Statistical Reasoning

Critical Thinking

Three Kinds of Inductive Arguments

- Inductive Generalization:
  - From particular to general: A is an X and has property P. B is an X and has property P. … Therefore, all (most) X's have property P.
- Inductive Syllogism
  - From general to particular: Most X's have property P. A is an X. Therefore, A has property P.
- Analogical Reasoning
  - From particular to particular: A and B share certain properties. A has property P. Therefore, B has property P.

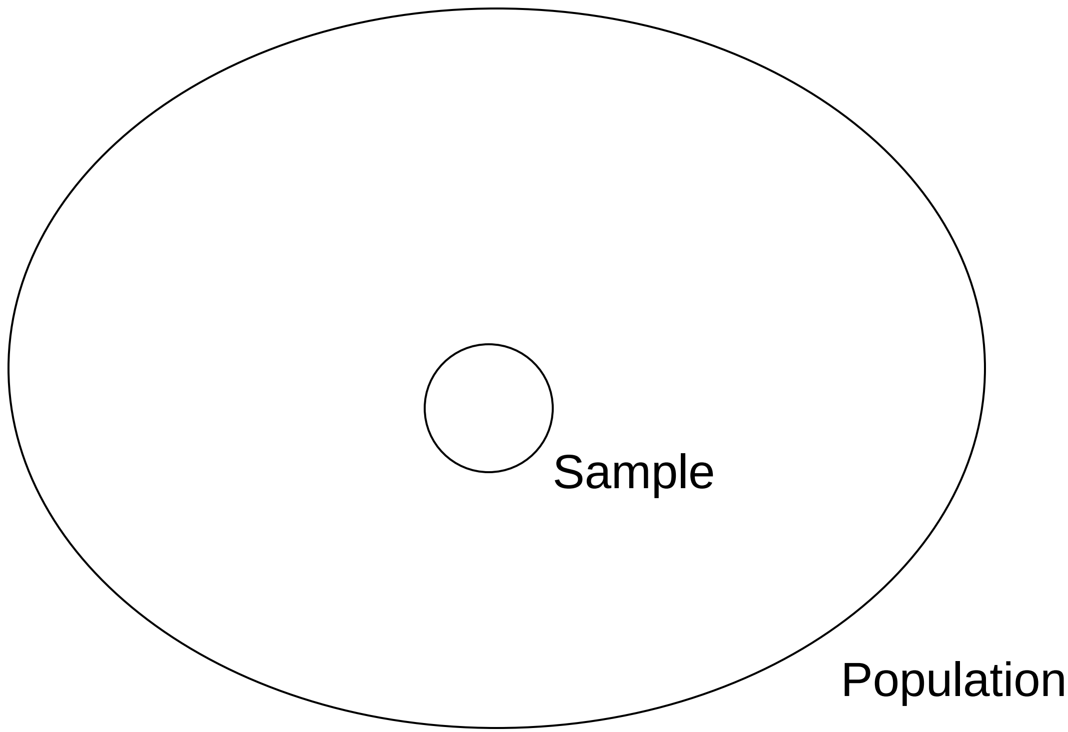
Inductive Generalizations

- Let’s focus on Inductive Generalizations
- Examples:
  - All my German friends like sauerkraut. Therefore, all Germans like sauerkraut.
  - Of 1500 Americans polled, 35% like what Obama is doing. Therefore, 35% of all Americans like what Obama is doing

From Samples to Populations

f: Observed frequency in sample = percentage of sample that has property P

p: actual percentage of population that has property P



Inference:

I observe f

Therefore, p is about f

Sample Size and Sample Representativeness

- How good the inference from a sample to a population is depends on 2 things:
  - Sample Size: The larger the sample size, the more reliable the inference is.
  - Sample Representativeness: The more likely it is that the make-up of the sample is like the make-up of the population, the more reliable the inference is.

The Law of the Big Numbers

- If you flip a coin 10 times, and you get 7 heads (i.e. f = 0.7), does that surprise you (assume the coin is fair, i.e. p = 0.5)? What about 70 out of 100? 700 out of 1000?
- The Law of the Big Numbers is that *the greater the sample size (n), the more likely it is that observed frequency (f) will approximate the actual percentage (p).*
  - So, if the coin is fair, f is more likely to approximate 0.5 for 1000 flips than for 10.
- Also, conversely, the greater n, the greater the likelihood that, given f, the actual value of p is in the neighborhood of f.
  - So, getting 700 heads out of 1000 flips makes it more unlikely that the coin is fair than getting 7 out of 10.

