Population Genetics with Statistical

by Yuejian Mo

12-21-2017

Abstract

Population genetics is the subfield of genetics, which deals with genetic differences inside and between population, and is primarily found by Sewall Wright, J. B. S Haldane and Ronald Fisher. Fisher also is describled as "a genius who almost single-handedly created the foundations for modern statical science". So, population genetics has strong relation with statistical. It will be enjoyable to review the work from the devolopment of population genetics and foundation of modern statistical to nowdays. Major fundamental and meaning topics of population genetics and modern statistical are reviewed, escpeically how to modern statical methods are being developed for answering population genetics proplem.

1 The population genetics and statistics before 1919

The story

In 1908, Hardy and Weiberg principle each published papers describing a mathmatical relationship between allele frequencies and genotype frequencies. This relationship, now called the **Hardy-Weinberg principle,** allows us to predict a population's genotype frequencies from its allele frequencies.

It suppose that in a population, a particular gene is segregating two alleles, A and a, and that the frequency of A is p and that of a is q.

2 Drawin's Evolution Theory and *Biometrika*

In 1859, **Charles Robert Drawin** published *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Reaces in the Struggle for Life*. This book introduce idea— **survival of the fittest**. It seem make sence. The similariy of some specials, show that model organism can evolve from pre-organism in the past. For example, different bird in isolative island still are common in anatomy, and different embryon has similary, even human embryo exist tail during the early development period.

However, Drawin can only explain why model organism exsite. In other words, Drawin can not find a new speicals, which evolutes from old specials. The problem inspired some mathematicians who appreciated evolution theory to find the answers. In 1901, Francis Galton, Karl Pearson, and Raphael Weldon established *Biometrika* to promote the study of biometrics. In the frist issuse, they annonuce that, this new jourist will collecct data from all over the world to meansure parameters. They expected that parameter varivtion relate with enviornment varivition. Here represents the mathemaical process.

> It is intended taht *Biomerika* shall serve as a means not only of collecting under one title biological data of a kind not systematically collected or pulished in any other peiodical, but also of spreading a knowledge of such statistical theory as may be requisite for their scientific treatment.

> A very few years ago, all those problems which depend for their slution on a study of the differences between individual members of a race or species, were neglected by most biologists. The complexity of organic structure is so great, and the number of distinguishable formes so enormous, that morphologists were obliged ot simplify their conceptions by constructing for every sepcies an ideal type, to which the individuals composing ti tconform with more or less eactness, and to neglect those deviations from the type which actually occur. Such simplification was ont only justifiable, but absolutely for many purposes; it has rendered enormous service to biology in the past, it does so still, and will continue to do so; nevertheless, there are many problems which cannot be dealt with by its aid.
>
> The starting point of Darwin's theory of evolution is precisely the exsitence of those differences between individual members of a race or species which morphologists for the most part rightly neglect. The first condition necessary, in order that any process of Natural Selection may begin among a race, or species, is the exsitence of differences among its members; and the first step in an enquiry into the possible effect of a selective process upon any character of a race must ge an estimate of the frequency with which individuals, exhibiting any given degree of abnormaylity whith respect to that character, occur. The unit, with which such an enquiry must deal, is not an individual but a race, or a statistically representaative sample of a race; and the result must take the form of a numerical statement, showing the relative frequency with which the varous konds od individuals composing the race occur.("(I.) The Scope of Biometrika," 1901)

3 Finsher and Argriculture

ANOVA

4 The 14 years in Rothamsted Experiment Station

Sir Ronald Aylmer Fisher FRS (17 February 1890 -29 July 1962), was a Brithsh statistician and geneticist. In 1919, he began working at the Rothamsted Experimental Station for 14 years, where he analysed its immense data from crop experiments since the 1840s., and developed the anaylsis of variance(ANOVA).

5 References

https://en.wikipedia.org/wiki/Charles_Darwin

(I.) The Scope of Biometrika. (1901). *Biometrika*, *1* (1), 1–2. https://doi.org/10.1093/biomet/1.1.1