Non – Sampling Errors

Dr. A.C. Kulshreshtha
U.N. Statistical Institute for Asia and the Pacific (SIAP)

Second RAP Regional Workshop on
Building Training Resources for Improving Agricultural & Rural Statistics
Sampling Methods for Agricultural Statistics-Review of Current Practices
SCI, Tehran, Islamic Republic of Iran
10-17 September 2013

Types of Non-sampling Errors (NSE)

- Specification errors
- Coverage errors
- Measurement or response errors
- Non-response errors
- Processing error
- Interpenetrating sub-sampling
- Evaluation of NSE

Specification Errors

- This occurs when the concept implied by the question is different from the underlying construct that should be measured
- A simple question such as whether a household is an agricultural household can be subject to different interpretations
- Meaning of the questions must be conveyed in an unambiguous way and must be properly understood by the respondent

Coverage errors

- Under coverage - Omissions
- External units included
- Duplication
- Deliberate and considered exclusions are **not** coverage errors
- Updating the frame necessary

Measurement errors

- What is observed or measured departs from the actual values of sample units
- These errors concern the accuracy of measurement at the level of individual units
- Also called ***response error***
- Mathematical treatment of measurement errors is available in the form of linear response error models (Refer Cochran, W. G. (1977))
- Models also used in the treatment of inter-penetrating net-work of sub-sampling which is used for estimating the enumerators' effect

Non-response errors

- Failure to get response from some of the sample units
- Population may be conceived as having two strata:
    i. Sample units for which measurements can be obtained
    ii. Sample units for which no measurements could be obtained

- Hansen and Hurwitz (1946) -Call-back methods
- Politz and Simon (1949) – No call backs but by asking to the respondent as to how many times he was at home during previous week to tackle non-response due to not at home

Non-response errors (Contd.)

Unit and Item Non-response

- ***Unit non-response*** - no information is obtained from certain sample units – may be due to refusal or no contact
- ***Item non-response*** - for some units, information not collected for some items
- Gaps in the data records for responding sample units. Reasons may be due to refusals, omissions by enumerators and incapacity
- Imputations

Processing errors

- Processing errors comprise of
  — Editing errors
  — Coding errors
  — Data entry errors
  — Programming erroors, etc.

Interpenetrating sub-sampling (Mahalanobis , 1946)

- Instead of selecting n units in a sample, k sub-samples of size m=n/k are selected independently
- Field work and processing of sample are planned ensuring no correlation between the errors of measurement of any two units in two different sub-samples
- Suppose the correlation arises solely from biases of enumerators. If each of k enumerators is assigned to a different sub-sample and if there is no correlation between errors of measurement for different interviewers, we have the clear use of this technique

Interpenetrating sub-sampling (Contd.)

- With a suitable model it is possible to estimate the relative amount which the correlated component (in this case due to interviewer's effect) of the response variance contributes to the total variance
- This technique has also been very helpful in estimation of variances for complex statistics

Evaluation of NSE

- **Consistency checks**
  - In designing survey instruments (questionnaires), care should be taken to include certain items of information that will serve as a check on quality of data to be collected
  - If additional items of information are easy to obtain, they may be canvassed for all units covered in the survey, otherwise, they may be canvassed only for a sub-sample of units
  - Desirable to follow some external consistency checks on salient results thorough comparable data sources ➔ important for validity and acceptability of estimates

Evaluation of NSE (Contd.)

- **Sample check/verification**
  - One way of assessing and controlling non-sampling errors in surveys is to independently duplicate the work at the different stages of operation with a view to facilitating the detection and rectification of errors
  - Sample check only on small samples using a smaller group of well-trained and experienced staff
  - it would be possible, not only to detect the presence of non-sampling errors, but also to get an idea of their magnitude

Evaluation of NSE (Contd.)

- **Post-survey checks**
  - To assess NSEs, select a sub-sample, or a sample in the case of a census, and re-enumerate it by using better trained/ more experienced staff than those employed for main investigation
  - Design check-survey to assess coverage / content errors. Re-enumerate all units in sample at high stages, e.g. EAs and villages to detect coverage errors
  - Re-survey only a sample of ultimate units ensuring proper representation for different parts of population which have special significance on non-sampling errors

Evaluation of NSE (Contd.)

- With the sample check, rectification work can only be carried out on the sample checked
- This difficulty can be overcome by dividing the output at different stages of survey, e.g. filled in schedules, coded schedules, computation sheets, etc., into lots and checking samples from each lot
- When error rate in a particular lot is more than specified level, the whole lot need to be checked and corrected for errors, thus improving quality of final results

THANKS