Design and Estimation for Domains

LESLIE KISH

Institute for Social Research, University of Michigan, Michigan, USA

---

1. Definitions and Classifications

Domains and subclasses are terms used similarly in the literature of survey sampling to denote divisions, usually partitions, both of the population and of the sample for separate estimation. Let us take advantage of that redundancy: let domains denote subpopulations, and let subclasses refer to their reflected subdivisions in the sample.

Estimates are required for a diversity of domains and the *types* of domains should influence the choices both of design and of estimation. A classification of three types of domains is therefore desirable, and I suggest the following terminology for them.

(a) There are *design domains* for which separate samples have been planned, designed and selected; their combination forms the entire sample, usually as a weighted sum of independent samples. For example: major regions or urban and rural domains, where these are composed of entire strata of primary sampling units; also geographical or alphabetical divisions when names are so selected individually from appropriately ordered listings.

(b) At the other extreme are *cross-classes* which cut across the sample designs, across strata, and across sampling units. These are the most commonly used kinds of domains and subclasses; e.g. age, sex, occupation, education and income classes, behaviour and attitude types, etc. They have not been separated into design domains because information was not available on these variables, or because they seemed less important than others, or because they were forgotten.

(c) Between the two extremes, but less commonly used than the two dominant types, are *mixed domains* of diverse kinds; they have not been separated by the design, but they tend to concentrate unevenly in the sampling units or in strata. For example: occupations such as fishermen, farming specialties, miners and lumberjacks, which are segregated by natural forces; or ethnic groups segregated by social forces; in both cases

the segregation may be prevalant but neither complete nor available as auxiliary data for design.

The *sizes* of domains also influence the choice of methods for design and estimation; hence a cross-classification of the above types with classes based on sizes of domains also seems useful. This classification is stated roughly to orders of magnitudes, with descriptive names assigned for ready reference.

(1) *Major domains* comprise perhaps $\frac{1}{10}$ of the population or more. Examples: major regions for design domains, and 10-year age groups or major occupational categories for cross-classes. For major domains reasonable estimates can be produced from probability samples with standard methods and essentially without bias; but variances are increased for cross-classes, and their estimation requires special attention, as in sections 2 and 3. Furthermore designing adequate sample sizes for them may lead to conflicts (section 7).

(2) *Minor domains* comprise perhaps from $\frac{1}{10}$ to $\frac{1}{100}$ of the population. Examples: populations of the 50 states of the USA or the 63 counties in the UK, or single years of age or two-fold classifications of major domains; like occupation by education; or regions (designed) by education (cross-class). (See Table I at end of article.)

(3) *Mini-domains* comprise perhaps from $\frac{1}{100}$ to $\frac{1}{1000}$ or even to $\frac{1}{10000}$ of the population. Examples: populations of the over 3 000 counties of the USA, or the 625 parliamentary districts in the UK, or a three-fold classification of age by occupation by education. For mini-domains usually (and often for minor domains also) the sample bases are too small for any usable reliability, hence standard methods of estimation are inadequate. New methods are needed and these are mentioned in section 8, and discussed by Purcell and Kish (1980).

Rare types, comprising less than $\frac{1}{10000}$ in the population are problems for which samples of an entire population are useless, and separate lists and methods are needed (Kish, 1965, section 11.4).

Domain means $\bar{y}_c = y_c/x_c$, the means for subclasses (c), are estimates most frequently used and most developed, and our discussion focuses on them, although some of it is relevant for other estimates. *Differences* (comparisons) *of pairs of subclass means* $\bar{y}_c - \bar{y}_b = y_c/x_c - y_b/x_b$ also represent most common uses of data from sample surveys and are noted. The *denominators* x_c, x_b of subclasses used in practice have several common characteristics:

(a) They generally represent count variables of the sample sizes of subclasses: 0, 1 variables denoting (non) membership in a specified subclass (domain). They may be unweighted n's or weighted counts. They may also

210

be other variables at times (e.g. live births per total births of women), and most analyses can deal with these also, perhaps with some modification.

