International Statistical Review, 45 (1977) 173–179 © International Statistical Institute/Printed in Great Britain

Optimum Allocation of Sample Size and Prior Distributions: a Review

T. J. Rao[1]
CSIRO Division of Mathematics and Statistics, Melbourne, Australia

Summary

In this review paper, the problem of optimum allocation of sample size to strata is examined in the light of *a priori* distributions. We first consider Neyman's optimum allocation in the case of the Simple Random Sampling (SRS) scheme and review the results concerning the justification for the assumption that the unknown proportionate values of within stratum variances σ_i^2's can be replaced by the known proportionate values of α_i^2's which are the corresponding within stratum variances based on the auxiliary information. We next review the problem of optimum stratification with proportional and optimum allocations in the case of SRS in the light of *a priori* distributions and discuss the role of arranging the auxiliary variate in increasing order for stratification. We then turn our attention to varying probability sampling schemes and present a review of the results based on a study discussing whether a stratified πPS (PPS) strategy with various non-optimal allocations is likely to be worth while in practice.

1 Introduction

Consider a finite population of size N divided into k strata of sizes N_i, $i = 1, 2, \ldots, k$. Let $\mathscr{Y}$ be the study variable parametric functions of which we are interested in estimating taking values Y_{ij} on the jth unit of the ith stratum; values X_{ij} of $\mathscr{X}$ (a positive auxiliary characteristic closely related to the characteristic $\mathscr{Y}$ under study) are available for all units, $j = 1, 2, \ldots, N_i$; $i = 1, 2, \ldots, k$. In the case of Neyman's optimum allocation (Neyman, 1934) of sample size to strata for Simple Random Sampling with replacement in each stratum, we have the allocation given by $n_{i,\text{opt.}} = nN_i\sigma_i/\sum N_i\sigma_i$, where n is the total sample size and σ_i^2 is the within variance for the ith stratum, $i = 1, 2, \ldots, k$. Computation of $n_{i,\text{opt.}}$'s requires at least the proportionate values of σ_i^2's which are unknown. In practice, some estimates α_i^2's (based on a pilot study or prior information) of σ_i^2's are substituted. These estimates, usually, are the within stratum variances of the auxiliary information $\mathscr{X}$ closely related to the characteristic under study.

Cochran (1946) showed that whenever auxiliary information on a characteristic $\mathscr{X}$ closely related to the characteristic $\mathscr{Y}$ under study is available, this information can be used to set up a criterion of optimality by regarding $\boldsymbol{Y} = (Y_{11}, Y_{12}, \ldots, Y_{kN_k})$ as a realization of an N-length random vector with distribution depending on $\boldsymbol{X} = (X_{11}, X_{12}, \ldots, X_{kN_k})$ and some unknown parameters. Given $\boldsymbol{X}$, we explicitly formulate our model $\theta(g)$ thus:

$$\begin{aligned} &\mathscr{E}_{\theta(g)}(Y_{ij} \mid X_{ij}) = a + bX_{ij} \\ &\mathscr{V}_{\theta(g)}(Y_{ij} \mid X_{ij}) = \sigma^2 x_{ij}^g \\ &\mathscr{C}_{\theta(g)}(Y_{ij}, Y_{i'j'} \mid X_{ij}, X_{i'j'}) = 0, \end{aligned} \tag{1.1}$$

where the script letters $\mathscr{E}$, $\mathscr{V}$ and $\mathscr{C}$ denote the conditional expectation, variance and covariance given X_{ij}'s respectively. In this realistic model of practical interest, it is noted that g is non-negative and in most of the practical situations is found to lie between 1 and 2 (cf. Foreman

---

[1] On leave from Indian Statistical Institute, Calcutta.

and Brewer (1971). This has been borne out by the empirical studies of Smith (1938), Jessen (1942) and Mahalanobis (1944). The equations (1.1) are also known as a "super-population model".

The justification for the assumption mentioned above that the unknown proportionate values of σ_i^2's are usually not very different from the proportionate values of the known α_i^2's is examined in the light of *a priori* distributions specified by (1.1). This problem was first considered by Hanurav (1965) and Rao (1968) for special cases of (1.1).

