Small-Sample Techniques in Quality Assurance and M&E

with focus on Lot Quality Assurance Sampling

Luis Crouch
9-26-2007

e-mail lcrouch@rti.org

- There is increasing pressure to monitor and evaluate
- Sampling is a powerful technique
  - <u>You seldom need censuses or full counts for monitoring anything</u>
  - <u>A good sample is 90% as accurate as a full census, at perhaps 10% (or less) of the cost</u>
- In some cases need to do full counts anyway, for political or administrative reasons
- But because sampling is so much cheaper, can often complement counts with samples, for particular issues, or to dig in depth
- But sampling is fine for many purposes even without full counts
- Sampling can be used in USAID projects, but more importantly should be taught to counterparts to monitor quality professionally

- Sampling is at the heart of all the most powerful social science, industrial quality control methods, and policy study methods
  - More used in industry than in the public sector
  - Within the public sector, more used in health than education
- Why does it work?
  - Essentially: units in a population typically do not differ that much from each other, and therefore small numbers can “represent” large numbers.
    - You can purposefully over-represent groups that you know differ from the rest.
  - The more units differ from each other, the larger a sample you need in order for it to be “representative.”
  - To put it in a folksy way: Do you need to drink the whole pot of soup before you can pronounce it sufficiently salty (or not salty enough)? Or will a sip do?
  - If the soup has been stirred (if the sampling is random), then no: <u>a sip will do</u>

Samples often need to be biggish… smaller than full count, but still biggish…

- This still imposes a cost
- Say you want to know “What percent of teachers are happy with the new curriculum?”
- For that, while you don’t need a full count, you may need to sample as many as 1000 teachers, if you want to be quite safe you’ve got good knowledge on their opinion.

But…

There are some ways to radically lower sample size…

- We can radically lower sample size if all we want to know is <u>whether</u> we are hitting a monitoring goal, but we don't care what the actual value is
- Very small samples: for "whether" issues as opposed to "how much" issues
- Folksy example:
  - You toss a coin 6 times and it turns out heads each time
  - You can be 98.5% sure it is not quite a fair coin, as there is only a 1.5% chance of this happening with a fair coin
  - But you can't be very sure about HOW unfair it is… It could be 80% weighted towards heads or 85% or 90%. 6 heads would be a pretty likely result under any of those scenarios

One particular useful technique: Lot Quality Assurance Sampling (LQAS)

- Used first in industry
- You can’t test every item, because it is expensive and in some cases destroys the items (firing bullets to see if they are good)
- But you have to test some, to have some control over quality
- So you sample-test the lots, count the number of defectives, and reject or accept the lot depending on the number of defectives
- Health sector has borrowed this and is beginning to make very good use of it…

Lot Quality Assurance Sampling (LQAS)

- Basic idea
- Suppose you want to see: “Are teachers using active techniques to teach reading?”
- If >80% of teachers are doing it in a district, you don’t intervene, as that is good enough
- If <50% are doing it, you want to be sure you intervene
- LQAS develops rules of thumb that say, for example:
  - Sample <u>just 26 teachers per district</u>. If more than 8 teachers are not using active teaching, classify the district as non-compliant and needing help. You’ll be right more than 90% of the time.

Lot Quality Assurance Sampling (LQAS)

- You won't know what the average level of use of active teaching is in each district
  - The sample size won't be large enough for that
- But if you aggregate up the district data, in most cases you can tell what the average level is in the province or country
  - The sample size will be large enough for that
- So you will be able to tell how far off you are on average, and, for each district (or school) you will be able to tell just whether they are hitting the goal or not
- In most quality assurance applications that is all we need to know: we just use this information to provide more support and demand more accountability from the non-performers
- All this for about 25 teachers (schools, or students, depending on the unit of analysis) for whatever management unit you want (nation, province, or district)