Epidémiol. santé anim., 1997, 31-32

SAMPLING STRATEGIES FOR DISEASE STUDIES IN TROPICAL COUNTRIES: EXAMPLES OF MULTI-STAGE CLUSTER SAMPLING AND ITS APPLICATION IN ZAMBIA AND KENYA

Kadohira M. [1], Kitala P.[2], Samui, K. [1],McDermott J.[2,3]

2

In order to obtain good information on disease parameters, data must be collected at random. Unfortunately, under tropical conditions, the lists of farms and animals required for random sampling are usually not available. It is usually feasible to get lists of larger clusters such as ecological areas and districts, but not smaller elements. In such circumstances, multi-stage cluster sampling is a practical alternative. Essentially, larger aggregates are divided into smaller sampling areas from which the compilation of sampling frames of herds or animals are logistically feasible.
For two-stage cluster sampling, the sample sizes for each stage depend on their relative contributions to the variability of disease occurrence and costs of sampling. However, extensions to multiple stages are more complicated and are often only informally done. Often, the problem can be reduced to a two-stage problem. For example, in the assessment of vaccination coverage in the WHO Expanded Programme of Immunization (EPI), the number of villages to be sampled within an area is estimated and then a fixed number of children per village, from q random starting household are sampled (Lemeshow and Robinson, 1985) . However, in some circumstances, there can be considerable correlation within these chosen clusters, such that a larger number of children per village need to be sampled to estimate rates with a desired precision. There are two potential solutions. One is to first stratify villages into smaller sampling areas to increase the number of household clusters per village and decrease the number of children per cluster (Levy and Lemeshow, 1991). A second method is to randomly sample households. This is not always feasible in larger towns but in our experience is not too difficult for sampling either herds or households in rural settings. The relative number of areas, farms per area and animals per farm in very large herds will then depend on the relative variability and costs of sampling at each level.
In Chongwe District of central Zambia, a two (area and farm) and three (area, farm and animal) stage sampling strategy was used for questionnaire and blood sample collection respectively. Preliminary estimates indicated that farm-to-farm variation was most important (and particularly for small farms), that areas were relatively homogeneous and that the costs of visiting farms once in an area is relatively small. Thus, only 2 of 5 areas but 36 farms per area were sampled. With 3 stage sampling, the price of sampling animals and within herd variability can be incorporated. In general, sampling costs per animal are low so cluster sampling is usually preferred. In Machakos District, Kenya a multi-stage cluster sampling strategy was employed to estimate dog ecology characteristics and collect serum for rabies testing. Again, the small number of dogs per household reduced the problem to a two stage sampling problem. In this case, their were larger area-to-area than farm-to-farm differences so a much higher sampling frequency of areas (60%) relative to farms (>20%) was used.

REFERENCES

Lemeshow S., Robinson D., 1985 Surveys to measure programme coverage and impact: a review of the methodology used by the expanded programme on immunization. World Health Statistical Quarterly, 8:65-75.
Levy P., Lemeshow S., 1991. Sampling of Populations: Methods and Applications, 2nd ed. Wiley, New York, 420 pp.

[1] Department of Disease Control, School of Veterinary Medicine, University of Zambia Box 32379, Lusaka, Zambia
[2] Department of Public Health, University of Nairobi, Kenya
[3] Deparment of Population Medicine, University of Guelph, Canada