Ericson (1965) studied the problem of optimum stratified sampling when prior information is taken to be expressible in the form of a multivariate Normal prior distribution. He had also discussed several methods of assessing prior distributions and presented computational algorithms for determining the optimum allocations. A good review of the work on the problem of optimum stratified sampling is given in Solomon and Zacks (1970) who discuss Bayes and minimax allocations as well.

In this review article we shall first study the case of Simple Random Sampling (SRS) in each stratum. We discuss Neyman's optimum allocation (for SRS) in the light of *a priori* distributions specified by (1.1). Next, we review the work (cf. Reddy, 1976) on the problem of optimum stratification with proportional and optimum allocations in the case of SRS in the light of *a priori* distributions specified above. During the last two decades, several attempts have been made to utilize auxiliary information available in constructing optimum sampling strategies. A host of research workers have considered varying probability sampling schemes and emphasis was laid on the construction of sampling schemes with probability of inclusion of a unit in the sample, π_i, Proportional to its Size measure (πPS schemes). In Section 3 we review the results on the utilization of prior distributions in the study of relative efficiencies of stratified πPS sampling strategies with various allocations and unstratified πPS sampling. Next, we note that the results for Probability Proportional to Size (PPS) with replacement sampling scheme follow from this.

2 Stratified Random Sampling

Neyman's Optimum Allocation and Prior Distributions

For Simple Random Sampling with replacement in each stratum, Neyman's optimum allocation (Neyman, 1934) is given by $n_{i,\text{opt.}} = nN_i\sigma_i/\sum N_i\sigma_i$, where

$$\sigma_i^2 = (\sum_j Y_{ij}^2 - Y_i^2/N_i)/N_i,\ Y_i = \sum_j Y_{ij}.$$

The corresponding allocation when SRS without replacement scheme is used in each stratum is given by $nN_iS_i/\sum N_iS_i$ where $S_i^2 = N_i\sigma_i^2/(N_i-1)$.

Under the prior distributions specified by (1.1) we have:

$$\begin{aligned}\mathscr{E}_{\theta(\varphi)}\sigma_i^2 &= [\sum_j \mathscr{E}(Y_{ij}^2) - \mathscr{E}(Y_i^2)/N_i]/N_i \\ &= [\sum_j \{(a+bX_{ij})^2 + \sigma^2 X_{ij}^g\} - \{\sigma^2 \sum_j X_{ij}^g + (\sum_j (a+bX_{ij}))^2\}]/N_i \\ &= [\sigma^2 \sum_j X_{ij}^g (1-N_i^{-1}) + b^2(\sum_j X_{ij}^2 - X_i^2/N_i)]/N_i\end{aligned}$$

where $X_i = \sum_j X_{ij}$

$$\simeq b^2\alpha_i^2 + \sigma^2 \sum_j X_{ij}^g/N_i,$$

assuming $(N_i-1)/N_i$'s are approximately equal to unity.

On the other hand

$$\mathscr{E}_{\theta(\theta)} S_i^2 = [\sigma^2 \sum_j X_{ij}^g (1 - N_i^{-1}) + b^2 (\sum_j X_{ij}^2 - X_i^2/N_i)]/(N_i - 1)$$
$$= b^2 \alpha_i'^2 + \sigma^2 \sum_j X_{ij}^g / N_i$$

where $\alpha_i'^2 = (\sum_j X_{ij}^2 - X_i^2/N_i)/(N_i - 1)$. Note that no assumption is needed here.

Thus σ_i^2's (S_i^2's) can be expected to be in the same proportion as α_i^2's ($\alpha_i'^2$'s) if α_i^2's ($\alpha_i'^2$'s) are proportional to $\sum_j X_{ij}^g / N_i$. In particular,

$$\sigma_i^2 \propto \alpha_i^2 \quad \text{if} \quad \alpha_i^2 \propto \sum X_{ij}^2 / N_i \text{ for } g = 2$$
$$\propto \bar{X}_i \text{ for } g = 1$$
$$\propto \text{constant for } g = 0.$$

