

and van Rooyen provides the basis for studies of this kind. While we have, in penicillin and allied substances, instances of the action of metabolic products of an organism on bacteria, we have, in the work here recorded, an instance of the reverse process—the action of bacterial metabolic products on an amoeba physiologically equivalent to the phagocyte, which is not, fortunately for us, affected by penicillin. The work of Castellani and van Rooyen brings us, in fact, nearer to the work on the opsonins and similar substances—work which aimed at rendering the invading bacteria more palatable to the amoebic phagocyte, or at any rate aimed at helping the ingestion of these bacteria by the phagocyte. It is possible that the work done with *H. Castellanii* might be best developed with this idea as its basis.

G. LAPAGE.

AGRICULTURAL SAMPLE SURVEYS

THE need for adequate statistics relating to our agricultural resources and requirements must have become obvious to all during the last few years, for the urgency of war problems has served to direct attention to the inadequacy of peace-time data and also to the methods of rapidly filling the deficiencies. Complete and reliable censuses are often impracticable and always make great demands on both time and skilled labour. Where the need is more for a quick and reasonably accurate estimate of crop acreage, yield, or whatever it may be, sample surveys will generally offer a better method of obtaining the data. These set out to arrive at an estimate of the whole from the collection of a limited sample of representative parts. The dangers of such an approach are as clear as the advantages, and only by conducting the sample surveys along sound statistical lines can biased or distorted estimates be avoided and a measure of the reliability of the estimate be secured.

The need for information of the kind given by sample surveys is, of course, confined neither to Great Britain nor to war-time. In 1937 a statutory body called the Indian Central Jute Committee initiated, as one of its first tasks, a five-year scheme for obtaining improved estimates of the area under jute in Bengal. After some hesitation it was decided to use the sample survey method, the earlier years being devoted to small exploratory surveys, with a complete survey of about 60,000 square miles in 1941. It was laid down that the final estimate of area under jute should have a margin of error not exceeding 5 per cent, that it should be ready early in the jute season and that the cost should not be excessive. P. C. Mahalanobis, who was statistical adviser to the scheme, has now published an account of the methods, both organizational and statistical, by means of which the task was successfully accomplished. The final estimate was within 2.8 per cent of an independent official estimate based on census data; it was ready a week or so before the latest useful date; it cost only about £8,500 as against £110,000 for a complete census. In view of this the Jute Census Committee recommended the adoption of sample surveying to the Indian Government.

Mahalanobis' paper ("On Large-scale Sample Surveys", *Phil. Trans. Roy. Soc.*, B, 231, 329–451) is divided into three parts. Part 1 describes the way in which the problem arose, outlines the method of approach and discusses production and mapping

surveys in addition to those concerning acreage. Part 2 is a mathematical treatment of the statistical theory of various methods of sample surveying. The concepts and principles are dealt with mainly in the abstract, but the results of model sampling experiments are also used. Part 3 concerns the application of this theory to crop area estimation, especially the jute survey of Bengal. The experimental results are summarized and numerical examples worked out.

The fields under jute in Bengal vary much in size, and furthermore any field may be only partly devoted to this crop. It was therefore decided to take as the sampling unit areas, termed grids, of a definite size, like four or twenty acres. The proportion of the land given to jute in each grid was ascertained, and by combining these proportions from all the grids, which were randomly located over the jute-growing area, an estimate of the jute acreage was obtained. Both the precision and the cost of this estimate depend on the area of each grid and the number of grids (that is, density per square mile) surveyed. Now for any given cost, the larger each grid is, the smaller is the density that can be used. The problem is then to adjust grid size, and with it density, so as to maximize the precision of the final estimate. The alternative procedure, which though discussed was not used for the jute survey, is to adjust grid size and number to minimize the cost for a given level of precision.

Two functions, relating cost and precision (variance) to grid size and number, were set up. The constants which they contained were estimated empirically from the data of the early exploratory surveys and by their aid the final survey was planned. The cost function was found to involve consideration of time necessary for enumerating the jute areas within each grid (which depends on grid size but not on density), of time necessary for journeying from grid to grid (which depends on density but not size), of miscellaneous time (independent of both size and density) and of time needed in the statistical laboratory. The precision, or variance, function was found to involve a parameter which took into account the correlation of cropping on adjacent fields. It was also shown that precision varied with proportion of land under jute in the grid, so that the adjustment of grid size and density best for one proportion would not be best for another. For this reason the area to be surveyed finally was divided into zones of more or less homogeneous proportions of jute land, and the best grid sizes and densities found for each zone separately.

Linked pairs of sub-samples, at constant distances apart but randomly orientated, were used to give the standard error of the final estimate. These were always surveyed by different groups of enumerators and at different times, so as to prevent collusion.

The laboratory methods of organizing the survey and randomizing the grids are described in detail, as are the kinds of errors arising from untrained and even dishonest enumerational labour. The means used to adjust the work to the very varied speeds of the enumerators are also mentioned. In the discussion of the planning of sample surveys it is emphasized that surveys of the kind undertaken are progressive. Each one adds to the information relating to the cost and precision functions and their changes with zone and time. So each enables a better survey to be planned for the next occasion. Finally, a detailed account of the work of others on survey sampling is appended.