INTRODUCTION TO PROBABILITY

AND STATISTICS

(Text for 18.05, Spring 1999)

Hartley Rogers, Jr.

Volume I

(CHAPTERS 1 - 9)

TABLE OF CONTENTS

INTRODUCTION

In this book, we first describe some of the basic ideas and methods of the theory of probability. Probability is part of applied mathematics. It uses mathematics to make mathematical pictures or models of happenings in the physical world. These models are used to help describe things that have happened and to predict things that will happen. The models used in probability theory are especially interesting, because we are never completely sure that a probability model is an exact picture of what we are studying in the physical world. The model may differ in various ways from the physical situation that we are describing. How to choose a good probability model, and how to use it to make predictions about physical happenings, are questions that are not always easy, but they make probability one of the most interesting and exciting parts of mathematics.

We shall try, in Chapters 1 through 9, to give a good understanding of some of the basic parts of probability theory. When one has such an understanding, it is easy to go on and study other parts of probability theory as well.

We then turn, in Chapters 10 through 19, to some of the basic ideas and methods of mathematical statistics. Mathematical statistics is also part of applied mathematics. It is concerned with questions of the following kind: given observed data from a physical situation and given a mathematical model for that situation, how well do the observed data agree with the given model?

In statistics, we ask and answer questions of this kind in order to choose good models and to make practical decisions based on those models. Usually, our answer about how well our data may agree with a model will be influenced by a consideration of the various practical consequences that may follow from various practical decisions based on correct or incorrect models. As we shall see, mathematical statistics makes much use of probability theory.

The notion of model is fundamental in our approach to both probability and statistics. In a given physical situation, various different models may be possible. Physical evidence may lead us to believe that some models are better than others, or that some single model is best of all. We must at all times remember the difference between a model and the physical reality that we want it to be a picture of. A model is something that we set up in our minds or on paper. Often it is a simplified picture of reality and leaves out details that are not important to us. As we work with a model in a given physical situation, we may make changes in the model from time to time in order to make the model as useful a picture of the given reality as we can. The kind of model that we shall use in probability and statistics is described in Chapter 2.

The following brief examples suggest some of the ways in which models are used in probability and statistics. In probability theory, given an experimental situation, we choose a model, explore what follows from that model, and make predictions about future observations. For example, if our experiment is the repeated tossing of a coin, we may take as our model a mathematical picture in which heads and tails occur equally often over the long run and in which the outcome of each toss is not influenced by the specific outcomes of previous tosses. Let us call this model μ. A typical probability problem would be to calculate, from the model μ, how long (in terms of tosses) we should expect to wait, on the average, for a run of 10 successive heads to appear.

In statistics, on the other hand, we try to decide from observed data whether a given model is reasonable, or to decide which among some given set of models are most reasonable, or to decide whether any among some given set of models is reasonable. In other words, we try to make <u>inferences</u> about possible models from observed data. For example, if we toss a bent coin 15 times and observe a run of 10 heads from the third to the twelfth toss, we might ask if the model μ (given above) is reasonable in the face of these data.

Inferences about models are the subject of what has come to be called <u>classical statistics</u>. In this book, Chapters 10 through 18 are largely concerned with topics from classical statistics.

Beginning in the 1940s, statisticians increasingly recognized that, in practice, statistical problems often arise from a need to make a practical decision in the face of incomplete information. They came increasingly to believe that it is sometimes better to formulate a statistical problem as a problem of making a best <u>decision</u> rather than as a problem of choosing a best <u>model</u>. This newer approach to statistics is called <u>statistical decision theory</u>.

We consider it in Chapter 19. As we shall see, it requires that we develop a means for numerically measuring the good and bad consequences of different possible decisions.

We suggest some of the ideas of decision theory in the following simple (and artificial) example. An art collector is attending a sale of paintings in a shop. He finds a painting for sale that is claimed to be an old master. He knows that it may be fake. He also knows that further study by experts over a period of months will reveal whether or not it is fake. But if he is to buy it, he must buy it now. He knows from previous experience that only about 25% of the paintings offered in this shop as old masters are in fact genuine. He also knows that there is a special kind of brushwork that tends to characterize old masters. About 90% of the genuine old masters sold in the shop have this special brushwork, but only about 5% of the fakes have it. He examines the painting and sees that it does indeed have this special brushwork. Should he buy the painting?

There are two possible models to consider. The first, call it μ_1, is the mathematical picture in which we assume that the painting is genuine, that 90% of such genuine paintings have the special brushwork, and that 5% of similar fake paintings have the special brushwork. The second, μ_2, is the mathematical picture in which we assume that the painting is fake, that 5% of such fake paintings have the special brushwork, and that 90% of similar genuine paintings have the special brushwork. We have, as data, the observation that the painting has

the special brushwork. Finally, we assume that 25% of the paintings offered are genuine and that 75% are fake. Classical statistics asks the question:
given the data, which of the two <u>models</u> μ_1 and μ_2 is more reasonable? Statistical decision theory asks the question:
what is the best <u>decision</u> to make? In this example, the answer to the latter question will obviously depend upon the rewards and penalties that may follow from the decision made. If the price for the painting is very high, this will count against purchase. If the price is low, this will count in favor. If the collector views the embarrassment of having bought a fake as harmful to his reputation and welfare, this will count against purchase, even if the price is low. If the collector views his own future regret at having missed an opportunity to buy a genuine painting as a serious matter, this will count in favor of purchase even if the price is high. In Chapter 19, we shall return to this example and see how statistical decision theory takes these various considerations into explicit account.

While certain practical problems, like the art collector problem above, cannot be properly treated by classical statistics; while the approach of statistical decision theory may be more fundamental conceptually than the approach of classical statistics; and while there are deep connections between classical statistics and decision theory; it is nevertheless simpler and perhaps more instructive to begin with classical statistics. We therefore put off statistical decision theory to the final chapter.