SAMPLING TECHNIQUES FOR FOODBORNE PATHOGENS IN ANIMALS AND ANIMAL PRODUCTS

M. SALMAN, B. WAGNER, AND J. GARDNER

INTRODUCTION. Monitoring for foodborne pathogens in U.S. animal populations prior to harvest has relied primarily on clinical isolates submitted during an outbreak investigation or as part of a diagnostic work-up. Some sampling for these pathogens also has occurred as part of the National Animal Health Monitoring System (Dargatz et al. 2000). Post-harvest detection of foodborne pathogens has utilized clinical samples and samples from compliance monitoring undertaken by regulatory agencies such as the USDA’s Food Safety Inspection Service (FSIS) and the Food and Drug Administration (FDA). Although there is value in linking pathogens with outbreak, diagnosis, and compliance events, there is also a need for increased efforts directed toward collection of reliable and representative samples for the precise and accurate assessment of pathogen levels in animals and in animal products. A well-designed monitoring system based on a sound sampling strategy makes it possible to make inferences concerning population prevalence or incidence levels and trends from a subset of the population. Associated risk factors and interventions can be assessed when the levels are reliably estimated.

SAMPLING METHODS. The purpose of this chapter is to review basic sampling techniques that can be used for estimating the prevalence or incidence of foodborne pathogens in animal or animal products. These techniques should account for the characteristics and distributions of the pathogens. Important considerations include expected prevalence, potential clustering or nonrandom distribution (geographic and biological), and pathogen level. Pathogens where the prevalence is extremely low, such as Salmonella enteritidis in eggs, may require a different sampling approach than a more common pathogen, such as Campylobacter spp. on processed poultry. Also, the phase of production can influence prevalence. Escherichia coli O157:H7, for example, is believed to occur in cattle on most feedlots (Chapman et al. 1993; Hancock et al. 1997) but is much rarer on processed beef (Chapman et al. 1993; Griffin 1995; Mahon et al. 1997). Sampling strategies also should address potential clustering of pathogens such as fecal contamination of beef carcasses. Last, the biological implications of pathogen levels in samples deserve special attention. For example, the presence of a few E. coli O157:H7 organisms on a food product can have serious human health implications, whereas Clostridium perfringens requires much higher spore numbers before human health is of concern. Four primary categories of conventional sampling techniques exist: (1) simple random sampling; (2) stratified random sampling; (3) systematic sampling; and (4) cluster sampling. The remainder of this chapter focuses on the presentation of these four basic sampling techniques and their effectiveness and practicality for sampling foodborne pathogens. The discussion of each technique includes a description of important assumptions, characteristics, and application to foodborne pathogens. The material presented here is intended to be an introduction. Many texts (Schaeffer et al. 1986; Kish 1995) are available for more detailed and theoretical presentations of sampling techniques.

Simple Random Sampling (SRS). Simple random sampling (SRS) is the primary sampling technique, which also serves as the basis for other techniques. Simple random sampling occurs when a sample size of n (for example, selected subset of eggs) from a population of size N (for example, total number of eggs produced on a farm on a given day) is taken in such a way that every possible sample of size n has the same chance of being selected. Ordinarily, this is interpreted as the equal probability of selection of each element from the population, but this would omit simple random sampling with probabilities that are proportional to size (for example, more eggs selected from large houses on the farm). The use of SRS depends on having a list of all elements that are part of the population (called a list frame) from which to select the sample. The selection of the sample must be conducted using a random procedure. In a random procedure, each element of the population is assigned a unique identification. Random numbers are then chosen from a random number table or by a similar method. Then the sample is chosen by selecting the elements that are associated with the selected random numbers. For example, if the objective were to sample feces from ten cattle in a feedlot pen of one hundred cattle, we would start by creating a list frame of animals in the pen. Each animal could be assigned a unique number between one and one hundred. Ten unique numbers between one and one hundred would be selected from a random number table. The ten cattle represented by these random numbers would then comprise the sample.
Estimates of the proportion (for example, percentage of cattle shedding bacteria that are resistant to tetracycline) is made using the following formula:

$$\hat{p} = \frac{x}{n}$$

where \(\hat{p}\) is the proportion of interest, \(x\) is the number of sampled elements that are positive for the condition of interest, and \(n\) is the sample size. So if four (\(x=4\)) of the ten (\(n=10\)) cattle sampled in this example are positive for resistance to tetracycline, then the estimated proportion of this resistance in this pen of cattle is 0.4.