(b) They denote two partitions (non-overlapping) based on two categories of the same variable. This feature can be used for simplifying some computing programs. If, on the contrary, they are overlapping categories, or one includes the other, or if they are subclasses from different variables, special care, notation and variance computations are needed and feasible.

(c) The pair of subclasses may be exhaustive partitions ($x_c + x_b = x$ for the entire sample), but often they are only two of k partitions, hence the pair represents two out of $k(k-1)/2$ possible pairs.

(d) Usually they are random variables, especially when they are cross-classes, even if the total sample size were fixed. This feature has consequences for the variances.

The *numerators* y_c, y_b represent the same variable for the pair. For proportions (percentages) they are analogous subsets of the count variables in the denominators. These features may be used for simpler computing.

The above represent what are commonly understood by domain and subclass means and differences (comparisons). It is helpful not to confuse them with other statistics, as has been done in some non-practical writing: (1) A proportion of x_c/x of the entire sample is not a subclass (domain); "shares" is a good name for these. (2) A difference $(x_c - x_b)/x$ based on the entire sample is not a comparison of domains, but a difference of shares. (3) Nobody uses comparisons $y_c/x_c - y_b/z_b$ based on different variables, hence they need no name. (4) Domains are not synonymous with strata, but they are often confused. Admittedly, it is useful to use strata to create design domains when we can; but usually strata can be made much more numerous than design domains. For example, we may use two selections per stratum, but would need perhaps ten or more of those strata to create stable bases for domains. Thus strata may be more numerous than design domains by orders of magnitude. Cross-classes are numerous but they cut across strata (and across each other).

2. Overview of Domain Effects

Sampling literature, theory and methods on design and on analysis all deal principally with entire samples selected to represent some specified population. Within that setting we take a brief overview of diverse effects that arise as subclasses of the sample are used for inferences about corresponding domains of the population. Our principal interest is in major cross-classes because they are of principal use and because these broad remarks are most pertinent to them. Selecting subclass members from a sample has the effect of assigning zero values to all variables of non-members. Hence

the effects are similar to non-members appearing as blanks in the frame during the selection process. More details are given and indicated in later sections, together with some justification and references for the sweeping generalizations boldly presented here.

A. *Selection probabilities are preserved for individual elements in subclasses.* The probabilities P_i assigned to member elements are unaffected by the zero values assigned to all variables of non-members in creating subclasses. This should be obvious but needs stating. Assigned weights proportional to $1/P_i$ are also used for the subclasses.

B. *Sample sizes become highly variable for cross-classes.* Zero for all non-members have the same effect as blanks in the selection process. These will be noted for stratum sizes (3) and for cluster sizes (4). Size variations also arise for the entire sample from imprecise measures of size and from non-responses; but as cross-classes become smaller the variability increases greatly as size controls become lost.

C. *Estimates of means and totals retain their forms for subclasses.* The unbiased nature of simple totals $\Sigma y_j/P_j$ is retained with the undisturbed selection probability (A). The ratio means retain their sturdy consistency until the variability of size (B) in the denominator becomes too high for small cross-classes (section 5).

D. *Variances of means and totals become greatly affected with decrease of size in subclasses.* Here we attempt a summary of separate treatments for stratified element sampling (section 3) and cluster sampling (section 4). For design domains the effects arise from the decreased sizes, especially in cluster sampling where the precision of variance estimates depends on numbers of primary selections.

