

3.1 Inductive arguments and statistical generalizations

As we saw in chapter 1 (section 1.8), an inductive argument is an argument whose conclusion is supposed to follow from its premises with a high level of probability, rather than with certainty. This means that although it is possible that the conclusion doesn't follow from its premises, it is unlikely that this is the case. We said that inductive arguments are "defeasible," meaning that we could turn a strong inductive argument into a weak inductive argument simply by adding further premises to the argument. In contrast, deductive arguments that are valid can never be made invalid by adding further premises. Recall our "Tweets" argument:

1. Tweets is a healthy, normally functioning bird
2. Most healthy, normally functioning birds fly
3. Therefore, Tweets probably flies

Without knowing anything else about Tweets, it is a good bet that Tweets flies. However, if we were to add that Tweets is 6 ft. tall and can run 30 mph, then it is no longer a good bet that Tweets can fly (since in this case Tweets is likely an ostrich and therefore can't fly). The second premise, "most healthy, normally functioning birds fly," is a statistical generalization. **Statistical generalizations** are generalizations arrived at by empirical observations of certain regularities. Statistical generalizations can be either universal or partial. Universal generalizations assert that *all* members (i.e., 100%) of a certain class have a certain feature, whereas partial generalizations assert that *most* or *some percentage* of members of a class have a certain feature. For example, the claim that "67.5% of all prisoners released from prison are rearrested within three years" is a partial generalization that is much more precise than simply saying that "most prisoners released from prison are rearrested within three years." In contrast, the claim that "all prisoners released from prison are rearrested within three years" is a universal generalization. As we can see from these examples, deductive arguments typically use universal statistical generalizations whereas inductive arguments typically use partial statistical generalizations. Since statistical generalizations are often crucial premises in both deductive and inductive arguments, being able to evaluate when a statistical generalization is good or bad is crucial for being able to evaluate arguments. What we are doing in evaluating statistical generalizations is determining whether the premise in our argument is true (or at least well-

supported by the evidence). For example, consider the following inductive argument, whose premise is a (partial) statistical generalization:

1. 70% of voters say they will vote for candidate X
2. Therefore, candidate X will probably win the election

This is an inductive argument because even if the premise is true, the conclusion could still be false (for example, an opponent of candidate X could systematically kill or intimidate those voters who intend to vote for candidate X so that very few of them will actually vote). Furthermore, it is clear that the argument is intended to be inductive because the conclusion contains the word “probably,” which clearly indicates that an inductive, rather than deductive, inference is intended. Remember that in evaluating arguments we want to know about the strength of the inference from the premises to the conclusion, *but we also want to know whether the premise is true!* We can assess whether or not a statistical generalization is true by considering whether the statistical generalization meets certain conditions. There are two conditions that any statistical generalization must meet in order for the generalization to be deemed “good.”

1. **Adequate sample size:** the sample size must be large enough to support the generalization.
2. **Non-biased sample:** the sample must not be biased.

A **sample** is simply a portion of a population. A **population** is the totality of members of some specified set of objects or events. For example, if I were determining the relative proportion of cars to trucks that drive down my street on a given day, the population would be the total number of cars and trucks that drive down my street on a given day. If I were to sit on my front porch from 12-2 pm and count all the cars and trucks that drove down my street, that would be a sample. A good statistical generalization is one in which the sample is **representative** of the population. When a sample is representative, the characteristics of the sample match the characteristics of the population at large. For example, my method of sampling cars and trucks that drive down my street would be a good method as long as the proportion of trucks to cars that drove down my street between 12-2 pm matched the proportion of trucks to cars that drove down my street during the whole day. If for some reason the number of trucks that drove down my street from 12-2 pm was much higher than the average for the whole day, my sample would not be representative of the

population I was trying to generalize about (i.e., the total number of cars and trucks that drove down my street in a day). The “adequate sample size” condition and the “non-biased sample” condition are ways of making sure that a sample is representative. In the rest of this section, we will explain each of these conditions in turn.

It is perhaps easiest to illustrate these two conditions by considering what is wrong with statistical generalizations that fail to meet one or more of these conditions. First, consider a case in which the sample size is too small (and thus the adequate sample size condition is not met). If I were to sit in front of my house for only fifteen minutes from 12:00-12:15 and saw only one car, then my sample would consist of only 1 automobile, which happened to be a car. If I were to try to generalize from that sample, then I would have to say that only cars (and no trucks) drive down my street. But the evidence for this universal statistical generalization (i.e., “every automobile that drives down my street is a car”) is extremely poor since I have sampled only a very small portion of the total population (i.e., the total number of automobiles that drive down my street). Taking this sample to be representative would be like going to Flagstaff, AZ for one day and saying that since it rained there on that day, it must rain every day in Flagstaff. Inferring to such a generalization is an informal fallacy called “hasty generalization.” One commits the fallacy of **hasty generalization** when one infers a statistical generalization (either universal or partial) about a population from too few instances of that population. Hasty generalization fallacies are very common in everyday discourse, as when a person gives just one example of a phenomenon occurring and implicitly treats that one case as sufficient evidence for a generalization. This works especially well when fear or practical interests are involved. For example, Jones and Smith are talking about the relative quality of Fords versus Chevys and Jones tells Smith about his uncle’s Ford, which broke down numerous times within the first year of owning it. Jones then says that Fords are just unreliable and that that is why he would never buy one. The generalization, which is here ambiguous between a universal generalization (i.e., all Fords are unreliable) and a partial generalization (i.e., most/many Fords are unreliable), is not supported by just one case, however convinced Smith might be after hearing the anecdote about Jones’s uncle’s Ford.