Estimates from SRS are unbiased. Bias is the difference between the mean of the estimate that would be obtained after repetition of the process necessary to get this estimate and the true population value. Thus, if the estimate is unbiased, then on average the estimate will not differ from the population value. If bias is present, then the mean of the estimation process will differ from the population value. Simply stated, estimates that are unbiased are properly aimed at the population target value.

In the preceding example, if we had selected ten different cattle, we may have found a different number of cattle to be shedding resistant bacteria. Measurements should incorporate the variability that arises because all the elements of the population were not measured for the outcome of interest. This variability is called the variance of the estimated proportion. The formula for the variance of a proportion, \(V_p\), under SRS is as follows:

$$V_p = \frac{\hat{p}(1-\hat{p})}{n} \left( \frac{N-n}{N} \right)$$

where \(N\) = population size. The second term in the equation is called the finite population correction factor (Scheaffer et al. 1986). If \(n \leq 0.05N\), the finite population correction factor can be ignored. In the cattle feedlot pen example, \(\hat{p}=0.4\) and \(n=10\), then the estimated variance is 0.022. The standard error, also referred to as the square root of the variance, is 0.15. This standard error can be used to represent the variability associated with sampling when making an inference to the population. This procedure is called a confidence interval. A confidence interval is interpreted as the proportion of times that the true parameter would lie within the bounds of the interval if the sampling procedure were repeated many times. Assuming a normal distribution, the confidence interval under SRS would be:

$$CI = \hat{p} \pm z_{\alpha/2} \cdot se(\hat{p})$$

where CI = confidence interval, \(z_{\alpha/2}\) = standard normal value for the \(\alpha\) level of significance, and \(se(\hat{p})\) = standard error of the proportion. To construct a 95 percent confidence interval for the cattle feedlot pen example, we use the \(z_{0.025}=1.96\) value to arrive at an interval of (0.11, 0.69). The product, \(z_{\alpha/2} \cdot se(\hat{p})\), is referred to as the error bound (Scheaffer et al. 1986). The interval can be interpreted as follows: 95 percent of the time, the true proportion of cattle shedding tetracycline-resistant bacteria in the pen would be between 0.11 and 0.69 if we maintain the SRS sampling with a sample size of 10.

An important step in designing a SRS study is to determine the sample size needed to be able to estimate the prevalence with a desired error bound. The sample size formula is as follows:

$$n = \frac{Np(1-p)}{B^2 + p(1-p)}$$

where \(N\) = population size, \(p\) = the proportion being estimated, and \(B\) = bound on the error. For example, suppose that a proposed study objective is to estimate the percentage of dairy cattle shedding *Mycobacterium paratuberculosis* in a five hundred-head dairy herd. Prior studies estimate the within herd prevalence to be about 5 percent (\(p\)), and we want the error bound to be 2 percent (\(B\)). Using the preceding formula, the required sample size is 244. If no prior information is available regarding the proportion being estimated, a conservative approach would be to use a value of 0.5 for \(p\).

Other parameters, such as mean (for example, number of colonies per plate) and total (for example, bacterial cell count in a gram of feces), also can be estimated from SRS. For example, baseline surveys by the USDA for *Clostridium perfringens* spores and cells in raw ground meat products are based on a direct plate count method in a 25-gram subsample of product (Marks and Coleman 1998). These samples are selected by randomly selecting beef processing plants and then randomly selecting a single sample of product from the plant. Because only one sample was taken from a plant, this is equivalent to SRS at the plant level. Count results were summarized as colony forming units (CFU)/g after adjustment for a 10-fold sample dilution. Of 453 samples, 389 (85.9 percent) were nondetectable for *C. perfringens*. Of the sixty-four positive samples, counts ranged from 10 to \(\geq 100\) CFU/g.

In addition to calculation of the percent of samples that have detectable organisms, the count data can be used to calculate the mean number of colony-forming units per gram. For purposes of calculation of the mean and standard deviation, we assume that the CFU/g measurements were 0 (\(n=389\)), 10 (30), 20 (14), 30 (4), 40 (3), 50 (1), 60 (4), 70 (1), 80 (3), 90 (1), and 100 (3).