The Law of the Big Numbers and Our Intuitions

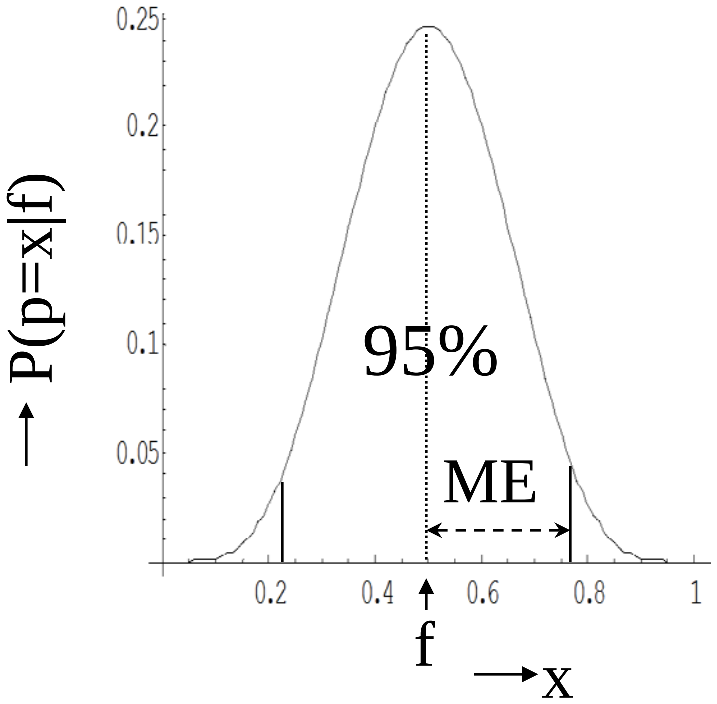
- The *qualitative* nature of the Law of the Big Numbers coincides with our intuitions.
- However, it *is* surprising to us *how quickly* f and p approximate each other as the sample size increases (see two slides from now for exact values).
- Thus, our intuitions are not very good in judging the *quantitative* nature of the Law of the Big Numbers.
- We use the Margin of Error to quantify this ‘likelihood of being in the neighborhood’.

Margin of Error

Given observed frequency f, what can we say about p?

Let P(p=x|f) be the probability that p=x given observed frequency f

p is most likely f, but of course p could also quite likely be a little smaller or a little larger than f



The margin of error is the percentage me such that there is a 95% chance that p is between f – ME and f + ME. This interval is called the confidence interval.

The greater the sample size, the more narrow this bell curve becomes (i.e. the smaller ME and hence a greater chance that p is approximately f)

How Margin of Error decreases as function of Sample Size

Sample Size (n)	Margin of Error* (ME)
10	30%
25	22%
50	14%
100	10%
250	6%
500	4%
1000	3%
1500	2%

*These values are using a confidence level of 95%.

How to Use the Margin of Error: An Example

- Suppose we poll 1,000 Americans and find that 35% likes how Obama is handling the situation in Afghanistan. I.e. f=35% (or 0.35)
- So, the sample size n = 1,000, giving a margin of error of 3% (see table)
- Hence, we can be 95% certain that the percentage of all Americans that like how Obama is handling the situation in Afghanistan (p) lies somewhere between 32% and 38%.

The Margin of Error is *not* some kind of magical, hard barrier

- Some people think that the margin of error gives us a lower and upper bound for the actual percentage, but that is *not* the case: due to the random nature of picking the sample it is of course still possible (though less than 5% likely) that the actual percentage is below 32% or above 38%.
- So, we can *not* say that we can be *100% certain* that somewhere between 32% and 38% like what Obama is doing in Afghanistan!
- Rather, we say that there is a 95% chance that the actual number (p) lies somewhere between 32% and 38%. Or: we can be *95% confident* that it is somewhere in that interval.

Why Polls usually involve some 1,000-2,000 people.

- Still, the above result is surprisingly precise: We have a population of over 300,000,000 Americans, and yet we only need to poll about 1,500 people to get an estimate of p with a margin of error of only 3%!
- Indeed, this is why typical polls use about 1,000 to 2,000 people:
  - Apparently, that is going to get one a fairly accurate snapshot of all Americans!
  - In fact, going beyond that really doesn't help all that much: polling 1,000,000 Americans really doesn't tell us much more!!
- We could call this the Law of the Not-So-Big Numbers

A Few Technical Notes I

- The values in the table reflect a 95% confidence interval
  - If we used a 99% confidence interval, the margins of error would need to be bigger.
  - In general, the greater the confidence we want, the greater the margin of error will be, and hence the weaker (i.e. less interesting) the claim becomes. In particular:
    - We can be 100% confident that p is somewhere between 0 and 100% (duh!)
    - We are about approximately 0% confident that p is exactly f.

A Few Technical Notes II

- The values have been calculated for f=0.5.
  - The more f deviates from 0.5, the less accurate the values become.
  - Now, they will still work just fine for f being anywhere between 0.2 and 0.8 (and, for reasonably sized n, between 0.1 and 0.9), but once f starts to approximate 0 or 1, the values become unreliable.
  - What happens is that the bell curve will start to get quite skewed for such extreme values.
  - Indeed, it doesn't make much sense to say that where n = 100, and f = 95%, p will be between 85% and … 105%?!

A Few Technical Notes III

- The values have been calculated for a population size of … infinity!
  - So, the Law of the Not-So-Big Numbers is even more surprising: even when the sample is only an infinitely small part than the population we can still obtain highly accurate results about the target population!
  - In fact, the smaller the population, the smaller the Margin of Error. Indeed, once the sample starts to become a significant proportion of the population, the listed margins of error become unreliable (as they are too high)!

A Few Technical Notes IV

- Still, as long as the population is significantly larger than the sample (say, 10 times as large or more), the numbers from the table *are* reliable.
  - And, this is what is normally the case, because if the population was only, say, 3 times the size of the sample, then we might as well have observed the population as a whole, rather than just a sample.
  - Indeed, under most circumstances the listed numbers are not only reliable, but apparently also *independent* of the size of the actual population, whether it is 300 thousand, 300 million, 300 billion, or 300 gazillion: in all cases, a ‘mere’ 1000 size sample gives margin of error of about 3%!

Other Inductive Arguments

- We’ve loked at inductive generalizations, but some of the same comments hold for inductive syllogisms and analogical reasoning:
  - How many cases are we looking at?
  - How representative are those cases?