Equivalently for $g = 2$, $\sigma_i^2 \propto \alpha_i^2$ when $\alpha_i^2 \propto \alpha_i^2 + \bar{X}_i^2$ or when the squares of coefficients of variation of the $\mathscr{X}$-characteristic are equal in all strata. In that case, Neyman's optimum allocation reduces to allocation proportional to $N_i \bar{X}_i = X_i$, the total of x-values in the ith stratum. For general g, the allocation is proportional to $(N_i \sum_j X_{ij}^g)^{\frac{1}{2}}$ when $\alpha_i^2 \propto \sum_j X_{ij}^g / N_i$. This condition can be re-written as

$$\alpha_i^2 \propto \alpha_i^2 + \bar{X}_i^2 + ((\sum X_{ij}^g - \sum X_{ij}^2)/N_i)$$

or

$$\alpha_i^2 \propto \alpha_i^2 + (\bar{X}_i^2 - \delta_i(g)/N_i^2)$$

where we set $\delta_i(g) = N_i (\sum X_{ij}^2 - \sum X_{ij}^g)$

or i.e. $\alpha_i^2 \propto \bar{X}_i^2 - \delta_i(g)/N_i^2$ (provided the r.h.s. > 0 for all i) and when this condition is satisfied, the allocation in the general case is equivalent to allocation proportional to $N_i(\bar{X}_i^2 - \delta_i(g)/N_i^2)^{\frac{1}{2}}$ or proportional to $(X_i^2 - \delta_i(g))^{\frac{1}{2}}$. Note that for $g = 2$, $\delta_i(g) = 0$ and the allocation reduces to allocation proportional to X_i. Elsewhere, the condition for the general case has been interpreted as the squares of the "corrected coefficients of variation" being equal in all strata and the results have been illustrated (Rao, 1968).

Optimum Demarcation of Strata and Prior Distributions

We shall now turn our attention to the problem of optimum stratification with proportional and optimum allocations in the case of simple random sampling in the light of an appropriate *a priori* distribution.

Given the number of strata, the equations for determining the optimum points of stratification under proportional and Neyman's optimum allocation have been worked out by Dalenius (1957) and other research workers followed this up with quicker and approximate methods. The rules for stratification were based on the assumption that stratification can be made on the values of the study variate $\mathscr{Y}$ itself. In practice, some auxiliary information on a variate $\mathscr{X}$ highly correlated with $\mathscr{Y}$ is available and can be used to construct the optimum points of stratification. Dalenius (1957) developed equations for the x-boundaries, assuming suitable regression models. Among others, Mahalanobis (1952), Dalenius and Hodges (1959), Ekman (1959) suggested approximations to theoretical solutions for use in practice, while Sethi (1963), Taga (1967), Singh and Sukhatme (1969) and others discussed further theoretical details. Some of the conjectures of Dalenius on stratification rules were supported by the investigations of Cochran (1961).

Reddy (1976) has considered the problem by assuming the prior distribution specified by (1.1) above and examined the optimum points of stratification which minimize the expected variance of the estimator of the total $\hat{Y}$ under (1.1). For Simple Random Sampling With Out Replacement (SRSWOR) of size n_i in each stratum, it is easy to obtain that

$$\mathscr{E}_{\theta(s)} \operatorname{var}(\hat{Y} \mid X) = \sum_{i=1}^{k} N_i^2 (n_i^{-1} - N_i^{-1})(b^2 \alpha_i'^2 + \sigma^2 \sum_j X_{ij}^g / N_i)$$

where

$$\alpha_i'^2 = \sum_j (X_{ij} - \bar{X}_i)^2 / (N_i - 1). \tag{2.1}$$

In the case of proportional allocation of sample size in each stratum, (2.1) reduces to

$$n^{-1}(N-n)\left(b^2 \sum_{i=1}^{k} N_i \alpha_i'^2 + \sigma^2 \sum_{i=1}^{k} \sum_{j=1}^{N_i} X_{ij}^g\right) \tag{2.2}$$

Reddy (1976) then defines S to be a "mis-stratification" (with respect to $\mathscr{X}$) if at least one of the strata contains an x-value which lies between the minimum and maximum of x in another stratum. Now let S^* be a stratification which minimizes (2.2). If there is a mis-stratification between any two strata, by proper interchange of units he then arrives at a contradiction to the assumption of S^* and thus proves

