# Remarks on the Concept of "Probability"

**Prerequisites**
None

*Inferential statistics* is built on the foundation of probability theory, and has been remarkably successful in guiding opinion about the conclusions to be drawn from data. Yet (paradoxically) the very idea of probability has been plagued by controversy from the beginning of the subject to the present day. In this section we provide a glimpse of the debate about the interpretation of the probability concept.

One conception of probability is drawn from the idea of **symmetrical outcomes**. For example, the two possible outcomes of tossing a fair coin seem not to be distinguishable in any way that affects which side will land up or down. Therefore the probability of heads is taken to be 1/2, as is the probability of tails. In general, if there are N symmetrical outcomes, the probability of any given one of them occurring is taken to be 1/N. Thus, if a six-sided die is rolled, the probability of any one of the six sides coming up is 1/6.

Probabilities can also be thought of in terms of **relative frequencies**. If we tossed a coin millions of times, we would expect the proportion of tosses that came up heads to be pretty close to 1/2. As the number of tosses increases, the proportion of heads approaches 1/2. Therefore, we can say that the probability of a head is 1/2.

If it has rained in Seattle on 62% of the last 100,000 days, then the probability of it raining tomorrow might be taken to be 0.62. This is a natural idea but nonetheless unreasonable if we have further information relevant to whether will rain tomorrow. For example, if tomorrow is August 1, a day of the year on which it seldom rains in Seattle, we should only consider the percentage of the time it rained on August 1. But even this is not enough since the probability of rain on the next August 1 depends on the humidity. (The chances are higher in the presence of high humidity.) So, we should consult only the prior occurrences of August 1 that had the same humidity as the next occurrence of August 1. Of course, wind direction also affects probability ... You can see that our sample of prior cases will soon be reduced to the empty set. Anyway, past meteorological history is misleading if the climate is changing?

For some purposes, probability is best thought of as *subjective*. Questions

such as "What is the probability that Ms. Garcia will defeat Mr. Smith in an upcoming congressional election?" do not conveniently fit into either the symmetry or frequency approaches to probability. Rather, assigning probability 0.7 (say) to this event seems to reflect the speaker's personal opinion --- perhaps his willingness to bet according to certain odds. Such an approach to probability, however, seems to lose the objective content of the idea of chance; probability becomes mere opinion.

Two people might attach different probabilities to the election outcome, yet there would be no criterion for calling one "right" and the other "wrong." We cannot call one of the two people right simply because she assigned higher probability to the outcome that actually transpires. After all, you would be right to attribute probability 1/6 to throwing a six with a fair die, and your friend who attributes 2/3 to this event would be wrong. And you are still right (and your friend is still wrong) even if the die ends up showing a six! The lack of objective criteria for adjudicating claims about probabilities in the subjective perspective is an unattractive feature of it for many scholars.

Like most work in the field, the present text adopts the frequentist approach to probability in most cases. Moreover, almost all the probabilities we shall encounter will be nondogmatic, that is, neither zero nor one. An event with probability 0 has no chance of occurring; an event of probability 1 is certain to occur. It is hard to think of any examples of interest to statistics in which the probability is either 0 or 1. (Even the probability that the Sun will come up tomorrow is less than 1.)

The following example illustrates our attitude about probabilities. Suppose you wish to know what the weather will be like next Saturday because you are planning a picnic. You turn on your radio, and the weather person says, "There is a 10% chance of rain." You decide to have the picnic outdoors and, lo and behold, it rains. You are furious with the weather person. But was she wrong? No, she did not say it would not rain, only that rain was unlikely. She would have been flatly wrong only if she said that the probability is 0 and it subsequently rained. However, if you kept track of her weather predictions over a long period of time and found that it rained on 50% of the days that the weather person said the probability was 0.10, you could say her probability assessments are wrong.

So when is it accurate to say that the probability of rain is 0.10? According to our frequency interpretation, it means that it will rain 10% of the days on which rain is forecast with this probability.