193

λ. This region has been chosen so that the probability of μ actually lying in the outer zone is only 5%; any other probability could equally have been chosen.

true

λ Without sufficient, or, more generally, exhaustive estimation, the mere fact that the ~~fixed~~ true value will lie between calculated limits with a ~~known~~ frequency, realisable by repeated sampling, is not equivalent to a probability statement about the unknown, for information included in the data may have been lost in ~~[illegible]~~ calculating these limits. True probability statements derived from the observations require all the information available, and this implies in particular ~~either~~ either (a) that no probability statements can be made prior to these observations or (b) that the estimation of the limits shall have been exhaustive and have included all that the data supply. When there really is exact knowledge a priori, Bayes' method is available.

112

hypothesis, which the experiment is capable of testing. In this case the data of the experiment, and the test of significance based upon them, have divided this continuum into two portions. One, a region in which μ lies between the limits 0·03 and 41·83, is accepted by the test of significance, in the sense that values of μ within this region are not contradicted by the data, at the level of significance chosen. The remainder of the continuum, including all values of μ outside these limits, is rejected by the test of significance. λ

It can now be seen that the t test is not only valid for the original null hypothesis that the mean difference is zero, but is particularly appropriate to an experimenter who has in view the whole set of hypotheses obtained by giving μ different values. The reason is that the two quantities, the sum and the sum of squares, calculated from the data together contain all the information supplied by the data concerning the mean and variance of the hypothetical normal curve. Statistics possessing this remarkable property are said to be *sufficient*, because no others can, in these cases, add anything to our information. The peculiarities presented by t, which give it its unique value for this type of problem, are :—

(i) Its distribution is known with exactitude, without any supplementary assumptions or approximations.

(ii) It is expressible in terms of the single unknown parameter, μ, together with known statistics only.

(iii) The statistics involved in this expression are sufficient.

In fact, it leads to an exact specification of μ as a Random Variable; probability statements about the unknown are correct, in the light of the observation, at all levels.

λ