For cross-classes drastic effects, due to loss of controls represented by the sample design, arise in three distinct ways. (a) The variances depend on pairwise joint probabilities of selection of elements. The sample designs represent controls over those joint probabilities, and in cross-classes those controls tend to be lost. Hence the design effects approach 1, and variances for self-weighting means of cross-classes approach those of simple random sampling, that is s^2/n. The approaches are from below 1 as the gains of proportionate stratified sampling are lost (section 3), and from above 1 as the losses of cluster sampling are regained (section 4). (b) For differences of means of cross-classes the approach toward deft $=1$ is more rapid than for the means alone; this is true for the deft > 1 in cluster sampling, and in proportionate sampling deft $=1$ is virtually attained in practice. (c) Some of those losses from the gains of stratification can be regained with post-stratification and ratio estimation; but the auxiliary data are often not available at the detail needed for small domains (sections 3 and 8). (d) The

controls introduced with measures of size in PPS selection tend to be weakened to the extent that those measures have less relevance for the cross-classes. This loss occurs first to the extent that members and non-members of cross-classes are randomly distributed in clusters. But the loss is worse to the degree that the clustering of the domains differ from that of the population. (e) Optimal allocation for the entire population can be very different from those for diverse domains. For many cross-classes it may be close enough, but for many other domains it may be far from optimal and even counterproductive (Kish, 1961, 1965, 4.5C)

E. *Computing estimates for totals, means and their variances can be done by assigning zero values to all non-members of the subclass,* with needed modifications for sample size. However, more convenient and economic procedures can also be used usually, omitting the non-members from computations.

F. *Designs for domains induce conflicts with the entire sample and with each other.* Conflicts arise in allocating sample sizes and rates to provide adequate sample bases for diverse design domains. This can be further complicated by considerations of optimal allocations to strata. Furthermore diverse sampling rates for cross-classes can require some screening procedure that may be expensive. Thus reducing variances for some domains can often increase variances for the entire sample and for most other domains (section 7).

G. *Estimates for small domains require special techniques* (section 8). These can improve estimates for minor domains; and they may facilitate estimates for mini-domains which would not be feasible at all from the sample alone. They may be used even for major domains, although their estimates can also be made from the sample alone.

3. Stratified Sampling of Elements

In this section we assume simple random selection (s.r.s.) of elements within strata. Estimates for design domains present no new problems because we use the usual formulae of stratified element sampling, provided that the domains' sample sizes n_c are large enough for stable estimates. When the domain contains a single stratum we are back to simple random sampling.

However, domains are often cross-classes and the presence of non-members as blank selections loosens the control of stratification over sample sizes. In the hth stratum s.r.s. selection of n_h from N_h elements with $f_h = n_h/N_h$ yields a sample size m_{hc}; this number is a dichotomous variable with expectation $f_h M_{ch}$ and variance $(1-f_h)n_h \bar{M}_{ch}(1-\bar{M}_{ch})$; here M_{ch} is the size and $\bar{M}_{ch} = M_{ch}/N_h$ the proportion of domain members. The

effects of this variation in increased variances of cross-class means, totals and differences we note below. They have been first described as "domains that cut across strata" by Yates (1953, sections 9.1–9.5) then by others (Cochran, 1963, 1977; Kish, 1961, 1965).

The increases in the cross-class variances arise essentially because of the factors $F_h = 1/f_h = N_h/n_h$ in the estimates, based on the entire population, instead of M_{ch}/m_{ch} for the subclass. This amounts to using for stratum weights $w_{ch} = F_h m_{ch}/\Sigma F_h m_{ch}$ instead of $W_{ch} = M_{ch}/\Sigma M_{ch}$. When W_{ch} values can be found and used for post-stratification (or ratio estimation, section 8), the increases in variances noted below can be avoided.

