INTERPOLATION, EXTRAPOLATION, AND THE LIMITS OF SCIENCE

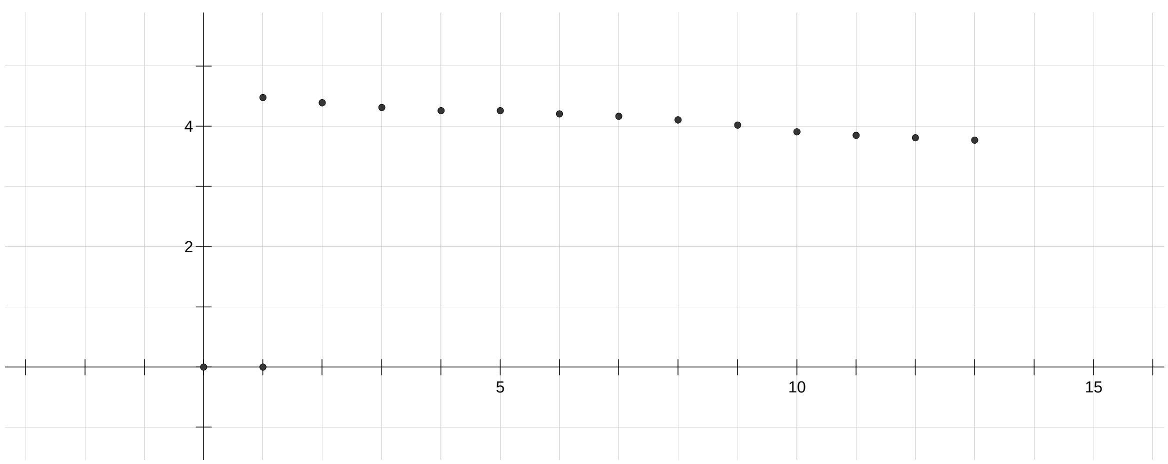
BY JAMES D. NICKEL

Interpolation is the method of is guessing another value of a variable *between* values that are known. *Extrapolation* is the method of is guessing (the word literally means "conjecture") another value of a variable that is *beyond* values that are known.

To illustrate these methods, consider the data by decades for the world record for the mile run (5280 feet):

Date	Runner	Time	Place
1870	W. C. Gibbs	4:28.8	England
1880	Walter George	4:23.2	London
1890	Walter George	4:18.4	Birmingham, England
1900	Thomas Conneff	4:15.6	Travers Island, NY
1910	Thomas Conneff	4:15.6	Travers Island, NY
1920	Norman Taber	4:12.6	Cambridge, MA
1930	Paavo Numi	4:10.4	Stockholm
1940	Sydney Wooderson	4:06.4	London
1950	Gunder Hägg	4:01.4	Goteborg, Sweden
1954	Roger Bannister	3:59.4	Oxford, England
1960	Herb Elliot	3:54.5	Dublin, Ireland
1970	James Ryun	3:51.1	Bakersfield, CA
1980	Steven Ovett	3:48.8	Oslo, Norway
1990	Steve Cram	3:46.3	Oslo, Norway
2000	Hicham El Guerrouj	3:43.1	Rome, Italy

Plotting these points on a graph (each unit on the x-axis represents a decade [unit 1 = 1870] and the y-axis represent time in minutes), we get:



To do this, we converted seconds and tenths of a second into decimal minutes. Here are two examples of how to do this:

