The Big Bang and the Age of the Universe

In 1929, Edwin Hubble discovered the cosmos. His 1929 groundbreaking paper, *A Relation Between Distance and Radial Velocity Among Extra-Galactic Nebulae*, published in the **Proceedings of the National Academy of Science**

http://antwrp.gsfc.nasa.gov/diamond_jubilee/1996/hub_1929.html

was the first step what was ultimately to be known as the Big Bang Theory. In his paper, Hubble presented data for 24 nebulae illustrating the relationship between the recession velocity, in kilometers per second, and the distance from the earth, measured in megaparsecs (one megaparsec is about 30.9 million trillion kilometers). The recession velocity was measured very accurately by the red shift in the spectrum of light while the distance was measured somewhat inaccurately by comparing mean luminosities of the nebulae to luminosities of known star types. The data are given in Table 1 and the scatterplot in Figure 1 illustrates the relationship.

Velocity	170	290	-130	-70	-185	-220	200	290	270	200	300	-30
Distance	0.032	0.034	0.214	0.263	0.275	0.275	0.450	0.500	0.500	0.630	0.800	0.900

Velocity	650	150	500	920	450	500	500	960	500	850	800	1090
Distance	0.900	0.900	0.900	1.000	1.100	1.100	1.400	1.700	2.000	2.000	2.000	2.000

Table 1: Recession Velocity and Distance for Hubble's Original 24 Nebulae

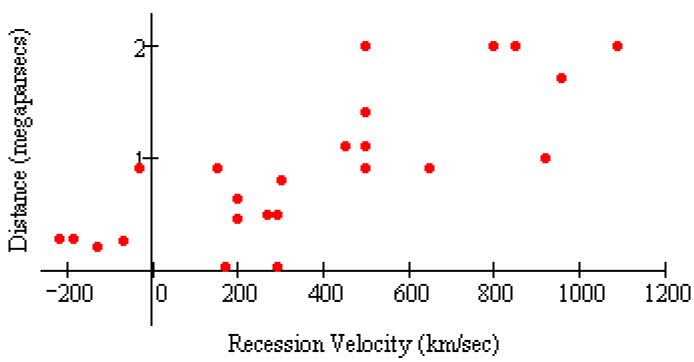


Figure 1: Scatterplot of Hubble's Original Data

We can use these data to estimate the age of the universe. First, we must develop a mathematical model that describes the positions of the earth and the observed nebulae assuming the Big Bang Theory is correct.

According to the Big Bang Theory, both the earth (E) and the nebula (N) have been moving directly away from the site of the Big Band (*BB*) at velocity *v* for time *T*, where *T* is the age of the universe. As these objects move away from *BB*, they also move away from each other. This is called recession. The rate at which they recede from each other is *R*, where *R* is the component of *v* in the direction of the segment connecting *E* and *N*. The distance between earth and the nebula at time *T* is *D*, as shown in Figure 2.

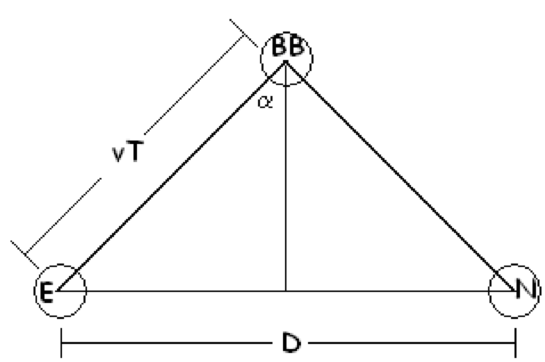


Figure 2: Schematic Drawing of Recession of Earth and Nebula

1) If the Big Bang Theory is correct, then the data in Table 1 should approximate the line $D = R \cdot T$. If so, then the slope of this line is an estimate of *T*, the age of the universe measured in megaparsecs/(kilometer/second) or megaparsecs-seconds/kilometer. We need to fit a line with no intercept, so we can't just plug the data in our calculator. We need a line of the form $D = R \cdot T$, not $D = R \cdot T + b$.

a) First, use your calculator to fit a line $D = R \cdot T + b$. Use the formula for the standard error of the intercept, $s_b = s_e\sqrt{\frac{1}{n}+\frac{(\bar{x})^2}{\sum(x_i-\bar{x})^2}}$, to determine if the intercept is significant. If it is, then this data contradicts the Big Bang Theory.

b) Now, lets fit a line without an intercept. This is a job for CALCULUS! To find the least square line without an intercept, define $S = \sum_{i=1}^{24}(D_i - T \cdot R_i)^2$. Find the value of T that minimizes S in terms of D_i and R_i.

2) Use the data to compute the least squares estimate of *T*. How well does it appear to fit the data. If one megaparsec-sec/km is about 979.8 billion years, how old is the universe according to Hubble's original data? The accepted age of the universe is around 13 billion years. How well did Hubble do? (Of course, the universe is older now since this data was collected in 1929.)

You may have noticed that the no-intercept line appears to lie too low in the data. The model with a non-zero intercept fits better. This is a result of the poor measurements of distance made by Hubble in 1929. If the model truly has a non-zero intercept, then the Big Bang Theory could not be correct. It turns out that Hubble's original data, which ultimately gave rise to the Big Bang Theory, does not actually support the Big Bang Theory.