Theorem 2.1 (from Reddy, 1976). *For any stratification of the population with simple random sampling without replacement in each stratum and with proportional allocation of sample size, it is necessary that the $\mathscr{X}$-variate be arranged in increasing order of magnitude for $\mathscr{E}_{\theta(s)} \operatorname{var}(\hat{Y})$ to be minimum.*

When optimum allocation is used in each stratum, the arrangement of the auxiliary variate in increasing (or decreasing) order of magnitude is not necessary. However, for the special case $g = 2$, when the coefficients of variation (w.r.t. auxiliary variate) are equal in all strata for SRSWOR scheme in each stratum Reddy (1976) has established the necessity for arranging the $\mathscr{X}$-variate in ascending or descending order of magnitude.

3 Stratified πPS and PPS Sampling

In this section we review the results on the utilisation of prior distributions in the study of relative efficiencies of unstratified and stratified πPS and PPS sampling strategies with various allocations.

First, let a πPS (π_λ, the probability of inclusion of the λth unit in the sample, Proportional to its Size) sample of size n_i be selected from the ith stratum such that $\sum n_i = n$ where the $\mathscr{X}$ characteristic is used as the size measure. Let π_{ij} be the probability of inclusion of the jth unit of the ith stratum in the sample, $j = 1, 2, \ldots, N_i$; $i = 1, 2, \ldots, k$. As an estimator of the population total $Y = \sum\sum Y_{ij}$, consider the Horvitz–Thompson (1952) estimator

$$\hat{Y}_S = \sum_i \sum_j Y_{ij}/\pi_{ij} = \sum_i \sum_j Y_{ij}/(n_i X_{ij}/X_i),$$

where $\sum_j$ denotes summation over the sampled units in the ith stratum.

We now derive the allocation of sample size which minimizes the expected variance of $\hat{Y}_S$ under the class of prior distributions $\textcircled{H} = \{\theta\}$ for which

$$\begin{aligned} &\mathscr{E}_\theta (Y_{ij} \mid X_{ij}) = b X_{ij} \\ &\mathscr{V}_\theta (Y_{ij} \mid X_{ij}) = v(X_{ij}) \\ &\mathscr{C}_\theta (Y_{ij}, Y_{i'j'} \mid X_{ij}, X_{i'j'}) = 0. \end{aligned} \tag{3.1}$$

Under (3.1) we have

$$\mathscr{E} \operatorname{var}(\hat{Y}_S) = \mathscr{E}\left[\sum_{i=1}^{k} \left\{\sum_{j=1}^{n_i} \left(\frac{1}{\pi_{ij}} - 1\right) Y_{ij}^2 + \sum_{j \neq}^{N_i} \sum_{j'}^{N_i} \left(\frac{\pi_{ijj'}}{\pi_{ij}\pi_{ij'}} - 1\right) Y_{ij} Y_{ij'}\right\}\right],$$

where $\pi_{ijj'}$ is the probability of joint inclusion of the pair of jth and j'th units of the ith stratum in the sample. Thus

$$\mathscr{E}\operatorname{var}(\hat{Y}_S)=\sum_{i=1}^{k}\sum_{j=1}^{N_i}(\pi_{ij}^{-1}-1)\,v(X_{ij})+\operatorname{var}\left(\sum_i\sum_j\frac{x_{ij}}{\pi_{ij}}\right)$$

$$=\sum_{i=1}^{k}\sum_{j=1}^{N_i}(\pi_{ij}^{-1}-1)\,v(X_{ij})$$

the second term being zero because $\pi_{ij}\propto x_{ij}$.

