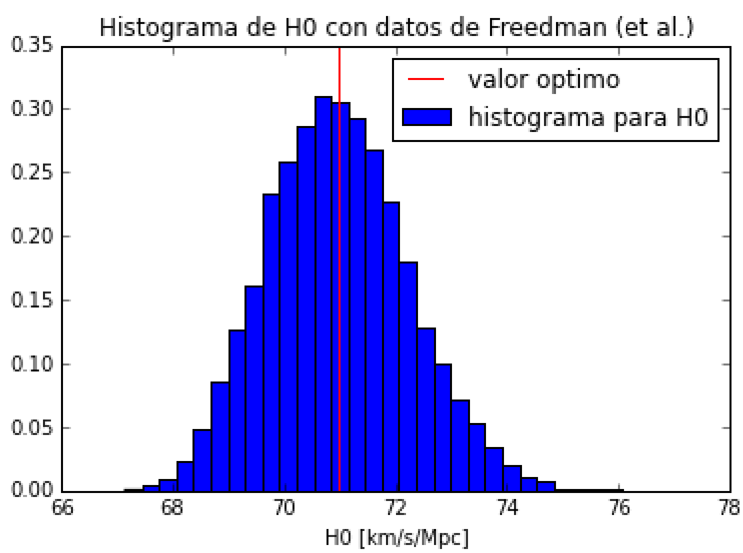

Figure 4:

1.4 Conclusiones

- Por una parte, los valores obtenidos con el ajuste lineal parecen ser correctos tanto para los datos de Hubble como para los de Freedman, pues se tienen intervalos de confianza acotados que aseguran una buena precisión de los valores. Aunque el rango del intervalo es bastante mayor para los datos de Hubble. Esto es esperable ya que se utilizó una muestra pequeña y existían grandes errores de calibración y medición en esa época. El valor estimado con datos recientes para la constante de Hubble es de $\sim 70\ km/s/Mpc$, que coincide con el actualmente aceptado para la tasa de expansión del universo.

- Por otra parte, comparando el valor de H_0 encontrado por Hubble con los encontrados más recientemente, es claro que Hubble sobre-estimó ampliamente esta constante. Como ya se dijo, eso fue principalmente por la mala calibración del método con el cual se medían las distancias. Interesante notar que $\frac{1}{H_0}$ es una medida aproximada del tiempo de universo, y con un valor de $H_0 \sim 472 km/s/Mpc$

resulta un tiempo muy corto para la edad de universo que se contradice con los datos de las edades de las estrellas más antiguas.

2 Relación entre el flujo en la banda i y la banda z de cuásares

2.1 Introducción

- Entre los datos en el catálogo de cuásares del Data Release 9 del SDSS (Sloan Digital Sky Survey) se encuentran valores de flujos en la banda i y la banda z, con sus respectivos errores.

- Se quiere utilizar nuevamente regresión lineal para encontrar una recta que se ajuste los datos, y ademas obtener intervalos de confianza del 95% para los parámetros de tal recta.

2.2 Procedimiento

- Primero, las unidades de los flujos y los errores son convertidas de nmaggies a Jy usando que:

$$1\ nmaggie \sim 3.631 * 10^{-6} Jy$$

- Siguiendo lo ya explicado en el problema anterior, una ajuste lineal para la relación entre las bandas $Banda_i$ y $Banda_z$ es:

$$Banda_z = a * Banda_i + b$$

  Donde a y b son las constantes a determinar. Para encontrarlas se utiliza $numpy.polyfit$ de python.

- Para construir un intervalo de confianza se usa una simulacion de Monte-Carlo. Similar al método bootstrap se producen varias muestras del tamaño de la original, esta vez suponiendo que los errores dados se comportan de manera gausseana, de tal manera que cada una de las nuevas muestras se obtiene como:

$$muestra_i = muestra_{original} + \epsilon * r$$

  Donde ϵ son los errores de la muestra original y r es una variable aletoria que se distribuye como Normal estándar (0,1). En cada muestra se encuentran valores para la pendiente a y la constante b usando $numpy.polyfit$. Finalmente los valores estos parametros con puestos en un arreglo ordenado. Como antes, al límite inferior del intervalo de confianza le corresponde 100(alpha), mientras que al límite superior 100(1-alpha/2). Para un intervalo de confianza del 95% se tiene $alpha = 0.05$.

2.3 Resultados

- Para el ajuste lineal $Banda_z = a * Banda_i + b$, se encontró:

  $a = 1.10$

  $b = 3.14$

  Intervalo de confianza para a = $[0.94, 1.14]$

  Intervalo de confianza para b = $[2.31, 7.91]$

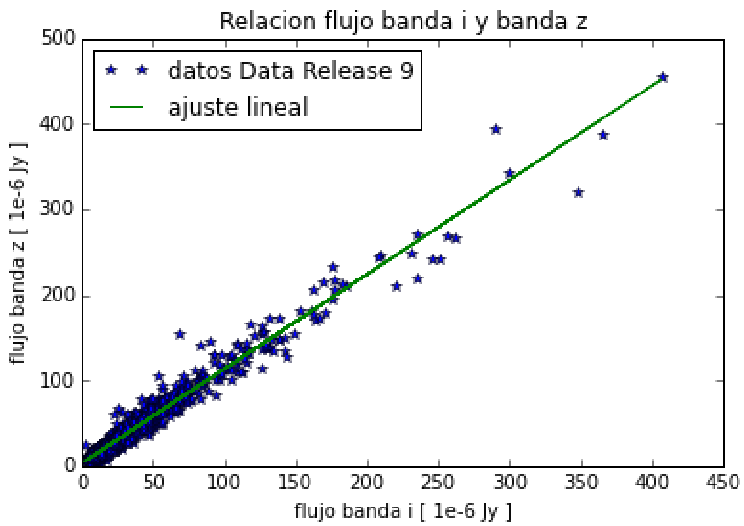

Figure 5: Gráfico flujo banda z versus flujo banda i. Se muestran datos del Data Release 9 del SDSS, y ajuste lineal a los datos

2.4 Conclusiones

- Puede decirse que los valores para los parámetros encontrados a y b, son razonablemente buenos, tanto cualitativamente al observar que en el gráfico de la figura 5 la recta parece ajustarse bien a los datos, como cuantitativamente si se tienen en cuenta que los extremos de los intervalos de confianza se mantienen cerca de los valores obtenidos. Al parecer para la pendiente a el valor encontrado es más preciso que para la constante b, pero de cualquier forma la relación entre los flujos para cuásares en estas dos bandas se ajusta bien con una función lineal.