38th SYMPOSIUM ON THE SEARCH FOR EXTRATERRESTRIAL INTELLIGENCE (SETI) – The Next Steps (A4)
SETI I : SETI Science and Technology (1)

Author: Dr. Claudio Maccone
International Academy of Astronautics (IAA), Italy, clmaccon@libero.it

SETI AND SEH (STATISTICAL EQUATION FOR HABITABLES)

Abstract

The statistics of habitable planets may be based on a set of ten (and possibly more) astrobiological requirements first pointed out by Stephen H. Dole in his book "Habitable planets for man" (1964). In this paper, we first provide the statistical generalization of the original and by now too simplistic Dole equation. In other words, a product of ten positive numbers is now turned into the product of ten positive random variables. This we call the SEH, an acronym standing for "Statistical Equation for Habitables".

The mathematical structure of the SEH is then derived. The proof is based on the Central Limit Theorem (CLT) of Statistics. In loose terms, the CLT states that the sum of any number of independent random variables, each of which may be ARBITRARILY distributed, approaches a Gaussian (i.e. normal) random variable. This is called the Lyapunov Form of the CLT, or the Lindeberg Form of the CLT, depending on the mathematical constraints assumed on the third moments of the various probability distributions.

In conclusion, we show that:

1) The new random variable NH, yielding the number of habitables (i.e. habitable planets) in the Galaxy, follows the LOGNORMAL distribution. By construction, the mean value of this lognormal distribution is the total number of habitable planets as given by the simplistic Dole equation. But now we also derive the standard deviation, mode, and all the moments of this new lognormal NH random variable.

2) The ten (or more) astrobiological factors are now positive random variables. The probability distribution of each random variable may be ARBITRARY. The CLT in the so-called Lyapunov or Lindeberg forms (that both do not assume the factors to be identically distributed) allows for that. In other words, the CLT "translates" into our SEH by allowing an arbitrary probability distribution for each factor. This is both astrobiologically quite realistic and very useful for any further investigation.

3) An application of our SEH then follows. The (average) DISTANCE between any two nearby habitable planets in the Galaxy may be shown to be inversely proportional to the cubic root of NH. Then, in our approach, this distance becomes a new random variable. We derive the relevant probability density function, apparently previously unknown and dubbed "Maccone distribution" by Paul Davies.

4) DATA ENRICHMENT PRINCIPLE. It should be noticed that ANY positive number of random variables in the SEH is compatible with the CLT. So, our generalization allows for many more factors to be added in the future as long as more refined scientific knowledge about each factor will be known to the scientists. This capability to make room for more future factors in the SEH we call the "Data Enrichment Principle", and we regard it as the key to more profound future results in the fields of Astrobiology and SETI.

5) A practical example is then given of how our SEH works numerically. We work out in detail the case where each of the ten random variables is uniformly distributed around its own mean value and has a given standard deviation. The conclusion is that the average number of habitable planets in the Galaxy should be around 100 millions, and the average distance in between any couple of nearby habitable planets should be about 88 light years.

6) Finally, we match our SEH results against the results of the Statistical Drake Equation that we introduced in our 2008 IAC presentation. As expected, the number of currently communicating ET civilizations in the Galaxy turns out to be much smaller than the number of habitable planets (about 10,000 against 100 millions, i.e. one ET civilization out of 10,000 habitable planets). And the average distance between any two nearby habitable planets turns out to be much smaller that the average distance between any two neighbouring ET civilizations: 88 light years vs. 2000 light years, respectively. This means an ET average distance about 20 times higher than the average distance between any couple of adjacent habitable planets.