

CONTENTS

Preface

page v

I LONG RUN FREQUENCIES

I

Statistical inference is chiefly concerned with a physical property, which may be indicated by the name *long run frequency*. The property has never been well defined. Because there are some reasons for denying that it is a physical property at all, its definition is one of the hardest of conceptual problems about statistical inference—and it is taken as the central problem of this book.

II THE CHANCE SET-UP

13

The long run frequency of an outcome on trials of some kind is a property of the *chance set-up* on which the trials are conducted, and which may be an experimental arrangement or some other part of the world. What the long run frequency is, or was, or would have been is to be called the *chance* of the outcome on trials of a given kind. Chance is a dispositional property: it relates to long run frequency much as the fragility of a wine glass relates to whether the glass does, or did, or will break when dropped. The important notion of a distribution of chances is defined, and chances are shown to satisfy the well-known Kolmogoroff axioms for probability. These are the beginnings of a postulational definition of chance.

III SUPPORT

27

Although the Kolmogoroff axioms help to define chance they are not enough. They do not determine when, for instance, an hypothesis about chances—a statistical hypothesis—is well supported by statistical data. Hence we shall need further postulates to define chances and to provide foundations for inferences about them. As a preliminary, the idea of support by data is discussed, and axioms, originally due to Koopman, are stated. Much weaker than the axioms commonly used in statistics, they will serve as part of the underlying logic for further investigation of chance.

IV THE LONG RUN

39

Some connexions must be established between chance and support. The simplest is as follows: given that on some kind of trial an event of kind *A* happens more often than one of kind *B*, then the proposition that *A*

occurs on some individual trial is, lacking other data, better supported than the proposition that *B* occurs at that trial. Something like this feeble connexion is widely held to follow from its long run success when used as the basis for a guessing policy. But no such long run defence is valid, and the connexion, if it exists at all, must have an entirely different foundation.

V THE LAW OF LIKELIHOOD 54

A principle named the *law of likelihood* is proposed as an explication of the banal connexion stated in the preceding chapter, and then is used as a postulate to be added to Koopman's and Kolmogoroff's axioms as part of the postulational definition of chance. It relies on mere comparisons of support, justifying conclusions of the form, 'the data support this hypothesis better than that'. But it is relevant not only to guessing what will happen on the basis of known frequencies, but also to guessing frequencies on the basis of observed experimental results. Any such suggested law must be regarded as a conjecture to be tested in terms of its consequences, and this work is commenced.

VI STATISTICAL TESTS 74

The traditional problem of testing statistical hypotheses is examined with a view to discovering general requirements for any theory of testing. Then it is shown how one plausible theory is in exact agreement with the law of likelihood.

VII THEORIES OF TESTING 89

A rigorous theory of testing statistical hypotheses is developed from the law of likelihood, and other theories are contrasted with it. The Neyman-Pearson theory is given special attention, and turns out to be valid in a much narrower domain than is commonly supposed. When it is valid, it is actually included in the likelihood theory.

VIII RANDOM SAMPLING 118

The theory of the preceding chapters is applied to inferences from sample to population. Some paradoxes about randomness are resolved.

IX THE FIDUCIAL ARGUMENT 133

So far in the essay there has been no way to measure the degree to which a body of data supports an hypothesis. So the law of likelihood is strengthened to form a *principle of irrelevance* which, when added to other standard axioms, provides a measure of support by data. This measure

may be regarded as the consistent explication of Fisher's hitherto inconsistent theory of fiducial probability; at the same time it extends the postulational definition of chance within the pre-established underlying logic. It is proved that any principle similar to the principle of irrelevance, but stronger than it, must lead to contradiction. So the principle completes a theory of statistical support.

X ESTIMATION

161

A very general account of estimation is provided in preparation for a special study of estimation in statistics.

XI POINT ESTIMATION

174

The traditional theory of point estimation is developed, so far as is possible, along the lines of the preceding chapter, and relying on the theory of statistical support. Unfortunately the very concept of estimation seems ill adapted to statistics, and unless other notions are imported, it is impossible to define a 'best estimate' for typical problems. At most *admissible estimates* can be defined, which are not demonstrably worse than other possible estimates. Usually the currently popular theories of estimation provide estimates which are admissible in the sense of the present chapter, but where popular estimates diverge, there is seldom any way of settling which is best without some element of arbitrary convention or whim.

XII BAYES' THEORY

190

Because of the occasional limitations in the theory of statistical support, the bolder theories of Bayes and Jeffreys are examined, but each is rejected for reasons which by now are entirely standard.

XIII THE SUBJECTIVE THEORY

208

Bayes' ideas have naturally led to a subjective theory of statistical inference, which is here explained sympathetically, and shown to be in principle consistent with our theory of statistical support. Neo-Bayesian work claims to analyse a much wider range of inferences than our theory attempts, but ours gives the more detailed account of the inferences within its domain, and hence it has the virtue of being more readily open to refutation and subsequent improvement.