Minimization of $\mathscr{E}\operatorname{var}(\hat{Y}_S)$ subject to the condition $\sum n_i=n$ would lead to the minimization of

$$\sum_{i=1}^{k}\sum_{j=1}^{N_i}(n_i^{-1}X_{ij}^{-1}X_i-1)+\lambda\left(\sum_{i=1}^{k}n_i-n\right),$$

where λ is the Lagrange multiplier. Now, differentiating w.r.t. n_i and equating to zero, we have after some simplification

$$n_{i,\,\mathrm{opt.}}=n\left(X_i\left(\sum_{j=1}^{N_i}v(X_{ij})/X_{ij}\right)\right)^{\frac{1}{2}}\Big/\sum_{i=1}^{k}\left(X_i\left(\sum_{j=1}^{N_i}v(X_{ij})/X_{ij}\right)\right)^{\frac{1}{2}}$$

We call this the θ-optimum allocation. For this θ-optimum allocation we have a generalization of Rao's (1968) theorem the proof of which follows on the same lines as in Rao (1968):

Theorem 3.1. *In the sense of expected variance under the above model, unstratified* πPS *sampling strategy* (*with Horvitz–Thompson* (HT) *estimator*) *is inferior to stratified* πPS *sampling strategy* (*with the corresponding* HT *estimator*) *with* θ*-optimum allocation.*

Proof. For the HT estimator of the population total in the case of unstratified πPS sampling, we have

$$\mathscr{E}_\theta\operatorname{var}(\hat{Y}_U)=\sum_i\sum_j((\pi'_{ij})^{-1}-1)\,v(X_{ij})$$

where $\pi'_{ij}=nX_{ij}/X$. Further, we have

$$\mathscr{E}_\theta\operatorname{var}(\hat{Y}_S)=\sum_i\sum_j(\pi_{ij}^{-1}-1)\,v(X_{ij})$$

where $\pi_{ij}=n_{i,\,\mathrm{opt.}}X_{ij}/X_i$

Therefore

$$\mathscr{E}_\theta\{\operatorname{var}(\hat{Y}_U)-\operatorname{var}(\hat{Y}_S)\}$$

$$=\frac{1}{n}\left[\sum_i\sum_j\frac{v(X_{ij})}{X_{ij}}\left(X-\frac{(\sum_i(X_i\sum_j v(X_{ij})/X_{ij})^{\frac{1}{2}})\,X_i^{\frac{1}{2}}}{(\sum_j v(X_{ij})/X_{ij})^{\frac{1}{2}}}\right)\right]$$

$$=\frac{1}{n}\left[\sum_i(\sum_j v(X_{ij})/X_{ij})\sum_i X_i-(\sum_i(X_i\sum_j v(X_{ij})/X_{ij})^{\frac{1}{2}})^2\right]$$

$$\geq 0,\ \text{with equality if } X_i\propto\sum_j v(X_{ij})/X_{ij}.$$

Vijayan (1971) also considered this and studied the condition further for special cases of $v(X_{ij})$.

It is, however, not known from this under what conditions unstratified πPS sampling is still inferior to stratified πPS sampling when one deviates from the θ-optimum allocation. With this aim, Ramachandran and Rao (1974a) investigated whether stratified πPS sampling with various non-optimal allocations is likely to be worthwhile and whether it should at all be attempted in practice. They considered deviations from the $\theta(g)$-optimum allocation (i.e.

θ-optimum allocation with $v(X_{ij}) \propto X_{ij}^{g}$) which is realistic and of practical interest. The mathematical content of this result is given in

Theorem 3.2 (Ramachandran and Rao, 1974a): *Let* $0 < p_{11} \leqq p_{12} \leqq \ldots \leqq p_{kN_k}$ *and allocation* $\mathbf{n} = (n_1, n_2, \ldots, n_k)$ *be such that* $a_i = n_i^{-1} p_i - n^{-1}$, $i = 1, 2, \ldots, k$ *are non-decreasing and not all equal where* $p_{ij} = X_{ij}/X$ *and* $P_i = \Sigma p_{ij}$. *Further, let* $a_i \leqq 0$ *for* $i \leqq t$ *and* $a_i > 0$ *for* $i > t$ *and not all* p_{ij}*'s for* $i > t$ *are equal to* p_{iN_i}. *Let* $f(g) = \mathscr{E}_{\theta(g)}\{V(\hat{Y}_S) - V(\hat{Y}_U)\}/\sigma^2 = \sum\sum a_i p_{ij}^{g-1}$. *Then (a) if* $f(1) < 0$, *there exists a unique* g_0 *in* $(1, 2]$ *such that* $f(g) \leqq 0$ *or* > 0 *according as* $g \leqq g_0$ *or* $> g_0$; *(b) if* $f(1) \geqq 0$, $f(g) > 0$ *for all* g *in* $(1, 2]$.