The most drastic effect is on the variance of the *simple* estimate of the cross-class aggregate:

$$\hat{Y}_c = \Sigma_h y_{ch} F_h$$

with

$$\text{var}(\hat{Y}_c) = \Sigma(1-f_h)F^2{}_h\left(\frac{n_h}{n_h-1}\right)m_{ch}[v^2{}_{ch} + (1-\bar{m}_{ch})\bar{y}_c{}^2{}_h] \qquad (3.1)$$

Here $\bar{m}_{ch} = m_{ch}/n_h$ the cross-class proportion and $\bar{y}_{ch} = \Sigma y_{chi}/m_{ch}$ the sample mean in the sample from the hth stratum. The element variance is denoted by

$$v^2{}_{ch} = \frac{1}{m_{ch}}\sum \hat{m}_{ch}(y_{chi} - \bar{y}_{ch})^2 = \frac{m_{ch}-1}{m_{ch}}s_{ch}{}^2 \qquad (3.2)$$

In the discussion below the distinctions between $s_{ch}{}^2$ and $v_{ch}{}^2$ and the factors $n_h/(n_h-1)$ and $(m_{ch}-1)/m_{ch}$ are neglected in order to concentrate on main issues. From the given estimates of variances the population variances can be constructed by substituting population parameters for sample statistics. The references above and others quoted in them give the formulae and derivations. These derivations depend on using zeros for all values for the $(n_h - m_{ch})$ non-members of cross-class c. Computing formulae can also be based on this feature. But it is usually more convenient to omit non-members and is meaningful to do so in the discussions below.

We must note in the brackets of (3.1) that the element relvariance $v_{ch}{}^2/\bar{y}_{ch}{}^2$ is increased by $(1-\bar{m}_{ch})$. This increase approaches unity for small cross-classes and would be a drastic increase, but it is usually avoided in practice by using other estimators, such as $\Sigma M_{ch}\bar{y}_{ch}$ when the M_{ch} are known.

214

The estimates for the mean of a cross-class and for its variance are denoted by

$$\bar{y}_c = \frac{\Sigma_h y_{hc}}{\Sigma_h m_{ch}}$$

and

$$\text{var}(\bar{y}_c) = \Sigma_h (1-f_h) \frac{w_{ch}^2}{m_{ch}} [v_{ch}^2 + (1-\bar{m}_{ch})(\bar{y}_{ch} - \bar{y}_c)^2] \qquad (3.3)$$

$$= \Sigma_h (1-f_h) \frac{w_{ch}^2}{m_{ch}} [t_{ch}^2 - \bar{m}_{ch}(\bar{y}_{ch} - \bar{y}_c)^2] \qquad (3.4)$$

Here $t_{ch}^2 = v_{ch}^2 + (\bar{y}_{ch} - \bar{y}_c)^2 = \Sigma m_{ch} (y_{chi} - \bar{y}_c)^2/m_{ch}$, the element variance containing both the within and between stratum components, thus computed as if without stratification. Thus for small cross-classes ($\bar{m}_{ch}$) the element variances are increased by most of the gains from stratification (3.3), and tend to approach the variances without stratification (3.4). Hence *the gains due to proportionate stratification tend to vanish for small subclasses.*

This tendency is even stronger for differences of means of cross-classes, where a good approximation for most empirical situations can be given by:

$$\text{var}(\bar{y}_c - \bar{y}_b) = \Sigma_h (1-f_h) \left[\frac{w_{ch}^2 t_{ch}^2}{m_{ch}} + \frac{w_{bh}^2 t_{bh}^2}{m_{bh}} \right] \qquad (3.5)$$

For *proportionate sampling the variances for differences of cross-class means tends to become as if unstratified s.r.s.* Furthermore, this also happens to ordinary $k \times m$ chi-square tests which can be regarded as combinations of such difference. (See Kish, 1961, 1965, section 4.5, 1969; Kish and Frankel, 1974.)

In the special case when each domain is confined to a single stratum, the situation for differences (3.5) and for means (3.3) reverts essentially to familiar formulae for s.r.s., because with $\bar{y}_{ch} = \bar{y}_c$ the second terms in the brackets disappear. But the increased variances for cross-classes for the simple estimate for totals (3.1) remain.

