Stat 136 (Bayesian Statistics)

Lesson 1.1: Main Approaches to Statistics

There are two main philosophical approaches to statistics. The first is often referred to as the **frequentist** approach. Sometimes it is called the **classical** approach. Procedures are developed by looking at how they perform over all possible random samples. The probabilities do not relate to the particular random sample that was obtained. In many ways this indirect method places the *cart before the horse*.

The alternative approach that we take in this course is the **Bayesian** approach. It applies the laws of probability directly to the problem. This offers many fundamental advantages over the more commonly used frequentist approach.

Frequentist Approach to Statistics

The frequentist approach to statistics is based on the following ideas:

- Parameters, the numerical characteristics of the population, are fixed but unknown constants.
- Probabilities are always interpreted as long-run relative frequency.
- Statistical procedures are judged by how well they perform in the long run over an infinite number of hypothetical repetitions of the experiment.

Probability statements are only allowed for random quantities. The unknown parameters are fixed, not random, so probability statements cannot be made about their value. Instead, a sample is drawn from the population, and a sample statistic is calculated. The probability distribution of the statistic over all possible random samples from the population is determined and is known as the sampling distribution of the statistic. A parameter of the population will also be a parameter of the sampling distribution. The probability statement that can be made about the statistic based on its sampling distribution is converted to a confidence statement about the parameter. The confidence is based on the average behavior of the procedure over all possible samples.

Bayesian Approach to Statistics

Bayesian approach to statistics put forward the ideas that:

- Since we are uncertain about the true value of the parameters, we will consider them to be random variables.
- The rules of probability are used directly to make inferences about the parameters.
- Probability statements about parameters must be interpreted as *degree of belief.* The *prior* distribution must be subjective. Each person can have his/her own prior, which contains the relative weights that person gives to every possible parameter value. It measures how *plausible* the person considers each parameter value to be before observing the data.
- We revise our beliefs about parameters after getting the data by using Bayes' theorem. This gives our posterior distribution which gives the relative weights we give to each parameter value after analyzing the data. The posterior distribution comes from two sources: the prior distribution and the observed data.

This has a number of advantages over the conventional frequentist approach. Bayes' theorem is the only consistent way to modify our beliefs about the parameters given the data that actually occurred. This means that the inference is based on the actual occurring data, not all possible data sets that might have occurred but did not! Allowing the parameter to be a random variable allows us make probability statements about it, posterior to the data.

This contrasts with the conventional approach where inference probabilities are based on all possible data sets that could have occurred for the fixed parameter value. Given the actual data, there is nothing random left with a fixed parameter value, so one can only make confidence statements, based on what could have occurred.

Bayesian statistics also has a general way of dealing with a nuisance parameter. A nuisance parameter is one which we do not want to make inference about, but we do not want them to interfere with the inferences we are making about the main parameters. Frequentist statistics does not have a general procedure for dealing with them. Bayesian statistics is predictive, unlike conventional frequentist statistics. This means that we can easily find the conditional probability distribution of the next observation given the sample data.

Both Bayesians and frequentists seek to learn from data, using data to fit models, make predictions, and evaluate hypotheses. Moreover, when working from the same data, Bayesians and frequentists will typically arrive at a similar set of broad conclusions. Yet there are key differences in the logic behind, approach to, and interpretation of Bayesian and frequentist analyses.

Concept	Frequentist interpretation	Bayesian interpretation
probability	the long-run relative *frequency* of a repeatable event (hence “frequentist”")	a measure of the relative plausibility of an event
role of data	data alone should drive our outgoing information	data should be weighed against our incoming information
questions asked	If the hypothesis isn’t correct, what are the chances I’d have observed these data?	In light of these data, what are the chances the hypothesis is correct?