SOLUTIONS:

1) If the Big Bang Theory is correct, then the data in Table 1 should approximate the line $D = R \cdot T$. If so, then the slope of this line is an estimate of *T*, the age of the universe measured in megaparsecs/(kilometer/second) or megaparsecs-seconds/kilometer. We need to fit a line with no intercept, so we can't just plug the data in our calculator. We need a line of the form $D = R \cdot T$, not $D = R \cdot T + b$.

a) If we use our calculator to fit a linear model, we find that $D = 0.001373R + 0.3991$ is our least squares regression model. Is the intercept 0.3991 significantly different from 0? If so, this data contradicts the Big Bang Theory. To answer this question, we compute a 95% confidence interval for the intercept and see if it contain 0. All 95% confidence intervals are created by using the equation

$$\text{point estimate} \pm t^*_{df}\left(\text{standard error}\right).$$

The standard error for this estimate is $s_b = s_e\sqrt{\frac{1}{n} + \frac{\left(\bar{x}\right)^2}{\sum\left(x_i - \bar{x}\right)^2}}$.

Since there are 24 data points, there are 22 degrees of freedom, so $t^*_{22} = 2.074$. The standard error of the residuals reported by the calculator is $s_e = 0.4050$. The mean recession velocity is 373.125 km/sec. With this information, we can compute s_b .

$$s_b = 0.4050\sqrt{\frac{1}{24} + \frac{\left(373.125\right)^2}{3170090.625}} = 0.1185 .$$

Our confidence interval, then, is 0.3991 $0.3991 \pm \left(2.074\right)\left(0.1185\right)$ or (0.153, 0.645). Based on this data, the true intercept is somewhere between 0.153 and 0.645. Zero is not a plausible value. This data does not support the Big Bang Theory!

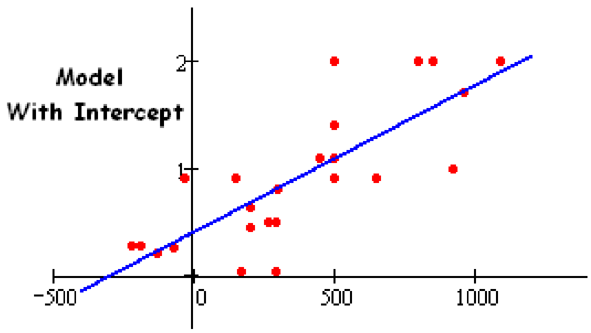
b) To find the least square line without an intercept, define $S = \sum_{i=1}^{24}\left(D_i - T \cdot R_i\right)^2$. Find the value of *T* that minimizes *S* in terms of D_i and R_i .

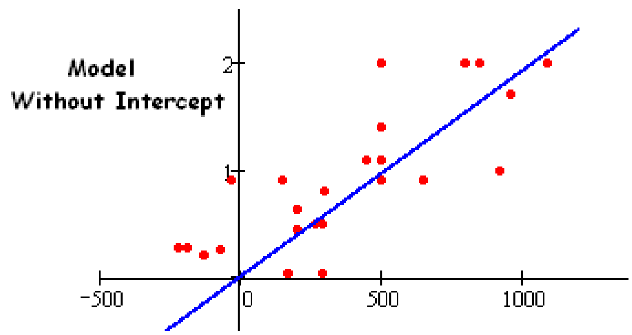
If $S = \sum_{i=1}^{24}\left(D_i - T \cdot R_i\right)^2$, then $\frac{dS}{dt} = \sum_{i=1}^{24} -2R_i\left(D_i - T \cdot R_i\right) = 0$ will find the value of *T* that minimizes *S* (and is therefore our estimate of the age of the universe). Solving for *T*, we find

$$\sum_{i=1}^{24} -R_iD_i + \sum_{i=1}^{24} T \cdot R_i^2 = 0 \text{ so, } \sum_{i=1}^{24} R_iD_i = T\sum_{i=1}^{24} R_i^2 \text{ , and } T = \frac{\sum_{i=1}^{24} R_iD_i}{\sum_{i=1}^{24} R_i} .$$

2) Use the data to compute the least squares estimate of *T*. How well does it appear to fit the data. If one megaparsec/km is about 979.8 billion years, how old is the universe?

With this data, we have $\sum_{i=1}^{24} R_iD_i = 12513.695$, $\sum_{i=1}^{24} R_i^2 = 6511425S$, so $T = \frac{12513.695}{6511425} = 0.0019218$ megaparsecs/kilometer. This is 1.883 billion years.





You may have noticed that the line appears to lie too low in the data. A model with a non-zero intercept appears to fit better. This is a result of the poor measurements of distance made in 1929. If the model truly has a non-zero intercept, then the Big Bang Theory could not be correct. It turns out that Hubble's original data, which gave rise to the Big Bang Theory, does not actually support the Big Bang Theory. By such steps and missteps, science progresses.

Reference: Ramsesy, Fred L. and Daniel W. Schafer, *The Statistical Sleuth*, Duxbury Press, 1997.