

Ch. 10: Inductive Arguments

I. Inductive arguments are arguments that try to apply what is known about known objects or situations to unknown objects or situations.

- a. Almost all inductive arguments fit the following pattern.
 - i. A known thing X has certain properties (a, b, c, etc.).
 - ii. Another thing Y that is not known as well as X has the same properties.
 - iii. X also has some additional property (p).
 - iv. On the basis of these three premises the argument concludes that Y also has the additional property (p).
 - v. The general idea is that if Y is like X in some ways (relevant ways of course) then it could also be like X in other ways.
- b. If the premises of an inductive argument are all true then the conclusion has some likelihood or probability of being true.
 - i. Conclusions of inductive arguments are more or less likely to be true; the arguments are correspondingly stronger or weaker.
 - ii. These arguments are never valid or invalid, for they are not deductive arguments.
- c. This general idea of an inductive argument can be stated more precisely with the help of certain key concepts:
 - i. The **property** under investigation – the "p" of the argument – is the property or feature in question.
 - ii. The **sample** is a group of things that are already known or believed to possess the property in question.
 - iii. The **target**, target class, or target population is an individual or a group about which one is asking: Does this target possess the property in question?
 1. Whether the target is a single thing or a larger group of things, the goal of an inductive argument is to show that it possesses the property in question.

2. Every inductive argument's conclusion attributes the property in question to the target.

II. Inductive arguments may be either analogical arguments or generalizations.

- a. The similarities between these two forms of argument are far important than any differences between them.
 - i. Both forms of argument begin with a sample, identify a property of the members of that sample, and conclude that the property is also shared by one or more items outside the sample.
 - ii. Because of these similarities, the two forms of argument are evaluated in exactly the same way.
- b. The difference between the two consists in the nature of their targets.
 - i. In generalizations the target is a class, i.e. a group of things.
 1. You might generalize from the fact that humans feel sleepy after big meals to the conclusion that all animals feel sleepy after big meals.
 2. In every case, the sample is part of the target class. The target class contains the sample.
 3. After all, you are generalizing to a larger group.
 - ii. In analogical arguments the target is an item, a single thing.
 1. Because Tom is a clown like Dick and Harry and they both wear big shoes, you conclude that Tom must wear big shoes too.
 2. In every case, the target item is not part of the sample. The target lies outside the sample.
 3. After all, by virtue of what an analogy is, you are stretching from what you know to what you don't know. The sample by definition is what you know; the target can't be part of it.

III. The strength of an inductive argument depends on the similarity between sample and target, and on the size of the sample.

- a. The similarity is known as the **representativeness** of the sample.
 - i. The sample represents the target in so far as it possesses all the features of the target that are relevant to the property in question.
 - ii. The less you can trust the sample's resemblance to the larger class it's in or to the target item, the less you can trust an argument based on that sample.
 - iii. If half the cats in your town never go outdoors, then before claiming any conclusions about all your town's cats you should consider a group in which about half never go out.
- b. A biased sample is one that fails to represent the target in some relevant respect.
 - i. The trick is to choose a sample that reflects all relevant features of the target, even though you don't always know which features will turn out to be relevant.
 - ii. If you're generalizing about the methods that auto thieves use, do you make your sample represent the educational levels of auto thieves in general? That makes sense, because more educated ones might use more sophisticated tools. What about left- or right-handedness? Would that be relevant?
 - iii. In the case of analogical arguments a related question comes up: Suppose you don't know whether the target does or does not possess some property?
 - iv. In both cases bias is best avoided by the selection of a varied sample.
- c. A large sample makes for a better inductive argument, as long as it is representative.
 - i. Typically, larger samples will also be more representative, all other things being equal.
 - ii. When the larger sample is not more representative, the representativeness is more important than the sample's size.
 - iii. Homogeneous classes are adequately represented even by tiny samples. One sip of the ocean anywhere is enough to show you the whole thing is salty, because ocean water is homogeneous. Heterogeneous classes on the other hand call for larger samples.

IV. Inductive arguments can be either formal or informal.

- a. The popularity of polls and large scientific studies might make you think that inductive arguments are only the domain of scientists. But such arguments turn up, informally, in daily life as well.
 - i. The general principles at work are the same for both formal and informal arguments.
 - ii. Even when you don't realize that you are looking for a representative sample, for instance, you still trust generalizations about human behavior that hold for many different ethnic groups and ages more than generalizations based on your own ethnic group, or based only on people your age.
- b. Two differences are worth noting.
 - i. Formal arguments tend to be generalizations and not arguments from analogy.
 - ii. Formal arguments pay attention to details, and rely on mathematical analysis, in ways that informal arguments don't.

V. Formal inductive arguments tend to follow rigorous guidelines for sample selection and size, and their accuracy is evaluated by means of precisely defined mathematical concepts.

- a. In a poll (for instance) the sample being polled must be representative of the target class it belongs to.
 - i. To represent a target class a sample must possess all the features of the target class that are relevant to the property in question, and must possess those features in the same way (meaning: to the same degree, in the same proportion) that the target class does.
 - ii. Speaking as accurately and precisely as possible, we say that a feature or property P is relevant to another property Q if it is reasonable to suppose that the presence or absence of P could affect the presence or absence of Q.
 - iii. It's not always clear which features will be relevant. On top of that, it can be exceedingly hard to select an exactly representative sample even when all the relevant features are known.
- b. **A random sample** can still be trusted even when the sample is not deliberately chosen so as to be representative.

