3. Multi-stage sampling.

P. C. Mahalanobis, Calcutta.

A technique of large scale sample survey in which the area under investigation is divided into a number of zones from each of which a suitable number of sampling units are selected at random (for observation or treatment) was found by the author to be appropriate for a certain type of enquiry (e.g. the census of acreage under a particular crop). The corresponding theory was developed in certain memoirs and also discussed at previous sessions of the Science Congress. For other types of enquiry (e.g. estimate of the total yield of a particular crop) this procedure would be extremely expensive and a different technique has to be adopted. The whole area under investigation is divided into a number of zones (say K) each of which is cut up into a number of units N_1, each of which again into a number of sub-units N_2 each of which again is divided into a number of smaller sub-units N_3 and so on till we get to the last stage. This is the population to be sampled. From each zone a number of units n_1 is selected at random from each of which again a number of sub-units n_2 are selected from each of which again a number of smaller sub-units n_3 are selected and so on till we get the last stage where we have actual observations in terms of which the total quantity under investigation has to be estimated. The process of selection at each stage is random. This technique has been called one of multi-stage sampling. An estimate of the total quantity or value under investigation (e.g. yield of crop) has been worked out in terms of the last stage observations and variance of this estimate has also been obtained. Subject to a given cost level and certain simple types of cost function (which, however, take into account

the journey cost) the best design (one leading to minimum variance) has been worked out giving the appropriate choice of n_1, n_2, etc. at different stages of this multi-stage sampling.