The mean is calculated as:

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{1.180}{453} = 0.00396\text{CFU/g}$$

Standard deviation is calculated as:
Chapter 28 / SAMPLING TECHNIQUES FOR FOODBORNE PATHOGENS IN ANIMALS AND ANIMAL PRODUCTS / 277

\[ s = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}} = \sqrt{\frac{\sum_{i=1}^{11} n(x_i - \bar{x})^2}{n-1}} = \frac{88,867.99}{452} = 14.022 \text{CFU/g} \]

where \( n \) = number of CFU units on a plate and \( n = \) number of samples (plates). The standard error is equal to the standard deviation divided by the square root of the sample size (SE= 0.659). The colony-forming unit counts often are transformed to meet the assumption of normality. A common method is to use the log transformation (natural log \( \ln(x+1) \)). The log transformation of the count results in a mean of 0.432 and a standard deviation of 1.10 (SE=0.052). The 95 percent confidence interval for the transformed data is (0.330, 0.533). The confidence interval for the data at the original scale is obtained by first creating a confidence interval with the transformed mean and standard error and then transforming (exponentiating) the confidence limits back to the original scale. The resulting confidence interval is (0.391, 0.704).

Advantages and disadvantages of SRS are listed in Table 28.1. Because elements are chosen at random, estimates of parameters for subgroups, such as cattle in different weight groups, can be calculated. The practicality of SRS is limited by a couple of issues, especially for large surveys for foodborne pathogens. Simple random sampling depends on the existence of a list frame. Typically, list frames do not exist for elements of any type of production animal or animal products, nor is it practical to produce such lists. The populations are simply too large and in such constant fluctuation that it is an impractical task. Even building a list of operations with the specified production animal is costly, time consuming, and would require constant updating. When there are substantial differences in means within groups of the data, there is no method within SRS to alter sampling to reduce variability.

**Table 28.1—Advantages and Disadvantages of the Four Primary Sampling Techniques**

<table>
<thead>
<tr>
<th>Sampling technique</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>STRS</td>
<td>Unbiased estimates of subgroups. Sample size estimates relatively simple. Smaller variance than SRS for fixed sample size when measurements are homogeneous within strata—gain in reliability. Convenient when need estimates of subgroups (strata). Potential cost savings.</td>
<td>May produce larger variance than SRS if little variation between strata and more within strata. Subgroup estimates are biased estimates of population parameters. Sample size computations depend on within-strata variance, which might be difficult to obtain.</td>
</tr>
<tr>
<td>SS</td>
<td>Can post-stratify. Unbiased estimates. Estimation formulas are familiar to most. Easy and convenient (cheaper per unit cost than SRS). Can be implemented without a sampling list frame. Formulas for estimation are the same as SRS.</td>
<td>Performs poorly if there is systematic structure in the data (for example, linear trend or periodicity). Estimates can be biased, especially with small population sizes. The intraclass correlation can impact the variance.</td>
</tr>
<tr>
<td>CS</td>
<td>Cost effective in terms of lowered travel cost and list development. Can be done if there is an inadequate list frame of population elements. May be better, in terms of error, than SRS or STRS if measurements within a cluster are heterogeneous and cluster means (or other estimate) are similar. Can perform single-stage or multi-stage sampling.</td>
<td>High standard errors (variance) if measurements are similar within a cluster. Nontraditional estimators that, if ignored, could lead to improper estimates. Sample size determination and actual design can be complex. Need to estimate within cluster sample size and the number of clusters to sample.</td>
</tr>
</tbody>
</table>

SRS=Simple Random Sampling; STRS=Stratified Random Sampling; SS=Systematic Sampling; CS=Cluster or Two-Stage Sampling.
simple random sample is taken with distinct subgroups (strata) of the population. The principle behind stratified random sampling is to take advantage of similarities within strata and differences between strata to make both population and strata level estimates. Stratified random sampling may produce a more precise estimate than a simple random sample of the same size, especially if the measurements within strata are homogeneous (Scheaffer et al. 1986). Suppose that the objective were to estimate pathogen prevalence in two regions of a country where one region is low prevalence and the other is relatively high. In STRS, random samples would be chosen for each region for determination of pathogen status. Estimates of prevalence could then be made for the regions individually and in combination. If there were relatively large within-strata variability and little between-strata variability, then the variance would tend to be larger than that of SRS with a fixed sample size. A measure of this difference (the ratio of the variance from a design relative to a SRS) is called the design effect.

