

PROBABILITY AND STATISTICS

MANJUNATH KRISHNAPUR

Contents

Probability

1. WHAT IS STATISTICS AND WHAT IS PROBABILITY?

Sometimes statistics is described as *the art or science of decision making in the face of uncertainty*. Here are some examples to illustrate what it means.

Example 1. Recall the apocryphal story of two women who go to King Solomon with a child, each claiming that it is her own daughter. The solution according to the story uses human psychology and is not relevant to recall here. But is this a reasonable question that the king can decide?

Daughters resemble mothers to varying degrees, and one cannot be absolutely sure of guessing correctly. On the other hand, by comparing various features of the child with those of the two women, there is certainly a decent chance to guess correctly.

If we could always get the right answer, or if we could never get it right, the question would not have been interesting. However, here we have uncertainty, but there is a decent chance of getting the right answer. That makes it interesting - for example, we can have a debate between *eyeists* and *nosists* as to whether it is better to compare the eyes or the noses in arriving at a decision.

Example 2. The IISc cricket team meets the Basavanagudi cricket club for a match. Unfortunately, the Basavanagudi team forgot to bring a coin to toss. The IISc captain helpfully offers his coin, but can he be trusted? What if he spent the previous night doctoring the coin so that it falls on one side with probability $3/4$ (or some other number)?

Instead of cricket, they could spend their time on the more interesting question of checking if the coin is *fair* or *biased*. Here is one way. If the coin is fair, in a large number of tosses, common sense suggests that we should get about equal number of heads and tails. So they toss the coin 100 times. If the number of heads is exactly 50, perhaps they will agree that it is fair. If the number of heads is 90, perhaps they will agree that it is biased. What if the number of heads is 60? Or 35? Where and on what basis to draw the line between fair and biased? Again we are faced with the question of making decision in the face of uncertainty.

Example 3. A psychic claims to have divine visions unavailable to most of us. You are assigned the task of testing her claims. You take a standard deck of cards, shuffle it well and keep it face down on the table. The psychic writes down the list of cards in some order - whatever her vision tells her about how the deck is ordered. Then you count the number of correct guesses. If the number is 1 or 2, perhaps you can dismiss her claims. If it is 45, perhaps you ought to be take her seriously. Again, where to draw the line?

The logic is this. Roughly one may say that *surprise* is just the name for our reaction to an event that we *a priori* thought to have low chance of occurring. Thus, we approach the experiment with the belief that the psychic is just guessing at random, and if the results are such that under that random-guess-hypothesis they have very small probability, then we are willing to be surprised, that is willing to discard our preconception and accept that she is a psychic.

How low a probability is surprising? In the context of psychics, let us say, $1/10000$. Once we fix that, we must find a number $m \leq 52$ such that by pure guessing, the probability to get more than