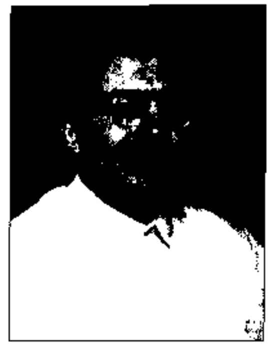
A Look at Raking for Weight Adjustment



Willard C. Losinger

The National Animal Health Monitoring System (NAHMS) is a relatively new program of the United States Department of Agriculture (USDA) (Hueston, 1990). The first NAHMS national study was the 1990 National Swine Survey, which provided information on farm biosecurity practices, facility characteristics, swine diseases, and routine preventive/treatment practices (USDA, 1992). Since then, we have generally done two national studies per year, revisiting a livestock species every five years. The 1995 National Swine Study concentrated on health management practices in the grower/finisher phase of production (Losinger et al., 1998). The NAHMS Swine 2000 study was the third NAHMS national study of swine producers, and provided not only new information on swine diseases and management, but served as a basis for profiling changes in the swine industry (based on information from the previous two NAHMS surveys). More information on the NAHMS is available at the web site *http://www.aphis.usda.gov/vs/ceah/cahm/*.

In many U.S. livestock industries, the population of animals tends to be highly concentrated on a small percentage of the farms (Losinger, 1997). In the U.S. swine industry, in particular, increasing returns to scale have been associated with a rapidly increasing concentration of pigs onto fewer, larger operations (Losinger et al., 1999). Some producers have farrow-to-finish operations, raising pigs from birth until they are ready for slaughter. Some producers specialize in the farrowing phase of production, sending pigs (after they are ready to leave their mothers) to other farms that specialize in fattening pigs for market. In selecting participants for NAHMS national studies, there is a certain trade-off between representing animals and representing livestock producers. To select farms for participation in NAHMS national studies, farms are generally grouped into strata based on farm-size (i.e., number of animals on the farm) within states. Larger farms (which account for the majority of animals) are sampled at a higher rate than smaller farms (which are more numerous, but have a much smaller fraction of the animals) (Losinger et al., 2000).

Willard Losinger (wlosinger1@netscape.net) is a statistician with the US Department of Energy's New Brunswick Laboratory, and has been a federal statistician for 12 years. Other statistical experience includes a couple of years in the pharmaceutical industry, and two years as a US Peace Corps volunteer, where he served as Chief of the Survey Division of Tonga's Statistics Department.

Initial Sampling Weights

From the farms that participate in a survey, we generate estimates that apply to all of the farms and to all of the animals in the participating states. Since farms have different probabilities of selection based on their size, we can't just take the average value from among the respondents and say that this is what average farms do or that this is how average pigs are managed. Generally, each small farm in your sample represents a lot more farms than each large farm in the sample and so must be assigned a sample weight equal to the number of farms in the population that a farm in your sample represents for estimation purposes. Initially, this is just the inverse of the sampling fraction within each stratum. For example, if your sampling rate for a particular stratum is one in ten, then each sampled farm in this stratum represents a total of ten farms in the population (itself, plus nine other farms). If your sampling rate in another stratum is one out of two, then each sampled farm in this stratum represents two farms in the population. Since large farms have a higher sampling rate than small farms, large farms receive lower sample weights than small farms.

As a simple hypothetical example, suppose that Stratum A consists of 20 small farms and that two of them are sampled (call them A1 and A2), for a sampling rate of one-in-ten. Suppose further that one of these two farms has a sick pig. Finally, suppose that Stratum B consists of two large farms, that one farm is sampled (call it B1), and that it does have a sick pig. Thus, within the sample, two out of three farms (66.7%) have sick pigs. However, each farm sampled in Stratum A represents ten farms in the population, and the one farm sampled from Stratum B represents two farms. Therefore, our adjusted estimate is that 12 farms (the ten represented by farm A1 and the two represented by farm B1) of the 22 in the population (54.5%) have sick pigs. Similarly, we could estimate the number of farms employing this or that practice (for example, using a particular vaccine or feed additive).

Animal-Level Weights

In addition to learning about practices of farms, we also want to be able to make estimates about the individual animals. Examples include the number of sick pigs, the number of pigs that receive a particular vaccine or feed additive, and death rates. Animal-level weights are created by multiplying the farm's weight by the number of animals that the farmer reported, in order to estimate the number of animals in the population that the animals on a participating farm are representing. Continuing the above example, suppose that farm A1 has ten pigs, one of which is sick, that farm A2 has twenty pigs, none of which is sick, and that farm B1 has 100 pigs, 90 of which are sick. The two farms sampled in Stratum A would then have pig-level weights of 100 and 200, respectively (10 pigs times the ten farms represented by A1, 20 pigs times the ten farms represented by A2). The sample farm from Stratum B would have a pig-level weight of 200 (100 pigs times the two farms represented by B1). Thus, we would estimate a total of 500 pigs in the population. Applying the pig-level weights to the percentage of sick pigs on each sampled farm, we would estimate that 38% ($100 \times .1$ for farm A1, 200×0 for farm A2, and $200 \times .9$ for farm B1, divided by the total estimate of 500 pigs) are sick. If we had merely estimated the percentage of sick pigs based on the animals sampled, we would have calculated $(1 + 0 + 90)/(10 + 20 + 100)$, which is 91/130, or about 70%. This is much different from the previous estimate, indicating that pig-level adjustments are necessary to obtain a more accurate view of the animal populations.

