Chapter 7
Sampling Distributions and Estimations

Objectives

After completing this chapter, you will be able to:

- Understand parameters and statistics
- Learn the principles of a good estimations
- Learn sampling distribution

A. Parameters and Statistics

In statistical estimation we use a statistic (a function of a sample) to estimate a parameter, a numerical characteristic of a statistical population. In the preceding discussion of the binomial distribution, we discussed a well-known statistic, the sample proportion pb, and how its long-run distribution over repeated samples can be described, using the binomial process and the binomial distribution as models. We found that, if the binomial model is correct, we can describe the exact distribution of pb (over repeated samples) if we know N and p, the parameters of the binomial distribution. The situation can be diagrammed as in the top of Figure 1 below

A.1. Reversing the Information Flow in Statistical Inference

Probability theory tells us about the long run behavior of pb, but this requires speciÖcation of precisely what we do not know, i.e., p, the proportion in the population. However, what we would like to have is something different from what probability theory provides directly. Generally what we have is one sample, of size N, and what we would like probability theory to provide for us is knowledge about p on the basis of the information in our data. But it doesnít, and so probability theory, as important as it is, provides just the beginning point for statistical inference.

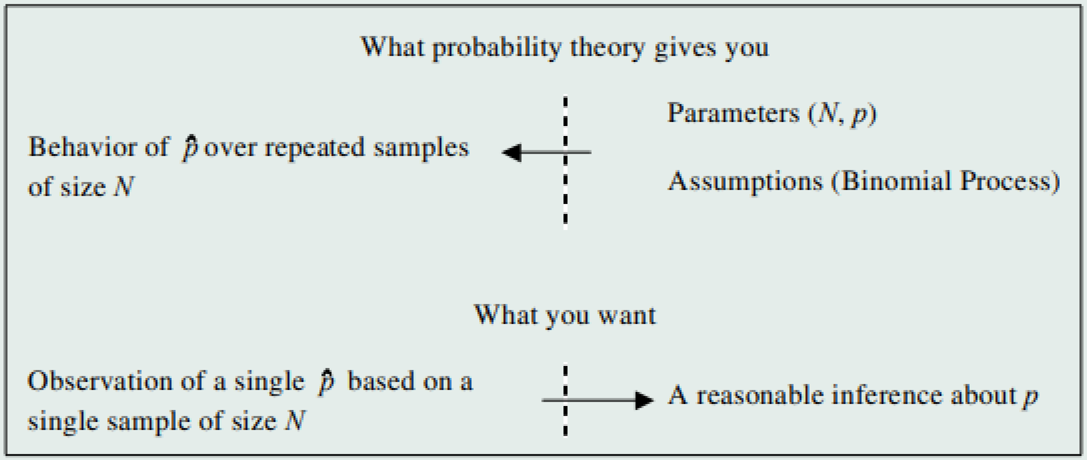

Figure 1: *Reversing the information flow in statistical inference*

1. Looking back at Figure 1, what we would like is to turn around the direction of information flow. In probability theory, knowledge of p and N leads to knowledge about the long run behavior of pb. In statistical inference, we would like something else ó a method to use knowledge of pb and N to lead to knowledge of p.

B. Sampling Distributions

A statistic is any function of the sample. Over repeated samples, statistics will almost always vary in value. So, over repeated samples, a statistic will have a sampling distribution. Sampling distributions have several characteristics:

1. Exact sampling distributions are difficult to derive
2. They are often diffirent in shape from the distribution of the population from which they are sampled
3. They often vary in shape (and in other characteristics) as a functicon of N.

C. Sampling Error

Consider any statistic ø used to estimate a parameter ø. For any given sample of size N, it is virtually certain that ø will not be equal to ø. We can describe the situation with the following equation in random variables.

ø = o+ ∈

where ∈ is called sampling error,and is deÖned tautologically as

D. Principles of Good Estimation

A statistic that is used to estimate a particular parameter is called an estimator of that parameter. A realized value of the estimator is called an estimate of the parameter. Although some of the estimators that we use (like the sample mean X and the sample proportion p) are the sample analogs of the population quantities they estimate, many other estimators (for example, s^2 , the sample variance) are not. For any parameter, there are many possible estimators. Generally, an estimator in wide use has achieved popularity because it satisfies one or more optimality criteria, i.e. qualities that a good estimator is supposed to have. Below, we discuss a number of commonly used criteria for a good estimator.

D.1. Unbiasedness

Ideally, we would like the positive and negative errors of an estimator to balance out in the long run, so that, on average, the estimator is neither high (an overestimate) nor low (an underestimate).

D.2. Consistency

We would like an estimator to get better and better as N gets larger and larger, otherwise we are wasting our effort gathering a larger sample. If we define some error tolerance ∈, we would like to be sure that sampling error " is almost certainly less than ∈ if we let N get large enough. Formally we say the following.

D.3. Efficiency

All other things being equal, we prefer estimators with a smaller sampling errors. Several reasonable measures of ìsmallnessîsuggest themselves: (a) the average absolute error, and (b) the average squared error. Consider the latter. The variance of an estimator can be written

For an unbiased estimator, the sampling variance is also the average squared error, and is a direct measure of how inaccurate the estimator is, on average. More generally, though, one can think of sampling variance as the randomness, or noise, inherent in a statistic. (The parameter is the signal. Such noise is generally to be avoided. Consequently, the efficiency of a statistic is inversely related to its sampling variance,

$$Efficiency(\widehat{\theta}) = \frac{1}{\sigma^2_{\widehat{\theta}}}$$

The relative efficiency of two statistics is the ratio of their efficiencies, which is the inverse of the ratio of their sampling variances.

D. 4. Sufficiency

An estimator is sufficient for estimating O if it uses all the information about O available in a sample. The formal definition is as follows:

The fact that once the distribution is conditionalized on O it no longer depends on O, shows that all the information that O might “reveal in the sample” is captured by bO.

D.5. Maximum Likelihood

The likelihood of a sample of N independent observations is simply the product of the probability densities of the individual observations. Of course, if you donít know the parameters of the population distribution, you cannot compute the probability density of an observation. The principle of maximum likelihood says that the best estimate of a population parameter is the one that makes the sample most likely. Deriving estimators by the principle of maximum likelihood often requires calculus to solve the maximization problem, and so we will not pursue the topic here.

E. Practical vs. Theoretical Considerations

In any particular situation, depending on circumstances, you may have an overriding consideration that causes you to ignore one or more of the above considerations ó for example the need to make as small an error as possible when using your own data. In some situations, any additional error of estimation can be extremely costly, and practical considerations may dictate a biased estimator if it can be guaranteed that a bias can reduce " for that sample.

Video Links:

- https://www.youtube.com/watch?v=z0Ry_3_qhDw
- https://www.youtube.com/watch?v=0ZstEh_8bYc
- https://www.youtube.com/watch?v=Zbw-YvELsaM

References

- http://www.statpower.net/Content/310/Introduction%20to%20Sampling%20Distributions%20and%20Estimation.pdf
- https://crumplab.github.io/statistics/probability-sampling-and-estimation.html
- https://web.njit.edu/~dhar/math661/IPS7e_LecturePPT_ch05.pdf