I highly recommend the book as an up to date record of research in a rapidly moving field, and should be required reading for all muscle biologists and also clinicians interested in the workings of skeletal muscle.

F S WALSH

Methodology in Medical Genetics

2nd edition. By Alan E H Emery. (Pp 192; £24.00.) Edinburgh: Churchill Livingstone. 1986.

This is, in the best sense of the word, a cook book. It aims to provide the collector of raw data with ways of analysis which will make it more digestible. The purchaser of recipe books rarely has access to a full range of raw materials, and will judge it by those parts of interest. The reviewer has a more difficult problem, and has to read it from beginning to end, and assess each recipe by such criteria as to whether cooking improves the raw material, and, if so, whether the recipe is the best, and, if it is, if the presentation is clear. If, as in this case, the reviewer also reviewed the first edition, a comparison is needed. This at least is easy: it covers new territory, it is more elegantly produced, and it is more expensive.

The raw data cover a range sufficient for the most omnivorous geneticist, and is itself of value for its interest, variety, and reliability. This is to be expected from an author who has been productive both of raw data and of several lucid books, including a recent book on recombinant DNA techniques. The recipes are almost entirely unexplained, and their origin and reliability are a severe test on any reviewer. Such key references as those to Galton, Pearson, Robbins, Fisher before 1930, Muller, Malecot, Crow and Kimura, and Kimura after 1965 are missing. The frequent and necessary references to Cavalli-Sforza and Bodmer lack a page reference.

There are some errors and unnecessary obscurities. In the hope that their documentation will be productive, and of use to the many readers the book deserves, I will mention these in order.

The Hardy-Weinberg equilibrium is introduced with the interesting embellishment that "the mutation rate remains constant". Hardy did not mention this; unless zero it destroys the equilibrium. The treatment is an advance on Hardy for lucidity, though not brevity. Gene frequency estimation starts with non-codominant systems, rather than with simple gene counting.

The next chapter, on population structure, introduces, without explanation, the formula $100/2N_e$ for the percentage of loci which will become fixed or

eliminated in each generation. The sense is clear, but the formula is orphaned, and not helped by the '100'. In fact N_e , the effective population in genetical terms, is also the average life of a neutral allele in generations. Unfortunately the probability of loss is not independent of gene frequency, and, although the next sentence rescues this potential misinterpretation by implying an averaging over all loci, lucidity is hardly achieved.

The effective population size, N_e , is the product of the total population and the proportion which reproduce, weighted by its fecundity. It is not the product of the number and the proportion of their lives for which they are capable of reproduction. The effect of variance (p_{14}), due to Wright (1938), is not explained, although it is conceptually simple, since $(V + 2)$ is a 4 in a stable population with Poissonian fertility, when both the mean and variance are 2, the average number of children.

Under gene flow there is perhaps the only example in the book of a highly implausible explanation – the attribution of blood group B in Western Europe to the Mongolian invasions – which hardly reached Ireland, Wales, Iceland, or indeed anywhere beyond the Rhine. The use of absolute values in the equation on p23 is wrong, or would be if the signs differed: if they do not the ratio is always positive. If they do something impossible has happened.

An ingenious inversion of an equation featuring migration is used to estimate migration: however, this is clearly wrong, since it will estimate migration when applied to an isolate, as, indeed, it appears to have done in the example. Under selection (p27) the retention of 'q', when very small in the divisor $(1-q)$, leads to complications, including a graph with one logarithmic axis, basically showing two numbers to be identical. Tanaka's simple and rarely used method of estimating fitness is clearly described.

