

Applications of Probability models and Combinatorics to Biological sequences

Ourania Chryssaphinou

Prof. Emerita NKUA

ocrysaf@math.uoa.gr

Abstract: We present an elementary introduction to DNA sequences [3], [4], as well as some interesting thoughts of a few famous scientists about discovering them [1]. From the middle of 19th century till the middle of 20th century Statistics was the branch of Mathematics which had a relationship to Biology. That relationship was very productive for both sciences. We refer Karl Pearson (1857-1936) and his research students [6] and R. A. Fisher (1890-1962) for the application of Statistical methods. Before the announcement of double helix all the branches of Stochastic Processes played an important role in the study of Biological Sequences. In 1953 the researchers of Cambridge University James Watson (1928-) and Francis Crick (1916-2004) announced the structure of double helix of DNA publishing their conclusions in Nature [2]. They had used relative results of Rosalind Franklin (1920-1958) and Maurice Wilkins (1916-2004) researchers of Kings College of London. In 1962 Watson, Crick and Wilkins (1916–2004) jointly received the Nobel Prize in Physiology or Medicine. Rosalind Franklin (1920-1958) had been ignored since she had already died by cancer because she was using X-ray for the necessities of the research. Jacques Monod (1910-1976) was a French biochemist who discovered the messenger RNA, a crucial factor in the functioning of the cell. Monod shared the Nobel Prize in Physiology/Medicine in 1965. He wrote a famous book which title is: *Chance and Necessity* [2]. After the impressive development of computer sciences a new epoch for Biology has been started. Thus, J.E.Chen (1844-) wrote an article with the title «Mathematics Is Biology's Next Microscope, Only Better; Biology Is Mathematics' Next Physics, Only Better» [6]. Finally, Lothaire, M. (2005), *Applied Combinatorics on Words*, Encyclopedia of Mathematics and its Applications, **105**, [Cambridge University Press](#) [10] cover the following relative areas: Core algorithms, Natural language processing, Bioinformatics, Algorithms, Mathematics. We note that in the research field «Statistics on words with applications to biological sequences» some relative results obtained by O. Chryssaphinou, S. Papastavridis, E. Vaggelatou, and A. D. Barbour and O.Chrys.. are referred.

References

1. www.damtp.cam.ac.uk/user/tf227/tziotzios.pdf
2. Watson J.D. and Crick F.H.C., «A Structure for Deoxyribose Nucleic Acid», *Nature* **171**, 737-738, 1953.
3. Ζακ Μονό, « Η Τύχη και η Αναγκαιότητα», εκδόσεις Ράππα, 1971.
4. Richard Dawkins, «Το εγωιστικό γονίδιο», πρώτη Αγγλική έκδοση 1976, 30η έκδοση, Κάτοπτρο, 2010.
5. Wigner E. P., «The Unreasonable Effectiveness of Mathematics in the Natural Sciences», *Communications on pure and applied Mathematics*, Vol. XIII, 001-14, 1960.
6. Joel E. Cohen, «Mathematics is Biology's Next Microscope, Only Better; Biology is Mathematics'Next Physics, Only Better», *PLoS Biology*. **2** (12), 2012.
7. Michael S. Waterman, «Introduction to Computation Biology – Maps, sequences and genomes», Chapman & Hall, 1995
8. Lothaire, M., *Applied Combinatorics on Words*, *Encyclopedia of Mathematics and its Applications*, **105**, Cambridge University Press, 2005.