Our attention has been focused on cross-classes from proportionate sampling. What happens with optimal allocation to strata? For design domains separate optimal allocations may sometimes be feasible. But for cross-classes the situation is usually laden with confusing conflicts of multipurpose design (section 7). The effects of optimal allocation for the total sample can have opposite effects on different types of cross-classes (Kish, 1961, 1969). For some domains the gains are "inherited" in the cross-classes; perhaps even enhanced for some domains; for example

allocations by economic ratings may be inherited for age cross-classes, and even enhanced for cross-classes based on education and occupation. But for other types of domains losses may often result (Kish, 1961, 1969).

4. Cluster Sampling

The effects of domains in cluster sampling are drastic, but often neglected in the literature, in the designs of actual samples and in computing sampling errors. Yet we can limit ourselves to a brief discussion here and rely on some recent discussions for details and data (Kish, 1969; Kish and Frankel, 1974; Kish *et al.*, 1976; Verma *et al.*, 1980). We concentrate on how the effects of clustering on the variances of domain means and of differences of means differ from the effects on the variances of means of the entire sample. We know that typically the homogeneity of survey variables in sample clusters induce increases in their variances compared to those that an *s.r.s.* of the same size (n) would yield; this increase can be estimated by the *design effects*: as the ratios $deft^2$ of actual variance/s.r.s. variance; and the latter is s^2/n for means.

The sample variances depend on the nature of the variables and on their distributions in the population. These differ greatly and (contrary to widespread superstition) very large differences – ten- and hundred-fold – have been found frequently in the increases of variances for different variables from the same data base. These differences urge the computation of variances and $deft^2$ for most of the important survey variables. The sample variances also depend on the nature of the sample design: on the kind, size, numbers and stratification of sampling units used in the various stages of sampling, especially on the sizes of sample clusters achieved. In addition the variances often contain increases due to inverviewer variances, when effects are clustered in the sampling units.

Large-scale computation of variances must depend on simplifications, and commonly the overall variances are computed from "primary selections", which have also been called "ultimate clusters". These consist of the sample elements in units selected in the primary stage of selection; these are usually highly stratified, and computations are often based on "paired selections" from strata.

Design domains, such as regions, contain primary selections in separated strata, hence separate computations for each are feasible. They cause problems because they may be too numerous for separate computations and presentations; also because the number of primary selections, hence degrees of freedom, become small and this paucity results in unstable estimates of variances. For these reasons it may be better to use pooled estimates of variance for regions if these are somewhat similar in com-

216

position. However some domains – for example, the rural, the urban and metropolitan domains – may represent very different amounts of clustering, due to divers population distributions and to different selection methods, hence separate computations may be necessary. Differences between means of design domains, since they represent independent samples without covariances, should have variances that are the sums of the pairs of variances.

Cross-classes, which are most frequent, tend to cut across all strata and all clusters, hence to reduce the effects of clustering in drastic and fairly predictable manner. Creating cross-classes produces two principal effects on the sizes of sample clusters: it reduces their average sizes and it increases their relative variability. If the average sample size in a primary selections is $\bar{b} = n/a$, the average sample size for the cross-class is $\bar{b}_c = m/a = \bar{M}_c\bar{b}$, where $\bar{M}_c = M_c/N$, the proportion of the subclass among the N population elements. However, there exists a fair amount of variation in the sample sizes even for cross-classes like age, and sex and other demographic variables. For socio-economic cross-classes, like income and occupation, the variability is even greater, because of their partially clustered distributions. The controls exercised over subsample sizes (though stratification, PPS, etc.) tend to break down, less or more, for cross-classes.

