

Terminology

Sampling Concepts:

High-consumer – This is an individual identified as having a high daily rate of consumption over the time period of interest (either annually or life-time). A high-consumer may have some low consumption days over the time period, but likely has moderate to high consumption on a greater number of days or incorporates fish into the diet most days. A high consumer can be described as someone who has a usual intake in the upper half of the distribution.

Idaho Fish License holder - This is an individual who is over 18 years old and has either a current license or a license within the last 12 months. For calendar year 2012 the target population is estimated at 449,963.

Low-consumer – This is an individual identified as having a low daily rate of consumption over the time period of interest (either annually or life-time). A low-consumer may have some high consumption days over the time period, but when averaged out over the entire period the daily rate is still low. A low consumer may be described as someone who has a usual intake in the lower half of the distribution.

Non-consumer – Someone who does not consume fish. Typically such an individual is identified from self-reported data.

Operational definition – This is also known as a functional definition. This is a definition that is used to define a variable term or object.

Respondent – Defined as an adult of a household who answers the phone and agrees to take the survey.

Sample Panel - Sample population that will be contacted more than once and asked a set of same questions.

Sample size – This relates to the number of people to be picked for the study. Sample size depends on factors such as 1) the hypotheses or questions being asked, 2) the level of precision desired for the statistic, 3) the homogeneity of the sample population, 4) the sampling technique used, 5) monetary and personnel resources available, and 6) the amount of time available.

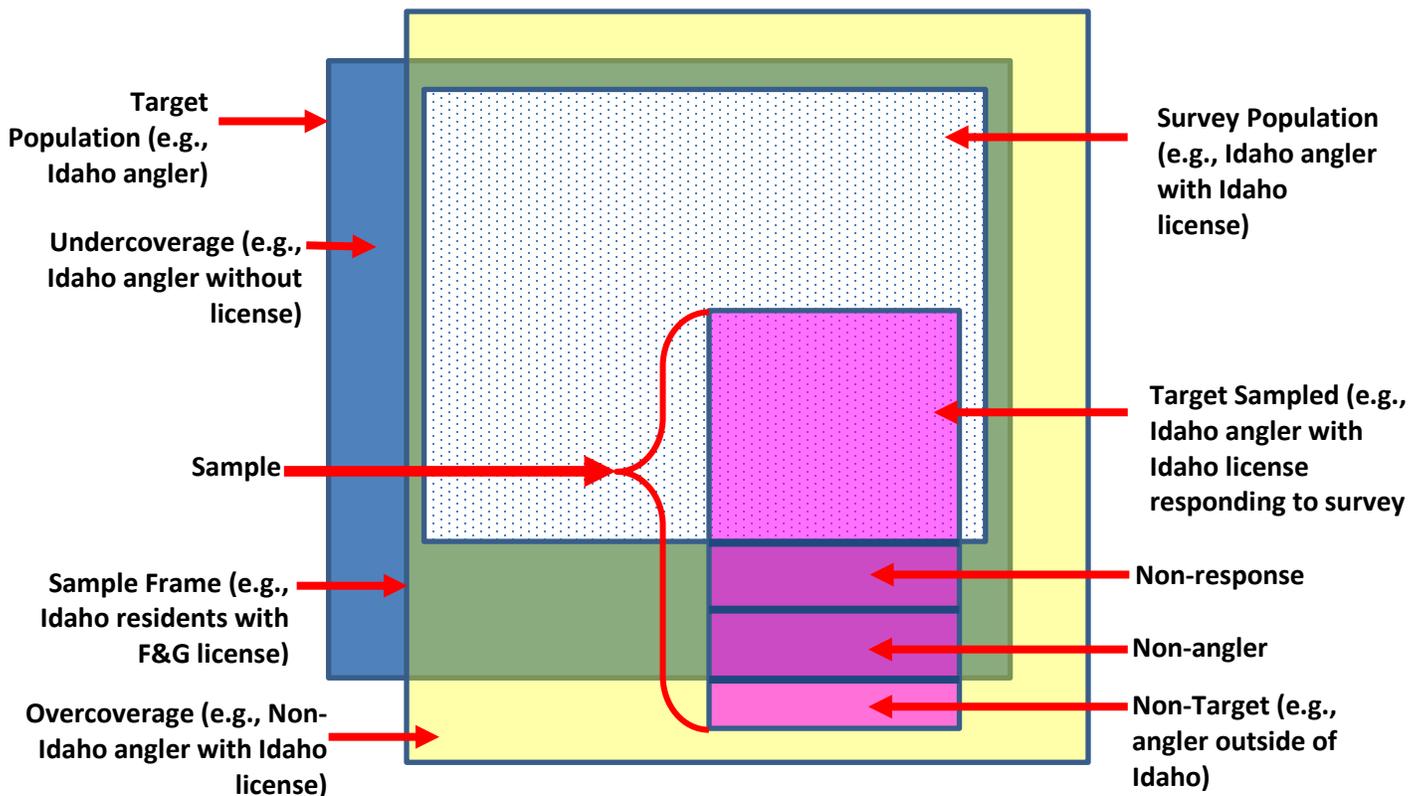
Sampling frame – this is the actual *list* of sampling units from which the sample is selected. It is, simply, a list of the study population (yellow box in figure). For the purposes of the fish consumption survey this list will be derived from records from the Department of Motor Vehicles as well as street addresses plus any other list that can be used to identify residents of Idaho. To create a list of Idaho resident anglers we will be using the Department of Fish & Game's database of licensed anglers. In reality the sampling frame and the target population may not completely match. In some cases there are some individuals who may be part of the target population but who are not captured in the sampling frame (e.g., residents of Idaho with no DMV record or street address). This would be considered the *undercoverage* of the sampling frame. In some cases there are individuals who are not part of the target population but who are captured in the sampling frame (e.g., individuals who hold a fishing license in Idaho but do not live in Idaho). This would be considered the *overcoverage* of the sampling frame.