The non-biased sample condition may not be met even when the adequate sample size condition is met. For example, suppose that I count all the cars on my street for a three hour period from 11-2 pm during a weekday. Let’s assume

that counting for three hours straight give us an adequate sample size. However, suppose that during those hours (lunch hours) there is a much higher proportion of trucks to cars, since (let's suppose) many work trucks are coming to and from worksites during those lunch hours. If that were the case, then my sample, although large enough, would not be representative because it would be biased. In particular, the number of trucks to cars in the sample would be higher than in the overall population, which would make the sample unrepresentative of the population (and hence biased).

Another good way of illustrating sampling bias is by considering polls. So consider candidate X who is running for elected office and who strongly supports gun rights and is the candidate of choice of the NRA. Suppose an organization runs a poll to determine how candidate X is faring against candidate Y, who is actively anti gun rights. But suppose that the way the organization administers the poll is by polling subscribers to the magazine, *Field and Stream*. Suppose the poll returned over 5000 responses, which, let's suppose, is an adequate sample size and out of those responses, 89% favored candidate X. If the organization were to take that sample to support the statistical generalization that "most voters are in favor of candidate X" then they would have made a mistake. If you know anything about the magazine *Field and Stream*, it should be obvious why. *Field and Stream* is a magazine whose subscribers who would tend to own guns and support gun rights. Thus we would expect that subscribers to that magazine would have a much higher percentage of gun rights activists than would the general population, to which the poll is attempting to generalize. But in this case, the sample would be unrepresentative and biased and thus the poll would be useless. Although the sample would allow us to generalize to the population, "*Field and Stream* subscribers," it would not allow us to generalize to the population at large.

Let's consider one more example of a sampling bias. Suppose candidate X were running in a district in which there was a high proportion of elderly voters. Suppose that candidate X favored policies that elderly voters were against. For example, suppose candidate X favors slashing Medicare funding to reduce the budget deficit, whereas candidate Y favored maintaining or increasing support to Medicare. Along comes an organization who is interested in polling voters to determine which candidate is favored in the district. Suppose that the organization chooses to administer the poll via text message and that the results of the poll show that 75% of the voters favor candidate X. Can you see what's wrong with the poll—why it is biased? You probably recognize that this polling

method will not produce a representative sample because elderly voters are much less likely to use cell phones and text messaging and so the poll will leave out the responses of these elderly voters (who, we've assumed make up a large segment of the population). Thus, the sample will be biased and unrepresentative of the target population. As a result, any attempt to generalize to the general population would be extremely ill-advised.

Before ending this section, we should consider one other source of bias, which is a bias in the polling questionnaire itself (what statisticians call the "instrument"). Suppose that a poll is trying to determine how much a population favors organic food products. We can imagine the questionnaire containing a choice like the following:

Which do you prefer?

- a. products that are expensive and have no FDA proven advantage over the less expensive products
- b. products that are inexpensive and have no FDA proven disadvantage over more expensive products

Because of the phrasing of the options, it seems clear that many people will choose option "b." Although the two options do accurately describe the difference between organic and non-organic products, option "b" sounds much more desirable than option "a." The phrasing of the options is biased insofar as "a" is a stand-in for "organic" and "b" is stand-in for "non-organic." Even people who favor organic products may be more inclined to choose option "b" here. Thus, the poll would not be representative because the responses would be skewed by the biased phrasing of the options. Here is another example with the same point:

Which do you favor?

- a. Preserving a citizen's constitutional right to bear arms
- b. Leaving honest citizens defenseless against armed criminals

Again, because option "b" sounds so bad and "a" sounds more attractive, those responding to a poll with this question might be inclined to choose "a" even if they don't really support gun rights. This is another example of how bias can creep into a statistical generalization through a biased way of asking a question.

Random sampling is a common sampling method that attempts to avoid any kinds of sampling bias by making selection of individuals for the sample a matter of random chance (i.e., anyone in the population is as likely as anyone else to be chosen for the sample). The basic justification behind the method of random sampling is that if the sample is truly random (i.e., anyone in the population is as likely as anyone else to be chosen for the sample), then the sample will be representative. The trick for any random sampling technique is to find a way of selecting individuals for the sample that doesn't create any kind of bias. A common method used to select individuals for a random sample (for example, by Gallup polls) is to call people on either their landline or cell phones. Since most voting Americans have either a landline or a cell phone, this is a good way of ensuring that every American has an equal chance of being included in the sample. Next, a random number generating computer program selects numbers to dial. In this way, organizations like Gallup are able to get something close to a random sample and are able to represent the whole U.S. population with a sample size as small as 1000 (with a margin of error of ± 4). As technology and social factors change, random sampling techniques have to be updated. For example, although Gallup used to call only landlines, eventually this method became biased because many people no longer owned landlines, but only cell phones. If some new kind of technology replaces cell phones and landlines, then Gallup will have to adjust the way it obtains a sample in order to reflect the changing social reality.