Response Adjustment

In fact, even these sample weight adjustments are not sufficient because, when we implement a survey, we invariably find that not every producer is still in business or willing to participate in the survey when visited by the enumerator. Therefore, weights need to be transferred from sampled farms that would have been eligible but refused to participate in the survey, to farms that participated. This is accomplished by creating a response adjustment equal to the sum of weights of eligible farms divided by the sum of weights of respondents, generally either within the original sampling strata or within poststrata (i.e., strata defined after the data have been collected) (Losinger et al., 1998). Typically, if a stratum has fewer than 20 respondents, then farms within this stratum are combined with farms from another stratum (in a similar region or farm-size group) to form a new poststratum. Weights of nonrespondents are set to zero, and weights of respondents are multiplied by the response adjustment. Sometimes, low participation rates in a few parts of the country can have a greater impact on the resulting weights than the initial sampling rates did.

In our hypothetical example, suppose that the farm with ten pigs (A1) had chosen not to participate. We would transfer its sample weight to the participating farm with 20 pigs, giving the participating farm with 20 pigs (A2) an adjusted weight of 20. Our estimate would then be that two (9.1%) of the 22 farms in the population have sick pigs, and that 30% ($20 \times 0 + 90 \times 2 = 180$ divided by $20 \times 20 + 100 \times 2 = 600$) of pigs are sick.

This small example illustrates that the sample weights can have a dramatic effect on the population estimates. However, we do have more information. The National Agricultural Statistics Service (NASS) publishes the number of operations and the number of pigs by state and size groups. We can use these inventory numbers as a way to verify our weights, by seeing if the resulting numbers "match." If they do not, we can make further adjustments to our weights.

Inventory Adjustment

For NAHMS surveys, the traditional inventory-adjustment method has been to force inventory estimates to match the NASS numbers by state and size groups. First, a NAHMS inventory estimate was computed (for each state-by-size group cell) by summing the animal-level weights (within each state-by-size group cell). Then, each participant's weight was multiplied by the ratio of the NASS published inventory to the NAHMS inventory estimate by state and size group (Losinger et al., 1998). Then, it was necessary to examine the distribution of the adjusted weights. At this stage in particular, we had to be extremely wary of the impacts that the inventory adjustments were having on both farm-level and animal-level estimates. Frequently, a small number of respondents ended up with extremely large weights (after the inventory adjustment) compared to the majority of the respondents in the sample. If we didn't do anything about it, then the population estimates would have been heavily dependent on the responses given by the respondents with large weights. Generally, respondent weights exceeding a particular value were truncated to a maximum value, and their excess weight was redistributed among all respondents within their poststratum (Losinger et al., 1998). Basically, within each poststratum, each participant's inventory-adjusted weight was multiplied by the ratio of the sum of the untruncated inventory-adjusted weights to the sum of the truncated inventory-adjusted weights. Thus, even the truncated inventory-adjusted weights received the adjustment

Outline of Steps in the Weight Creation Process for National Animal Health Monitoring System (NAHMS) data.

1. **Initial sample weight:** the inverse of the sampling fraction within each sampling stratum [initial sampling weight = 1/sampling rate]
2. **Response adjustment:** transfer weights from eligible non-respondents (i.e., farms that were selected in the sample and that would have been eligible to participate in the survey, but that refused or somehow failed to participate in the survey) to farms that participated in the survey [response adjustment = (sum of weights of eligible farms) / (sum of weights of poststratum)].
3. **Inventory adjustment:** force estimates of inventory (i.e., numbers of animals) to match figures published by the National Agricultural Statistics Service (NASS). We evaluated our traditional method and raking [inventory adjustment = (NASS estimate)/(NAHMS estimate) within each cell or raking between marginals].

The traditional method of performing the inventory adjustment had the following steps:

a. Compute NAHMS inventory estimates by cells based on state and size groups (using weights from Step 2 and inventory figures provided by participants).
b. Multiply each participant's weight by the ratio of NASS published inventory to NAHMS inventory estimates (within each cell).
c. Smooth excessively large weights by truncating to a maximum value, and redistributing excess weights to participants within the same poststratum.