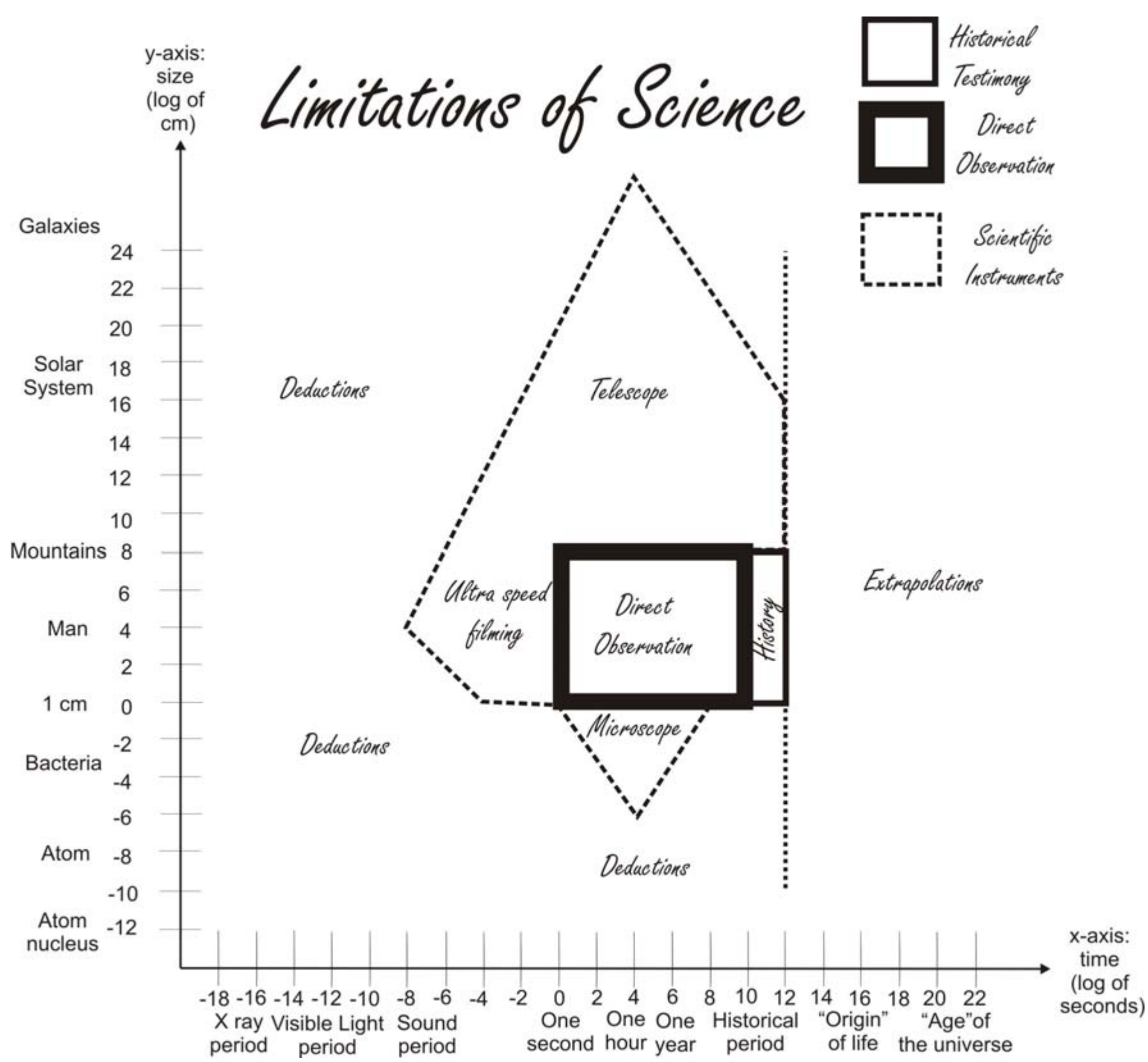
$$4{:}28.8 = 4\frac{28.8}{60} = 4.48$$

$$4.10.4 = 4\frac{10.4}{60} = 4.17$$

Note the apparent linear relationship. In other words, you can "fit the data" by a straight line. In statistics, we call this "line fitting" method *linear regression.* Since the line has a negative slope (or, negative rate of change or common difference that is negative), we can ask the question, "At what date might a 3:30 mile be run?" When we ask this kind of question, we engage in *extrapolation.* Based upon what the data tells us in the past, we are extrapolating in the future. Statisticians define extrapolation as an estimation (of the value of a variable) *outside* the tabulated or observed range. If the range is time-based, we can extrapolate into the past or the future.

We could also use the line to *interpolate* by estimating the records for some of the intervening years (e.g., 1935, 1942, 1978, etc.). Interpolate is introducing something additional or extraneous between other things or parts. In statistics, it means to insert, estimate, or find an intermediate term in a data sequence.

We use both methods with caution. In other words, we are making estimates, not stating matters of absolute truth! If there is too much separation between data points (x-values), interpolation may be in error. Extrapolation can sometimes lead to physically impossible situations. In the mile run, we can extrapolate to a conclusion that sometime, in the distant future, the line will intersect the x-axis (y = 0). This means that someone will be running a 0-minute mile, a physical impossibility!



Interpolation, Extrapolation, and The Limits of Science

by James D. Nickel

Note the *Limitations of Science* chart. The x-axis represents time (as the logarithm of seconds or *log* t) and the y-axis represents size (as the logarithm of centimeters or *log* cm). The y-axis, representing the measurement of size, ranges from the micro (atomic nucleus) to the macro (galactic dimensions). The chart reveals the limitations at which direct observations and scientific instruments (e.g., telescope, ultra speed filming, and microscopes, spectroscopes, betatrons, etc.) can reach. As instrumentation improves via technology, the better the measurement (thus, increasing the range in chart). Scientists extend these observations beyond technological limits using deductive methods. The reliability of these methods are always to be understood as correctable (i.e., not absolute) using, in some cases, statistical margins of error and confidence intervals. Orderly and unbiased scientific work should govern these deductions, but this is not always the case.

The x-axis (to the left of one second) represents fractions of a second for electromagnetic waves in terms of periods[1]). After one second, the time increases to one hour, one year, the length of recorded history, etc. The "etc." means "origin of life" or the "age of the universe" estimates are extrapolations.

The theory of evolution posits (as an *assumption*) *uniformitarianism* meaning all current physical processes (e.g., erosion, radioactive decay, etc.) have *always* been the same in the past. In other words, processes *that* operated in the remote geological past are *not* different from those observed now. Scientists take the *current rate* and *extrapolate* backwards in order to estimate starting dates. As we have seen, this type of extrapolation is merely conjecture and may lead to serious errors. These extrapolations are technically beyond the limits of careful and orderly scientific work, a discipline of study that is properly limited to the study of physical objects and/or events of the present.

---

[1] Period is defined as $\frac{1}{f}$ where f = frequency (number of oscillations per second).