Exercise 22: What kinds of problems, if any, do the following statistical generalizations have? If there is a problem with the generalization, specify which of the two conditions (adequate sample size, non-biased sample) are not met. Some generalizations may have multiple problems. If so, specify all of the problems you see with the generalization.

1. Bob, from Silverton, CO drives a 4x4 pickup truck, so most people from Silverton, CO drive 4x4 pickup trucks.
2. Tom counts and categorizes birds that land in the tree in his backyard every morning from 5:00-5:20 am. He counts mostly morning doves and generalizes, "most birds that land in my tree in the morning are morning doves."
3. Tom counts and categorizes birds that land in the tree in his backyard every morning from 5:00-6:00 am. He counts mostly morning doves and generalizes, "most birds that land in my tree during the 24-hour day are morning doves."

4. Tom counts and categorizes birds that land in the tree in his backyard every day from 5:00-6:00 am, from 11:00-12:00 pm, and from 5:00-6:00 pm. He counts mostly morning doves and generalizes, "most birds that land in my tree during the 24-hour day are morning doves."
5. Tom counts and categorizes birds that land in the tree in his backyard every evening from 10:00-11:00 pm. He counts mostly owls and generalizes, "most birds that land in my tree throughout the 24-hour day are owls."
6. Tom counts and categorizes birds that land in the tree in his backyard every evening from 10:00-11:00 pm and from 2:00-3:00 am. He counts mostly owls and generalizes, "most birds that land in my tree throughout the night are owls."
7. A poll administered to 10,000 registered voters who were homeowners showed that 90% supported a policy to slash Medicaid funding and decrease property taxes. Therefore, 90% of voters support a policy to slash Medicaid funding.
8. A telephone poll administered by a computer randomly generating numbers to call, found that 68% of Americans in the sample of 2000 were in favor of legalizing recreational marijuana use. Thus, almost 70% of Americans favor legalizing recreation marijuana use.
9. A randomized telephone poll in the United States asked respondents whether they supported a) a policy that allows killing innocent children in the womb or b) a policy that saves the lives of innocent children in the womb. The results showed that 69% of respondents choose option "b" over option "a." The generalization was made that "most Americans favor a policy that disallows abortion."
10. Steve's first rock and roll concert was an Ani DiFranco concert, in which most of the concert-goers were women with feminist political slogans written on their t-shirts. Steve makes the generalization that "most rock and roll concert-goers are women who are feminists." He then applies this generalization to the next concert he attends (Tom Petty) and is greatly surprised by what he finds.
11. A high school principal conducts a survey of how satisfied students are with his high school by asking students in detention to fill out a satisfaction survey. Generalizing from that sample, he infers that 79% of students are dissatisfied with their high school experience. He is surprised and saddened by the result.
12. After having attended numerous Pistons home games over 20 years, Alice cannot remember a time when she didn't see ticket scalpers

selling tickets outside the stadium. She generalizes that there are always scalpers at every Pistons home game.

13. After having attended numerous Pistons home games over 20 years, Alice cannot remember a time when she didn't see ticket scalpers selling tickets outside the stadium. She generalizes that there are ticket scalpers at every NBA game.
14. After having attended numerous Pistons home games over 20 years, Alice cannot remember a time when she didn't see ticket scalpers selling tickets outside the stadium. She generalizes that there are ticket scalpers at every sporting event.
15. Bob once ordered a hamburger from Burger King and got violently ill shortly after he ate it. From now on, he never eats at Burger King because he fears he will get food poisoning.

3.2 Inference to the best explanation and the seven explanatory virtues

Explanations help us to understand *why* something happened, not simply convince us *that* something happened (see chapter 1, section 1.3). However, there is a common kind of inductive argument that takes the best explanation of *why* x occurred as an argument for the claim *that* x occurred. For example, suppose that your car window is broken and your iPod (which you left visible in the front seat) is missing. The immediate inference you would probably make is that someone broke the window of your car and stole your iPod. What makes this a reasonable inference? What makes it a reasonable inference is that this explanation explains all the relevant facts (broken window, missing iPod) and *does so better than any other competing explanation*. In this case, it is perhaps possible that a stray baseball broke your window, but since (let us suppose) there is no baseball diamond close by, and people don't play catch in the parking garage you are parked in, this seems unlikely. Moreover, the baseball scenario doesn't explain why the iPod is gone. Of course, it could be that some inanimate object broke your window and then someone saw the iPod and took it. Or perhaps a dog jumped into the window that was broken by a stray baseball and ate your iPod. These are all possibilities, but they are remote and thus much less likely explanations of the facts at hand. The much better explanation is that a thief both broke the window and took the iPod. This explanation explains all the relevant facts in a simple way (i.e., it was the thief responsible for both things) and this kind of thing is (unfortunately) not uncommon—it happens to other people at other times and places. The