Estimation of the population proportion in STRS is similar to that of SRS except that the formulas are essentially weighted by stratum population size and sample size. The estimator of the population proportion is as follows:

$$\hat{p} = \frac{1}{N} \sum_{i=1}^{k} N_i \hat{p}_i$$

where $N =$ whole population size, $N_i =$ population size of the $i$th stratum, and $\hat{p}_i =$ is estimate of the within stratum proportion (same equation as SRS). Returning to the pathogen in the two populations example, suppose that there were 5,000 and 10,000 animals in the first and second population, respectively ($N=15,000$). If the estimated prevalence were 5 percent and 40 percent in the first and second populations, the estimate of the prevalence for the whole population would be 0.283. The variance is as follows:

$$\tilde{V}_n = \frac{1}{N^2} \sum_{i=1}^{k} N_i \tilde{V}_i(\hat{p}_i)(N_i - n_i) / n_i$$

where $\tilde{V}_i =$ the variance for the $i$th stratum, based on an SRS formula, and $k =$ number of strata. Confidence intervals can be constructed under the normality assumption, as was done with SRS with the use of this variance estimate. Sample size calculations are somewhat more complicated than those used for SRS because they depend on some knowledge of within-strata variability. Also, the sample can be allocated considering the cost of obtaining a sample in the strata. A researcher can then balance cost and variance to optimize the design. The formula is the following:

$$n = \frac{\sum_{i=1}^{k} N_i \beta_i (1 - p_i)/\omega_i}{N \beta_4 + \frac{1}{4} \sum_{i=1}^{k} N_i \beta_i (1 - p_i)}$$

where $\omega_i =$ the fraction of the sample size allocated to the $i$th strata. An approximate allocation to minimize cost to put in the preceding equation can be obtained using the following formula:

$$n_i = \frac{N_i \sqrt{p_i (1 - p_i)}/c_i}{\sum_{i=1}^{k} N_i \sqrt{p_i (1 - p_i)}/c_i}$$

where $c_i =$ the cost of obtaining a sample in strata $i$. To illustrate these equations, suppose that the preliminary estimate of prevalence in region 1 from the preceding example was 0.1 at a cost of $1.50 per sample, whereas in region 2 the estimated prevalence was 0.5 at a cost of $1.25. The total population remains the same ($N_1=5,000$ and $N_2=10,000$). The proportional allocation for region 1 would be 0.215 and region 2 would be 0.785. These are the $w_i$ values for the prior equation. Thus, if we want to calculate a sample size with a bound on the error of 0.05, the total sample size is 293, of which 63 (or round up to 64) would be taken randomly from the first population and 232 would be taken randomly from the second population.

Stratified random sampling (STRS) has the advantages listed in Table 28.1. Estimates for the strata or subgroups as well as the population parameters are unbiased. Post-stratification is possible if the experimenter is aware of sampling issues that may lead to biased estimation (for example, overrepresentation of a group). Subgroup estimates, although they are unbiased estimates for the group they represent, are biased estimates of the population parameters.

Systematic Sampling (SS). Systematic sampling (SS) is the process that consists of obtaining a sample by randomly selecting an element from the first part of a frame or ordered group and then sampling every element that follows in the list at a specified initial interval. For example, if the interval was twenty, one of the first twenty elements in the frame would be randomly selected and then every twentieth element in the frame would be selected. The interval is determined by the sampling fraction, which is the ratio of the required sample size to the total population. For example, if the objective were to sample the specific site of a slaughtered animal from 100 carcasses in a slaughterhouse that processes 1,000/day, we would start by calculating the required interval for the SS $(1000/100 = 10)$. The interval of ten indicates that every tenth carcass on that day will be sampled. The first sample carcass will be selected using a random procedure for those first ten carcasses that are processed on that day. At the end of the day, we should expect approximately 100 sampled carcasses.

Estimates of the proportion (for example, percentage of carcasses contaminated with bacteria that are resistant to tetracycline) and the variance for SS are calculated with the same formulas that are used for SRS. The confidence interval also is constructed in the same manner under the same assumption of a normal distri-
Similarly, the sample size calculation for SS is the same as it is for SRS.

Potentially, SS can have a lower variance than SRS if a negative correlation exists between elements, but SS will perform poorly if there is systematic structure to the data. This relationship can be seen by examining the following equation:

$$v(\overline{y}) = (\sigma^2/n)(1+(n-1)p)$$

where \( \sigma^2/n \) is the variance of SRS and \( p \) is intraclass correlation coefficient (a measure of the correlation between pairs of elements restricted between -1 and 1).

If the correlation is positive, the variance will increase, whereas the opposite is true if the correlation is negative. If the population is in random order, then \( p \) will be close to zero; thus, there will be little impact of using SS. If the elements are ordered, then we would expect the observations to be heterogeneous and \( p \) to be less than zero, which means that SS would be more effective than SRS for a fixed sample size. If the population is arranged in a periodic fashion, it is possible for the measurements to be more homogenous than one would expect if the SRS was implemented; thus, \( p \) will be greater than zero. In production animal sampling, the researcher should be aware of the potential impact of clustering and evaluate it for the situation.

