

BRK-012 · Discrete genes and continuous variation: a statistical reconciliation

BRK · Conceptual Breaks

Juan Segurolo · All Rights Reserved

Human height varies along an apparently continuous scale. Hereditary factors can nevertheless be transmitted in discrete alternatives. The reconciliation of these observations required more than choosing a winner between two descriptions. It required showing how one level of organization could generate the statistical pattern observed at another.

The conceptual change was to treat a distribution of measured traits as something that could be explained by combinations of inherited effects, environmental contributions and explicitly stated assumptions. Discreteness in transmission no longer implied a small number of sharply separated visible categories. Conversely, a smooth-looking distribution no longer established that hereditary material itself blended.

1. Two kinds of evidence, one apparent contradiction

Biometrical investigations measured continuous variation and resemblance among relatives. Mendelian investigations examined segregation of hereditary factors. In the early twentieth century, disagreement concerned whether the first class of observations could be reconciled with the second. Neither approach was simply an absence of quantitative thinking. They quantified different objects. [1]

It is especially misleading to assign every earlier investigator one uniform belief in blending inheritance. The historical positions were more varied, and attempts at reconciliation preceded Fisher's 1918 paper. Pearson, Yule and Weinberg had already contributed to the problem. Fisher's work became a major articulation of the statistical reconciliation, rather than an argument produced in an empty field. [1]

The observational contrast can be stated without retelling the familiar crosses. A factor with two alternatives may have a discrete transmission rule. A measured phenotype can depend on many such factors. The inference from "each contribution is discrete" to "the resulting distribution must have only a few categories" is invalid.

2. Constructing continuity from discrete contributions

Consider an illustrative model with ten independently inherited loci. At each locus, an individual carries two copies, each contributing either zero or one unit to a trait. Assume equal frequencies, random combination, equal additive effects and no linkage. The total genetic contribution then ranges from zero to twenty units.

There are many more ways to obtain a middle total than an extreme total. Under these deliberately simple assumptions, the total is binomial with twenty trials and probability one-half. Its mean is 10 and its variance is 5. A single extreme total has probability $1/1,048,576$, while the central total has probability $184,756/1,048,576$, approximately 0.1762.

These are properties of a constructed model, not measurements of human height or counts of height genes. More loci, unequal effects and environmental variation can further blur visible categories. The example establishes compatibility between discrete inheri-

tance and a distribution that can look increasingly continuous. It does not establish that every real quantitative trait has precisely this architecture.

Environmental effects can be included as a separate contribution. But simply adding a residual term is not a complete biological explanation of the environment. It is an accounting decision that must be related to how observations were made and which influences the model actually represents.

3. Resemblance becomes a problem about covariance

Fisher's 1918 paper addressed correlations between relatives under Mendelian inheritance. This goes beyond demonstrating that many discrete contributions can make a smooth distribution. Relatives share hereditary connections, so their trait values can covary in ways that a genetic model predicts. The paper also treated complications including dominance and assortative mating. [2]

Suppose, in a simplified population model, phenotype is represented as an additive genetic value plus an environmental residual. Its variance equals the genetic variance plus the environmental variance only if their covariance is zero. Otherwise the decomposition also contains twice that covariance. Writing a compact equation does not remove the need to justify the independence assumption.

Similarly, resemblance between relatives is not automatically a pure measurement of genetic sharing. Relatives may share environments, and mating patterns can affect genetic associations. A statistical reconciliation makes these problems analyzable; it does not abolish them.

The advance was to ask whether one set of assumptions could account jointly for variation within a population and resemblance across relationships. An explanation gained strength by surviving several such constraints. Fitting one correlation alone would leave more alternatives available.

4. A variance is not a percentage of a person

This framework encouraged a new language of components of variation. That language is easy to misuse. A statement about the proportion of variance associated with additive genetic differences in a specified population is not a statement that the same proportion of an individual's body was manufactured by genes.

Variance describes differences around a population mean. Its size depends on the population and conditions. If an environmental source of variation becomes more uniform, the relative contribution of genetic variation can increase even when no individual's biology has become less dependent on an environment.

Nor does a large genetic contribution to observed variation imply immutability. A change applied broadly across a population can alter the mean while leaving much of the rank ordering intact. The model of existing differences and the prediction of an intervention are related questions, but one is not a substitute for the other.

The distinction can be made without assuming that genes and environments act as physically separable ingredients. An additive statistical approximation can be useful even when the underlying developmental mechanisms involve interactions. Whether it is adequate for a particular inference must be checked against that inference.

5. What the reconciliation left open

The reconciliation did not reduce all inheritance to simple Mendelian ratios. Linkage, dominance, interactions, shared environments and different mating systems can alter the statistical relationships. Neither did it reveal the molecular identity of every contributing factor. It showed how a class of inheritance mechanisms could account for a class of population patterns.

Modern genomic work makes some of those contributing differences observable at a scale unavailable in 1918. The historical achievement should nevertheless be judged by the problem then being addressed: the apparent incompatibility of continuous phenotypic variation with particulate inheritance. Reading later molecular detail back into the original model would obscure what was and was not established.

6. A different relationship between mechanism and pattern

The enduring break was a change in the relation between a rule for individual transmission and a pattern across many individuals. Those descriptions no longer had to compete as mutually exclusive accounts. A statistical theory could connect them, while exposing the assumptions needed for the connection.

That connection also created new obligations. One must specify the population, the relationships, the environmental structure and the quantity being estimated. The reward is not a universal number explaining heredity. It is a framework in which disagreement about observed resemblance can become disagreement about explicit, testable models.

Related reading

BRK Article 014, “Evolution without an adaptive explanation for every change,” follows the use of population-level reasoning to distinguish the production of variation from its changing frequency.

References

- [1] Visscher, P. M., and Goddard, M. E. “From R.A. Fisher’s 1918 Paper to GWAS a Century Later.” *Genetics* 211, 1125–1130 (2019). <https://doi.org/10.1534/genetics.118.301594>
- [2] Fisher, R. A. “The Correlation between Relatives on the Supposition of Mendelian Inheritance.” *Transactions of the Royal Society of Edinburgh* 52, 399–433 (1918). <https://digital.library.adelaide.edu.au/items/5ed47714-2635-4af9-8409-cf7eb6f87c3b>