- i. A randomly selected sample is the type most likely to represent its target class. This is the great virtue of random selection: It is not a process esteemed for its own sake, but because it helps to make a sample representative.
 - ii. In a random sample, every individual is as likely to be selected as every other one.
- c. Sample size, error margin, and confidence level help evaluate the accuracy of generalizations from random samples.
- d. **Error margins** express the potential for error created by random variations.
 - i. Randomly chosen samples are bound to differ from their target class to some degree. One-tenth of all Americans may be left-handed, but a random assortment of 50 Americans might contain only one lefty, or none at all, or as many as 10 or 15.
 - ii. The occurrence of some property in any random sample will be close to its occurrence in the target class; this "closeness" is indicated by a percentage plus or minus some percentage points.
 - iii. Thus the error margin measures the range within which a generalization about the sample also describes the target – e.g., 5 ± 2 .
- e. The closely related concept of **confidence level** measures the likelihood that an inductive generalization has produced a true conclusion.
 - i. Confidence level indicates the percentage of all random samples in which the property in question occurs within the error margin.
 - ii. Suppose a study reports that 28 percent of Estonians have curly hair, with an error margin of \pm percent and a confidence level of 88 percent. Assuming this study looked at a random sample of some given size, this result means that 88 percent of all random samples of the same size will contain between 25 and 1 percent curly-haired people.
 - iii. In most scientifically produced studies, confidence level is set at 95 percent.
- f. The confidence level and error margin both depend on the size of the sample.
 - i. As the size increases, the error margin decreases, or the confidence level increases, or both.

- ii. A poll of 50 people (within a sufficiently large target class) has an error margin of ± 14 with a 95 percent confidence level. If the sample size doubles to 100, the error margin falls to ± 10 at the same confidence level.

VI. Informal inductive arguments follow the same principles as formal ones, but they express error margin and confidence level more loosely.

- a. The general principle remains the same. The smaller the sample of a heterogeneous population, the less reliable the conclusion of the argument will be.
- b. Error margins find their informal analogues in cautious conclusions.
 - i. You make a conclusion cautious by expressing it in less precise terms.
 - ii. Thus instead of saying that the temperature is 8 degrees lower in the middle of the lake, you might say it's somewhere around 10 degrees lower.
- c. Confidence levels in turn get translated into expressions of greater or lesser likelihood that a conclusion is true.
 - i. Ordinary registers of likelihood, from "you can bet that ..." through "it is pretty likely that ..." on down to "it's possible that ..." correlate to technical statements of confidence level.
 - ii. Thus instead of saying that a generalization holds only at a confidence level of 65 percent, you might say the conclusion is more likely to be true than false.

VII. A few fallacies most commonly arise with inductive arguments.

- a. A biased generalization or biased sample uses a sample that may be large enough, yet does not represent the target class.
- b. The fallacy of hasty conclusion occurs in arguments that use too small a sample. If the argument is an inductive generalization, this fallacy is also known as hasty generalization.
- c. The fallacy of anecdotal evidence is one kind of hasty generalization. In such arguments, one tells a story about one or two examples that is meant to demonstrate a general truth.

VIII. Besides the general information about statistical arguments that has been covered in this chapter, you should keep a few specific warnings in mind when evaluating public-opinion polls.

- a. Polls are based on self-selected samples when the people whose opinions are being measured have put themselves into the sample.
 - i. A television call-in poll overrepresents people with strong enough opinions to call and register their opinions.
 - ii. Although such polls may measure strength of feeling, they do not indicate what the general population believes.
- b. Advocacy groups might commission polls that are worded in such a way as to produce a more favorable response. These are polls with slanted questions.
 - i. Slanted questions introduce a form of bias that is separate from the representativeness of the sample.
 - ii. Loaded questions can pretend to solicit an opinion while suggesting a claim: "Do you favor reducing death and street crime by means of emergency housing?"
 - iii. At other times it is the order that questions come in that skews the results.

IX. The importance of large sample size in inductive arguments follows from a statistical principle, the law of large numbers.

- a. When an event occurs at random with predictable ratios of results, it will get closer to those predictable ratios the more often the event is repeated.
 - i. Roll a die 12 times and you might get a twice, but you might get it more times than that and you might not get it once.
 - ii. Roll the same die 12,000 times and it becomes more and more likely that occurs one-sixth of the time.
- b. The law of large numbers assures us that the individuals in samples, when those samples are large enough, behave with the regularity that characterizes the target class.
- c. Sometimes the law of large numbers is misunderstood; then it becomes the **gambler's fallacy**.

- i. The gambler's fallacy is the belief that past random occurrences influence the next occurrence, that in order to reach the predicted ratios the next events have to "catch up."
- ii. The gambler's fallacy occurs very often in nongambling contexts. You might think: "We've had four boys in a row. The next child is more likely to be a girl."
- iii. Although large samples do tend to produce the numbers one expects, those numbers do not influence any particular event. Each time is a new flip of the coin.