The important effects of cross-classes on the design effects have been given elsewhere with details and data, but they should be summarized briefly here. Denote the design effects in the entire sample generally as $\text{deft}^2 = [1 + \text{roh}(\bar{b} - 1)]$ for specified survey variables, and these will vary greatly between variables. The synthetic *roh* here summarizes the effects of several stages of clustering (with stratification, etc.) and averages them over the diverse parts of the sample. The design effects in the cross-classes can be denoted as $\text{deft}_c^2 = [1 + \text{roh}_c(\bar{b}_c - 1)]$: and these tend to

$$\text{deft}_c^2 = [1 + \text{roh}(\bar{b}_c - 1)] \tag{4.1}$$

That is the *roh$_c$* for the cross-classes tend to be similar to *roh* for the entire sample for the same survey variable, hence deft^2 *tends to be reduced as the crossclass proportion $\bar{M}_c$ decreases*. There tends to be a slight increase for small cross-classes and slightly more for those that are less truly cross-classes, such as socio-economic classes. But these increases pale to insignificance compared to differences of *roh* between survey variables, to decreases in $\bar{b}_c$ compared to $\bar{b}$, and to differences between diverse sizes of $\bar{b}_c$; also often relative to sampling fluctuations of computed values of deft$_c^2$. The relationships seem to stand up best fortunately when most needed: for larger values of deft$_c^2$, due to larger values of *roh* and of $\bar{b}_c$. They should not be used for values of $\bar{b}_c$ near 1, when they are less needed. In a word:

roh is much more "portable" for different cross-classes of a survey variable than deft², because it excludes the variations in $\bar{b}_c$.

For differences of pairs of cross-class means also a useful empirical rule has been found in numerous computations from large varieties of data:

$$\frac{s_c{}^2}{n_c} + \frac{s_b{}^2}{n_b} < \text{var}(\bar{y}_c - \bar{y}_b) < \text{var}(\bar{y}_c) + \text{var}(\bar{y}_b) \tag{4.2}$$

Covariances $2\text{cov}(\bar{y}_c\bar{y}_b)$ reduce the effects of the cross-class variances, but the variance remains greater than the sum of s.r.s. variances for the means. Results, though subject to sampling fluctuations, fall generally between the two extremes (which are often assumed naively), more often nearer the lower than the higher extreme. The covariances represent welcome and important reductions of the effects of clusters in sample designs.

5. Computing Sampling Errors

The importance of domain means and of their differences for survey analysis complicates greatly the task of computing and even more the challenge of presenting sampling errors. Each sample typically obtains data on many survey variables. Presenting errors on all variables for all relevant domains would involve typically a two-dimensional complexity, and another dimension would be added for differences between domain means. Those who have attempted to deal with these complexities have used several simplifications for computing and generalizations for presenting the results (Kish *et al.*, 1976, Verma *et al.*, 1980). These are the motives for the generalizations in the other sections and we shall use them for the discussions below. These generalizations, like models usually, involve risks, but without some such brave effort sampling errors are bound to suffer from continuation of the prevailing neglect.

For proportionate stratified element sampling it may suffice to compute variances for all important variables for the entire sample. Seldom are the gains of proportionate sampling large; but if the gains are large for some variables (as shown by deft values well below 1) it may be strategically wise to add computations for critical design domains. However for means of cross-classes the generalizations can probably provide adequate guidance without further computations. For differences of cross-class means the approach to s.r.s. (deft² = 1) can probably be accepted. For disproportionate stratified sampling (optimal allocation) the situation is too complex for brief comments here.

Cluster sampling deserves a more careful treatment, but for simplicity assume self-weighting samples based on epsem selection of elements. First we emphasize computing variance and deft values for the entire sample for means of most of the important and critical variables because

of great variability of those values between variables. Ten-fold and even higher variations in the values of deft^2 are not uncommon for large clusters. These occur because roh values often range for different variables in the same sample from zero to 0·2 and sample clusters of 100 would yield deft$^2 = 20 = [1 + 0·2(100–1)]$; with sample clusters averaging 10 that variation would be three-fold. Thus (deft$^2 - 1 +$ roh) can vary in proportion to the cluster size.