This essentially means that for a practical situation when the allocation is such that X_i/n_i (or a_i) are non-decreasing and not all equal, whenever $\sum N_i X_i n_i^{-1} - NXn^{-1} < 0$ (which is a simple condition to verify in practice) stratified πPS sampling with the given allocation is better than unstratified πPS sampling of size n up to a certain value g_0 in $(1, 2]$ and beyond that it is worse. Further, when $\sum N_i X_i n_i^{-1} - NXn^{-1} \geqq 0$, it follows from (*b*) that, stratified πPS sampling is worse than unstratified πPS sampling and hence is not recommended. A number of special cases and comparisons between various types of allocations are considered in Ramachandran and Rao (1974a).

Next consider the selection of units from each stratum with Probability Proportional to Size (PPS) with replacement where the auxiliary information $\mathscr{X}$ is used as the size measure. As an estimator of the population total consider $\hat{Y}_S = \sum_i (\sum_j y_{ij}/x_{ij}) X_i/n_i$. For the class of distributions specified by (3.1) Ramachandran and Rao (1974b) obtained the allocation of sample size which minimizes the expected variances of $\hat{Y}_S$ given by

$$n_{i,\text{opt.}} = n \left(\sum_j v(X_{ij})(X_i X_{ij}^{-1} - 1)\right)^{\frac{1}{2}} \Big/ \sum_i \left(\sum_j v(X_{ij})(X_i X_{ij}^{-1} - 1)\right)^{\frac{1}{2}}.$$

It is observed that under Ⓗ, unstratified PPS sampling is inferior to stratified PPS sampling with this allocation. Earlier, Raj (1963) has compared unstratified PPS sampling with stratified PPS sampling when X-proportional allocation is used in terms of exact variance. Observing that this comparison is between unstratified PPS sampling and stratified PPS sampling with a non-optimum allocation, Ramachandran and Rao (1974b) studied whether stratified PPS sampling with various non-optimal allocations is likely to be useful. Here we have $\mathscr{E}_{\theta(g)}\{\text{var}(\hat{Y}_S)\text{-var}(\hat{Y}_U)\}/\sigma^2 \leqq f(g)$ and it is immediate from Theorem 3.1 that whenever $f(1) < 0$, there exists a value g_0 such that stratified PPS sampling is better than unstratified PPS sampling for values of g *at least up to* this g_0. When $f(1) > 0$, even though $f(g) > 0$ for all g in $(1, 2]$ stratified PPS sampling might still be better than unstratified PPS sampling for values of g close to unity.

4 Conclusion

Neyman's optimum allocation for Simple Random Sampling within each stratum requires the values of stratum variances σ_i^2 which are generally unknown. In practice, rough estimates of σ_i^2 based on pilot studies or some auxiliary information are substituted for σ_i^2 and no theoretical justification is available for this. Whenever auxiliary information related to the study variate is available, we can express this in the form of a prior distribution and the allocations can be studied under this distribution, thus providing a justification for the near-optimum allocations. This was first recognized by Hanurav (1965) and Rao (1968) and later on extended to detailed studies on the utilization of prior distributions in various types of allocation and sampling strategies. The aim of this paper has been to review these specific studies and put them together which are scattered in the literature. The role played by prior distributions in studying the optimum allocation of sample size, relative efficiency of sampling strategies and related problems is now clear from these studies. It would be interesting to study optimum allocations further under similar set up when multiple auxiliary information is available.

References

Cochran, W.G. (1946). Relative accuracy of systematic and stratified random samples for a certain class of populations. *Ann. Math. Statist.*, **17**, 164–177.

Cochran, W.G. (1961). Comparison of methods for determining stratum boundaries. *Bull. Int. Stat. Inst.*, **38** (2), 345–358.

Dalenius, T. (1957). *Sampling in Sweden.* Contribution to the methods and theories of sample survey practice. Stockholm: Almqvist and Wicksell.

Dalenius, T. and Hodges, J.L. Jr. (1959). Minimum variance stratification. *J. Amer. Statist. Assoc.*, **54**, 88–101.