Segregation analysis, even if its interest is now mainly historical, is dealt with clearly, with some simple and little known estimators, such as the 'singles method'. The three pages of tables are unlikely to find much use against the personal computer, which makes it simple to use one-off trial-and-error approaches. The variances, on which so much algebraic ingenuity has been exercised, are of limited interest since they do not relate to symmetrical distributions, and exact graphical displays can be produced if needed. The chapter is a fine example of type setting and proof reading, but it would be easier if 'q' had not become promoted from its minority status earlier in the book, as is now usual. Heritability, following Falconer's lucid approximations, is clearly described but regression is not, and there is no help in the way of diagrams of

bivariate surfaces. The statement that the child-midparent regression (1.0) differs from the midparent-child regression (0.5) seems wrong. The relationship must be symmetrical, both values being the square root of $1/2$, their product being $1/2$, as is stated. Regression is difficult, but easier than heritability, and it would seem unwise to enter these shoal waters without an introduction through regression.

Linkage is introduced clearly, although, as the author acknowledges, no more clearly than by Race and Sanger in the recent editions of *Blood Groups in Man*. The use of probability as a synonym of likelihood, and the failure to explain why log likelihood is easier to work with, seem unnecessary, but in the absence of any good elementary description of linkage it is unfair to criticise one of the best attempts to clarify a field so fertile in error and false hope. Lod scores are given for the $z1$ situation, but only the e ascertainment corrections given; the more commonly needed c corrections are not mentioned. The figure on p78 confuses likelihoods and log likelihoods; three of the four curves given should be asymptotic to the y axis, and only curve a should show any upward convexities. All should be tangential to the X axis at $\theta = 0.5$.

Twins are introduced with a clear biological introduction, and a useful set of examples, including one on zygoty estimation. Holzinger's simple index of similarity gets short shrift, although it is difficult to criticise it merely on the grounds it is unrelated to heritability unless one finds heritability estimates more meaningful. In the next chapter on recurrence risks the author is very much on home ground, and includes the use of linked loci for predictive purposes, with special reference to both Huntington's chorea and Duchenne muscular dystrophy. However, the false optimism on reliability inseparable from applying the maximum likelihood estimates of recombination is not considered.

In the section on disease associations, linkage disequilibrium is explained in terms of alleles being "preferentially maintained in coupling", as was advanced by Fisher (1930) under the term 'linkage equilibrium'. This difficult matter is not discussed further. The table of HLA associations is unfortunately culled from before the era of HLA-DR typing.

A chapter on the resolution of genetic heterogeneity attacks a difficult, and poorly colonised,

field of enquiry and is more valuable for the lucidity with which problems are defined than for the remedies proposed. The use of cousin incidence to define heterogeneity leads to a mildly erroneous formula. The gamete from the aunt or uncle of a recessive disorder has a probability of conveying the affected allele of $1/4 + 3q/4$, and a probability q of meeting an affected gamete, giving a cousin incidence of $(1 + 3q) q/4$, or approximately $q/4$, rather than more complicated, but essentially equal, $pq/4$ given. The formula for numbers of loci, and the accompanying graph, assume equal gene frequencies at all loci.

Finally there is a most valuable pair of chapters on parental age and birth order, matters rarely considered in modern books, and on changes in disease frequency. The analysis of trend by traditional contingency methods is not considered. On p155 a formula is made unnecessarily difficult by χ^2 , the normal deviate, whose square is χ^2 for one degree of freedom. There is a simple introduction to CUSUMS, or cumulative sums, a valuable but inadequately used method of checking for systematic changes.

The book is well provided with tables, but the normal curve is not tabulated. The z transformation of the correlation coefficient is fully tabulated. Incidentally this can be used to convert Kosambi's mapping function.

This is a useful book, and has few competitors beyond its own first edition; such books are very difficult to write, as the depth of experience necessary to decide whether the raw material is worthy of either notice, or of cooking, leaves little time to develop the mathematical niceties. There are books with more explicit mathematical foundations and with better references to earlier work, but these are often preoccupied with the solution of problems of no biological interest, or advance solutions of such complexity that time is better spent collecting more data and analysing them less efficiently.

Most of the simpler numerical problems which present in practice are covered, with the exception of those related to radiation. Human geneticists, especially those working with problems of environmental effect, will find it a useful source of reliable procedures. Its many and well justified references to Cavalli-Sforza and Bodmer make this an almost obligatory companion.

J H EDWARDS