For design domains, especially those – like urban and rural domains – which may have distinct values of deft, the variances may need separate computations. Here especially we need the caution of computing CV(x), the coefficient of variation for the denominators of ratio means. This should be done routinely for variances of all ratio means; values over 0·1 or 0·2 signal danger for the approximations used for ratio means and for their variances.

Cross-classes are frequently used and important. We have the model deft$_c^2 = [1 + roh_c(\bar{b}_c - 1)]$ to deal with them; but caution dictates checking the approximation roh$_c = $roh, especially for variables with larger values of roh and deft, where departures could cause greater differences in values of deft$_c^2$.

For differences of cross-class means the variance computations can be combined into the same program with those for the means themselves. Here the models are subject to wider ranges and doubts, hence there is more need for empirical data and verification. It will be useful to publicize more methodological work on sampling errors for domain means and for their differences.

Sampling errors for analytical statistics from complex samples is an area in great need of further development, but too vast and complicated for a brief treatment here, beyond this mention of its relation to differences between pairs of cross-class means (Kish and Frankel, 1974).

6. Domain Considerations in Design

Treatments of design in the sampling literature concentrate almost wholly on the entire sample, whereas surveys often have domain statistics as principal purposes. Hence some explicit remarks are needed here, although they will remain brief and general, and although they appear implicitly in the preceding sections and in the next section.

Sample size is often dominated implicitly by domain considerations. Remarks have often been made that sample sizes are frequently too large, and expensive, that sampling precision is too good, compared both to large unknown biases of observation (response and non-response) and to the needs of the user. But the actual fate of surveys is to appear too small, too

imprecise when it comes to analyses of domain means and of their differences.

In addition to numbers of elements, the numbers of selected units, especially numbers of primary selections, can become critical for design domains in clustered samples. For better estimates for design domains the sample may need more sampling units, more spreading, smaller sample clusters(b) than is warranted by the entire sample. On the contrary, cross-class means justify smaller clusters. Design domains should also affect the choice of stratification. They are likely do so implicitly in frequent practice, but explicit planning may help. It is useful and efficient to relate strata to domains, although they need not and cannot coincide always.

In proportionate element sampling cross-classes raise no new problems. We know that the usually modest gains of proportionate sampling tend to diminish for cross-class means, and to vanish for their differences. However the costs of proportionate stratification are often so small that we can use them in any case for the entire sample and for design domains.

The most important consideration here may be the sizes of sample clusters when cross-classes and differences are the principal purposes. For these we should consider using larger sample clusters in fewer sampling units, when they result in cost savings that will buy more elements and more precision for the cross-classes. However, the number of sampling units must be large enough for stable variances.

The effects of weighting on domains need more consideration than they can receive here, especially since our discussions tend to assume self-weighting samples. Weights for disproportionate sampling (optimal allocation for the entire sample or for some design domains) may have diverse effects on cross-classes, some of them harmful (section 2). But weights may also be introduced for ratio- or post-stratification estimation. These may also have distinct effects on cross-classes; they may reduce selection biases, but also increase sampling variances which are often proportionately greater for subclasses than for the entire sample.

7. Multipurpose Design

This much neglected subject can receive no adequate treatment here, although its relevance for domains in sampling is paramount. But even in a brief treatment we can go well beyond the question "What is *the* principal statistic the sample must produce?" Displaying a single $\bar{y}$ or $\hat{Y}$ and its "desired precision" is misleading for most survey samples.

First, most samples must produce many survey variables that can have very different variances. Hence producing a "needed" sample size would be

largely fictitious. These distortions are not too painful if the principal survey variables are proportions within the range 0·2–0·8, where the element variances (PQ) fall within a narrow range; and if design effects do not introduce broad ranges for the variances.

However, even then means for domains and for their differences, considered as principal purposes, introduce conflicts into designs of sample sizes. Conflicts arise in other aspects of design, expecially in cluster sampling, where the numbers and sizes of sampling units for diverse stages of selection must also be considered.