Raking for inventory adjustment had the following steps:

a. Compute NAHMS inventory estimates by state (using weights from Step 2 and inventory figures provided by participants).
b. Multiply each participant's weight by the ratio of NASS published inventory to NAHMS inventory estimates (by state).
c. Use weights from step b. to compute inventory estimates by size group.
d. Multiply each participant's weight (from step b.) by the ratio of NASS published inventory to the new NAHMS inventory estimates (by size group).
e. Repeat steps a. through d. until weights change little from one iteration to the next.

and ended up with weights greater than the truncation limit. We referred to this procedure as "smoothing."

To perform the inventory adjustment for the Swine 2000 study, we decided to try an alternative weight adjustment method called "raking" (Deming and Stephan, 1940). We had 2,499 participating farms with 100 or more pigs in the 17 states included in the study, with a total of 8,024,131 pigs. Table 1 shows the marginal totals for the numbers of pigs and operations that we sought to represent with the sample (i.e., the population from which we were sampling). NASS provides the total number of operations, which often have multiple farm sites. We computed estimates for farm sites rather than for operations, and we use the terms "participant" and "respondent" to refer to a participating farm site.

Table 2 provides some summary statistics based on the traditional method of adjusting weights for inventory within each of the 85 state-by-size group breakouts. At this stage, this was the result of multiplying each participant's weight by the ratio of the NASS published inventory to the NAHMS inventory estimate within each of the 85 state-by-size groups. Within each state-by-size group, the weighted total numbers of pigs matched the NASS numbers exactly (no surprise there), but the weighted number of farms was often off by quite a bit. Three farms (within a group that had relatively low participation) wound up with extremely huge weights (more than 17,000—all other weights were less than 3,000). The next phase in the traditional weight adjustment method would be to "smooth" the more outlandish weights by truncating their weights to some maximum reasonable number, and then reallocating their excess weight to farms in a similar part of the country and in a similar size group. We always had to pay very close attention to what was happening with the weights and resulting estimates, and to make judgments about various tradeoffs involved in choosing one cutoff versus another. Then, if estimates of the numbers of farms were way off, some arbitrary compromises had to be made between misrepresenting the number of farms and misrepresenting the number of animals.

With raking, we do not adjust weights for cells individually. Instead, first we adjust all participant weights to match one set of marginal totals, and then the other set of marginal totals. Then, we go back and do it again ("raking" back and forth) until we achieve convergence (i.e., very little change from one iteration to the next).

Raking Analysis

In this case, first we adjusted all weights so that the weighted sum of pigs would match the NASS-published estimates by state (across all five size groups). Then, we adjusted the weights so that the weighted sum of pigs would match the NASS-

Table 1. Total number of operations and pigs (for operations with 100 or more pigs) on January 1, 2000, in the 17 states included in the NAHMS Swine 2000 Study.

State	Number of Operations	Number of Pigs
Arkansas	420	676,200
Colorado	110	885,550
Illinois	4,440	4,108,500
Indiana	3,800	3,250,500
Iowa	12,100	15,453,500
Kansas	1,000	1,376,100
Michigan	900	990,000
Minnesota	5,300	5,643,000
Missouri	2,100	3,004,250
Nebraska	3,550	2,905,750
North Carolina	1,800	9,552,000
Ohio	2,500	1,330,000
Oklahoma	300	2,255,650
Pennsylvania	950	1,013,250
South Dakota	1,800	1,205,400
Texas	100	809,100
Wisconsin	1,100	533,600
Size Group		
100–499 pigs	20,490	4,314,150
500–999 pigs	8,820	5,092,700
1,000–1,999 pigs	6,205	7,206,750
2,000–4,999 pigs	4,815	12,591,850
> 5,000 pigs	1,900	25,786,900
Total	42,230	54,992,350

Source: *http://www.usda.gov/nass*

published estimates by size group (across all seventeen states). Then, we went back and adjusted again by state, and then by size group, and so on, until we did a total of ten adjustments. Convergence happened pretty quickly (Table 3). We could have easily stopped prior to ten iterations, but went on to ten to see what would happen. The variability in the resulting weights was much less than with the traditional weight adjustment method; indeed, the maximum weight was only 232 instead of over 26,000! The raking method corrected for the number of pigs and had nothing to do with the number of farms during the process. However, an examination of the state-by-size group weighted number of farms and pigs showed that we were reasonably close to the NASS-reported numbers (much better than what we had usually experienced with our traditional weight adjustment methods). Moreover, we didn't have any extreme weights to smooth and fuss with at all.

One pitfall with raking is that you have to pay attention to falling weights—especially weights that fall below one. You know that a farm in your sample represents at least itself in the population—it cannot represent any less than itself. Therefore, one weight that fell slightly below one was rounded up to one. We used the resulting weights to estimate numerous survey parameters related to swine health and management, and to provide information that will ultimately be used to improve swine production practices in the United States.