Ekman, G. (1959). An approximation useful in univariate stratification. *Ann. Math. Statist.*, **30**, 219–229.

Ericson, W.A. (1965). Optimum stratified sampling using prior information. *J. Am. Statist. Assoc.*, **60**, 750–771.

Foreman, E.K. and Brewer, K.R.W. (1971). The use of supplementary information in standard sampling procedures. *J. Roy. Statist. Soc. B*, **33**, 391–400.

Hanurav, T.V. (1965). *Optimum sampling strategies and some related problems.* Ph.D. Thesis submitted to the Indian Statistical Institute.

Horvitz, D.G. and Thompson, D.J. (1952). A generalization of sampling without replacement from a finite universe. *J. Amer. Statist. Assoc.*, **47**, 663–685.

Jessen, R.J. (1942). Statistical investigation of a sample survey for obtaining farm facts. *Iowa Agr. Exp. Sta. Res. Bull.*, **304**.

Mahalanobis, P.C. (1944). On large scale sample surveys. *Royal Statist. Soc. Phil. Trans. B*, **231**, 329–451.

Mahalanobis, P.C. (1952). Some aspects of the design of sample surveys. *Sankhyā*, **12**, 1–7.

Neyman, J. (1934). On two different aspects of the representative method, the method of stratified sampling and the method of purposive selection. *J. Royal Statist. Soc.*, **97**, 558–606.

Raj, D. (1963). A note on stratification in unequal probability sampling. *Sankhyā*, **25B**, 215–216.

Ramachandran, G. and Rao, T.J. (1974a). Allocation to strata and relative efficiencies of stratified and unstratified πPS sampling schemes. *J. Royal Statist. Soc. B*, **36**, 292–298.

Ramachandran, G. and Rao, T.J. (1974b). Probability proportional to size sampling and stratification. Paper presented at the International Symposium on Recent Trends of Research in Statistics, Calcutta, December 1974.

Rao, T.J. (1968). On the allocation of sample size in stratified sampling. *Ann. Inst. Stat. Math.*, **20**, 159–166.

Reddy, V.N. (1976). Stratified simple random sampling and prior distributions. *Ann. Inst. Stat. Math.*, **28**, 445–459.

Sethi, V.K. (1963). A note on optimum stratification of populations for estimating the population means. *Aust. J. Statist.*, **5**, 20–35.

Singh, R. and Sukhatme, B.V. (1969). Optimum stratification. *Ann. Inst. Stat. Math.*, **21**, 515–528.

Smith, H.F. (1938). An empirical law governing soil heterogeneity, *J. Agric. Sci.*, **28**, 1–23.

Solomon, H. and Zacks, S. (1970). Optimal design of sampling from finite populations. *J. Amer. Statist. Assoc.*, **65**, 653–677.

Taga, Y. (1967). On optimum stratification for the objective variable based on concomitant variables using prior information. *Ann. Inst. Stat. Math.*, **19**, 101–129.

Vijayan, K. (1971). Personal communication.

Résumé

Dans le présent article, on etudie le problème de la répartition optimale d'un échantillon entre diverses strates, à la lumière de distributions a priori. Tout d'abord, on considère la répartition optimale de NEYMAN, dans le cas d'un plan de sondage simplement au hasard (dit SRS); et on passe en revue les résultats pouvant justifier ou non l'hypothèse selon laquelle les variances inconnues O_i^2 à l'intérieur des strates pourraient etre remplacées proportionnellement par des valeurs connues α_i^2 correspondant aux variances de l information auxiliaire dans les dites strates. On revoit ensuite le problème de la stratification optimale avec répartition de l'échantillon soit proportionnelle, soit optimale, dans le cas dit SRS, à la lumière de distributions a priori; on discute le cas ou la stratification est faite suivant les valeurs croissantes de la variable auxiliaire. On s interesse enfin au cas de plans de sondage à probabilités variables; et on passe en revue les résultats d'une étude en vue de savoir si, en pratique, il serait vraisemblable qu une stratégie combinant stratification et tirages avec probabilités inégales fût payante,—ceci pour diverses répartitions non-optimales de l échantillon.