A potential source of serious conflict, often unrecognized, arises from increased sample sizes or sampling rates assigned to improve estimates for some domain(s). When viewed from the increased total cost of the sample, this can often decrease the efficiency for estimates based on the entire sample, also for most of the other domains.

Several ingredients are needed for some rational solutions to these conflicts. First, foremost and most difficult are the non-technical obstacles to obtaining a mere listing of principal objectives, with some weighting of their importance; and to go beyond that to realistic specified precisions does not seem usually possible. Second, data on cost factors are not usually available in the needed detail, but these seldom vary critically. Third, data on variances are needed, but I believe (contrary to common supposition) that these are more easily guessed and modelled than the above (e.g. Kish, 1965, section 4.5C). Finally all of the above need to be put together in some acceptable formulation for a rational solution to the conflicts.

This last is a challenge for future research, but I am optimistic that reasonable solutions can be devised, with the aid of computers, provided that needs, the first requirement above, can be determined. Even some of the solutions now available would be helpful. It seems fairly clear that "some compromise is better than none. But optimal allocation . . . is considerably better" still (Kish, 1976, section 7.6).

8. Small Domain Estimation

Estimation for minor domains and for mini-domains has become recently a favourite subject of research. Needs, research, data and computing facilities combine to account for recent advances. These have some of their roots in demographic projections of statistical offices, and others in ratio estimation and in "synthetic methods". The needs and possibilities for minor domains and for mini-domains are neither entirely similar nor entirely different. Furthermore, some can also be found useful for major domains (as ratio estimation is now). This topic is treated by Purcell at this

conference, and by others; for references to these see Purcell and Kish (1979, 1980).

Table 1
Classification by types and sizes of domains/subclasses with examples

Types	Sizes			
	1 Major domains	*1/10* Minor domains	*1/100* Mini domains	*1/1000* or *1/10 000*
Design domains	5 Major regions	50 States of USA	3000 Counties of USA	Rare items
	City sizes	63 Counties of UK 22 Provinces	525 parl. dts in UK	
Mixed domains		Partial segregation: ethnics, resources, etc. Mixed types: regions × education		
Cross-classes	5 year age groups	Single years of age	Years of age × income	
	Major occupa-tions Income classes	Occupation × education	Age × education × income	

REFERENCES

COCHRAN, W. G. (1963, 1977). *Sampling Techniques*, 2nd and 3rd ed. John Wiley and Sons, New York and London.

KISH, L. (1961). Efficient allocation of a multipurpose sample. *Econometrica*, **29**, 363–85.

KISH, L. (1965). *Survey Sampling*, John Wiley and Sons, New York and London.

KISH, L. (1969). Design and estimation for subclasses, comparisons and analytical statistics, Ch. 21 in *New Developments in Survey Sampling*, ed. N. L. Johnson and H. Smith. John Wiley and Sons, New York.

KISH, L. and FRANKEL, M. R. (1974). Inference from complex samples. *Journal of the Royal Statistical Society (B)*, **36**, 1–37.

KISH, L., GROVES, R. M. and KROTKI, K. P. (1976). *Sampling Errors for Feertility Surveys*, World Fertility Surveys, Occasional Paper 7. London.

KISH, L. (1976). Optima and proxima in linear sample design. *Journal of the Royal Statistical Society (A)*, **139**, 80–95.

PURCELL, N. J. and KISH, L. (1979). Estimation for small domains. *Biometrics*.

PURCELL, N. J. and KIS, L. (1980). Postcensal estimates for local areas (or domains). *International Statistical Review*, **48**, 3–18.

VERMA, V., SCOTT, C. and O'MUIRCHEARTAIGH, C. (1980). Sample designs and sampling errors for the World Fertility Survey. *Journal of the Royal Statistical Society (B)*, **42**.

YATES, F. (1953, 1960). *Sampling Methods for Censuses and Surveys*, 2nd and 3rd edns. Chas. Griffin and Co., London.