With raking, convergence is not necessarily guar-

Table 2. Summary of results using the traditional inventory weight adjustment method of multiplying each farm's unadjusted weight by the ratio of the number of pigs (reported by NASS) to the sum of the weighted number of pigs (i.e., the sum of the products of each farm's unadjusted weights and number of pigs) within each of the 85 state-by-size group poststrata.

	Mean	Minimum	Maximum
Unadjusted Weight	10.19	1.10	124.19
Adjusted Weight	171.44	1.71	26,404.84
Adjustment Factor	16.64	0.74	777.60

Table 3. Summary of results using "raking" to adjust weights to match marginal totals of numbers of pigs by state and size group.

	Mean	Minimum	Maximum
Unadjusted Weight	10.19	1.10	124.19
Weight after:			
First Adjustment	14.08	1.21	162.38
Second Adjustment	16.05	1.05	232.19
Third Adjustment	16.13	1.00	232.46
Fourth Adjustment	16.07	0.99	232.09
Fifth Adjustment	16.10	0.99	232.61
Sixth Adjustment	16.07	0.98	232.22
Seventh Adjustment	16.08	0.98	232.34
Eighth Adjustment	16.07	0.98	232.25
Ninth Adjustment	16.08	0.98	232.28
Tenth Adjustment	16.07	0.98	232.26
Weight Adjustment Factors:			
First Iteration	1.37	0.59	2.08
Second Iteration	1.03	0.80	1.43
Third Iteration	1.00	0.93	1.07
Fourth Iteration	1.00	0.99	1.01
Fifth Iteration	1.00	0.99	1.01
Sixth Iteration	1.00	1.00	1.00
Seventh Iteration	1.00	1.00	1.00
Eight Iteration	1.00	1.00	1.00
Ninth Iteration	1.00	1.00	1.00
Tenth Iteration	1.00	1.00	1.00

anteed. It is possible that a given set of circumstances would lead to divergence (i.e., a bouncing back and forth between two or more points) rather than convergence. Thus, the statistician will have to pay attention to what is happening from one iteration to the next.

Conclusion

Raking weight adjustments (back and forth across marginal totals) demonstrated superiority to our traditional method (of adjusting weights to match totals within individual cells) for performing inventory-adjustments on our weights. Estimates of numbers of farms were better with raking than with our traditional method (which were often way off and required compromises between accurate estimates of numbers of farms and numbers of animals), and we didn't end up with enormous weights (due to low participation in a few cells) that had to be smoothed. Sometimes, our confidence with the results of smoothing was not high because of errors in judgment that might have occurred in deciding where to truncate and smooth. With raking, low participation in any particular cell is generally not a problem, as long as there are enough participants to contribute to the marginal total. With raking, you do have to watch out for falling weights, and you certainly don't want to allow any participant's final weight to remain below one. Whatever weight adjustment method is used, statisticians in this line of work do need to be aware of potential pitfalls, do need to examine distributions of weights at each stage of adjustment, and do need to pay close attention to impacts on estimates. While weight adjustment is a highly specialized branch of survey statistics, "raking" has been used for weight adjustment for over 60 years, and it would be fascinating to see whether raking might have some application to other branches of the statistics profession.

References

Deming, W.E., and Stephan, F.F. (1940), "On a Least Square Adjustment of a Sampled Frequency when the Expected Marginal Totals are Known," *Annals of Mathematical Statistics*, 11, 427–444.

Hueston, W.D. (1990), "The National Animal Health Monitoring System: Addressing Animal Health Information Needs in the USA," *Preventive Veterinary Medicine*, 31, 1–14.

Losinger, W.C. (1997), "The Lorenz Curve Applied to Livestock Populations," *Chance*, 10(2), 19–22.

Losinger, W.C., Bush E.J., Hill, G.W., Smith, M.A., Garber, L.P., Rodriguez, J.M., and Kane, G. (1998), "Design and Implementation of the National Animal Health Monitoring System 1995 National Swine Stud," *Preventive Veterinary Medicine*, 34, 147–159.

Losinger, W.C., Dalsted, N.L., Sampath, R.K., and Salman, M.D. (1999), "Returns to Scale in the Production of Finisher Pigs in the United States," *Investigación Agraria: Producción y Sanidad Animales*, 14, 71–84.

Losinger, W.C., Traub-Dargatz, J.L., Sampath, R.K., and Morley, P.S. (2000), "Operation-Management Factors Associated with Early-Postnatal Mortality of US Foals," *Preventive Veterinary Medicine*, 47, 157–175.

United States Department of Agriculture. (1992), National Swine Survey Technical Report. Document N106.692. USDA:APHIS:VS, CEAH, 555 South Howes St., Fort Collins, CO 80521.