Systematic sampling (SS) is easy and convenient, but its biggest asset is probably the ability to be implemented without a list (see Table 28.1). For production animal sampling, this characteristic may be of key importance, especially for situations such as sampling animals going through a chute or along a processing line.

**Cluster or Two-Stage Sampling (CS).** In the context of foodborne pathogens in domestic animals, the elements of interest (for example, a steer, sow, or chicken) often occur in groups such as herds and flocks. Animals in these groups, also referred to as clusters, are typically managed similarly and may have similar exposures to pathogens of interest. Cluster or two-stage sampling (CS) can be used as a strategy to investigate groups of animals. The primary sampling unit in these types of sampling is the cluster, such as a herd. In cluster sampling, all the units within a cluster are sampled, whereas with two-stage sampling, only a randomly selected subset of the units in a cluster is chosen. Clusters can be randomly sampled, like the SRS, or subject to other strategies, such as STS. Sampling clusters allows for estimation at both cluster level and within-cluster level for considering both herd and animal-level factors that may influence the presence of foodborne pathogens (McDermott and Schukken 1994).

Cluster sampling has the advantage of not needing a complete list of elements in the population. For animal production, this is an important advantage because of the inability to create an animal-level list, whereas it is possible, with work, to develop an operation-level list or some type of geographic cluster. If clusters are geographical in nature, like blocks in a city or operations in a county, then CS may result in substantial cost savings caused by reduced travel and interview time costs. Potentially, cluster sampling may have lower errors than SRS or STS if measurements within a cluster are heterogeneous and cluster means are similar. In animal production, though, we would expect the opposite: within-cluster homogeneity and heterogeneity between clusters. Thus, the variance may be higher for CS than for SRS in many applications.

The measure of the similarity of diseased (or healthy animals) within a cluster is the intracluster correlation (McDermott and Schukken 1994; Donald 1993). If \( p = 0 \), then the disease is randomly distributed within a population, and animals from within a herd are no more likely to be diseased than any animal selected from the entire population. When \( p = 1 \), then all animals within a herd have the same status. Thus, when using cluster sampling, as \( p \) approaches 1, the herd is behaving more like an individual, and fewer animals within a herd would need to be evaluated to determine herd status. Under this scenario, sampling more herds and fewer animals per herd would be an appropriate approach. When \( p \) is close to zero, then animals are behaving as though they were randomly distributed throughout the population, and it may be more appropriate to sample fewer herds and more animals per herd to optimize the design. McDermott and Schukken (1994) reviewed a number of papers from the veterinary epidemiology literature and found that, where estimable, \( p \) varied between 0.0017 and 0.46.

Estimation of the proportion under CS sampling is similar to that of SRS. The formula is as follows:

$$\hat{p} = \frac{\sum_{i=1}^{n} x_i}{\sum_{i=1}^{n} m_i}$$

where \( x_i \) is the number of positive elements in the \( i \)th cluster, \( l \) is the number of clusters, and \( m_i \) is the total number of elements in the \( i \)th cluster. The estimated variance is the following:

$$\hat{\sigma}^2 = \frac{\sum (x_i - \hat{p}m_i)^2}{N(n-1) - \sum m_i^2}$$

where \( M \) is mean number of elements in a cluster. The total number of elements to be sampled can be calculated with the following formula:

$$n = \frac{n\sigma^2}{\frac{lBM^2}{n} + \sigma^2}$$

where \( \sigma^2 \) is approximated by:
An example of the use of these formulae can be found in Scheaffer et al. (1986). Cluster and two-stage sampling and estimation can get very complex, especially in the presence of an imperfect diagnostic test. More detailed discussion of this topic can be found in Cameron and Baldock (1998a, 1998b), and Jordan and McEwen (1998).

**SUMMARY AND CONCLUSIONS.** Many statistical sampling tools are available to foodborne pathogen researchers. The discussion in this chapter demonstrates that, although only four methods were reviewed, sampling methods can be combined to meet study objectives and to address constraints such as list frame availability and resources. The complexities of pathogen distributions and the diversity of potential environments (for example, animal production systems, animal processing facilities, food, and humans) necessitates careful consideration of the sampling approach prior to study initiation to ensure that sampling effort and expense will provide interpretable results.

Study hypotheses should be carefully contemplated before a sampling technique is selected and the design is finalized. On the other hand, findings from a published study should be scrutinized in terms of the appropriateness of the sampling technique before the findings can be generalized. Although most of the common sampling techniques are presented in this chapter, investigators may use their judgment to consider selecting an alternative technique that can fit their situation. Regardless of the choice of techniques, extrapolation of the findings must be based on the implemented sampling technique.

**REFERENCES**