Simple random sampling – A sample design where each element in the frame has an equal probability of selection.

Stratified sampling - This organizes the random sample into separate strata to minimize within strata variability and maximize variability between strata. Sometimes researchers stratify after the survey is delivered.

Survey Population – the operational definition of the target population; that is the target population with explicit exclusions – for example the population accessible, excluding those outside the state.

Target Population – the entire set of units for which the survey data is to be used to make inferences. It can also be defined as the eligible population that is to be included in the research work. For the purposes of the fish consumption survey, the target populations are Idaho residents and Idaho residents who fish (blue box in figure).



Survey Concepts:

24hr recall method - This is a survey method that uses a short recall period for respondents to access their memory. Usually the researcher wants two 24-hour recall surveys to develop the within-person variability.

BRFSS – This is the Behavioral Risk Factor Surveillance System, is a CDC program which conducts an annual behavioral risk assessment survey nationwide. The CDC allows states to add in a number of their own questions. Idaho put in FFQ questions on the 2007 and 2012 surveys for Idaho residents.

Common core questions - A set of the same or similar questions used in different survey populations. As long as the samples are random, if the same questions are asked the data can be compared across survey instruments.

FFQ - This is the food frequency questionnaire that models how often respondents consume a food item.

Mixed survey method - This is a method that uses the same survey instrument (questionnaire) and collects responses by different methods, such as in person interviews and telephone.

NCI method - This is a statistical approach pioneered by the NCI that uses the assumption that the within person variance does not vary across consumption frequency distribution to create the food frequency distribution based on two 24-hr recall surveys.

SES - Social Economic Status

Survey Instrument – The set of questions developed to collect the survey information.

Survey Method – The way in which questions will be asked of those participating in the survey, e.g., telephone, mail, in-person interviews, etc.

Data Analysis:

Between-person variance - This is a term that describes the difference in the variation in the amount of fish between multiple respondents.

Confidence interval – This is the range about the sample statistic within which the population parameter value is likely to be. A 90% confidence interval means that 9 times out of 10, if we sample the same population in the same way, our interval will contain the population value.

Confidence level (or level of confidence) –This is a statement of how often we want our sample results to fall within a certain range if the survey were to be repeated. Stated another way, it is the degree of certainty of obtaining the similar results for repeat surveys.

Parameter – This is a summary description of a given variable in a population. There is only one value of a parameter for the target population.

Population distribution – This depicts the variation in values, such as usual fish intake or consumption rate, for all individuals in a target population. A parameter from the distribution of fish consumption rates is used to characterize the population for purpose of the development of water quality criteria, e.g., 90th percentile, 95th percentile, etc.

Recall bias - This is what occurs when a respondent provides a response that is shaped by an external factor not asked by the question. An example would be where a respondent wants to give the “right answer” to the interviewer and overstates the frequency that they consume fish.

Skewed distribution – This is a non-symmetrical bell shaped curve where the mean and median are not equal

Statistic – This is a summary description of a given variable in a sample. Each sample of a study population is likely to produce a different estimate of the population parameter.

Usual intake – This is an individual’s long-term average (e.g. annual or lifetime) rate of consumption of a food item, such as fish. It is often expressed in grams per day, as those units fit into human health criteria calculations. Although a rate may be expressed as a daily average, often that rate is estimated from data taken over a short period of time that is then extrapolated to a life-time consumption rate. It does not mean that an individual’s daily consumption is the same from day to day.

Within-person variance - This term describes the variation in the amount of fish consumed by the same person over time.