Abdullah Al Mahmud **web**: mahmud.statmania.info **Email**: almahmud.sbi@gmail.com

A Way to Estimate Corruption Rate through Coin Flipping

Abdullah Al Mahmud

It is really difficult to single out people involved in corruption or crime, because these people are also apt at hiding their true face. But, suppose we wish to know, for instance, the proportion of people evading income tax. We cannot just approach people and ask whether they do so, which is supposed to create an embarrassing situation. Also, we cannot always expect a just answer even though they are willing to face the question. But, statistics, the scientific field dealing with uncertainty, can surely show us a way out, providing us with a reliable estimate of the proportion of people evading income tax.

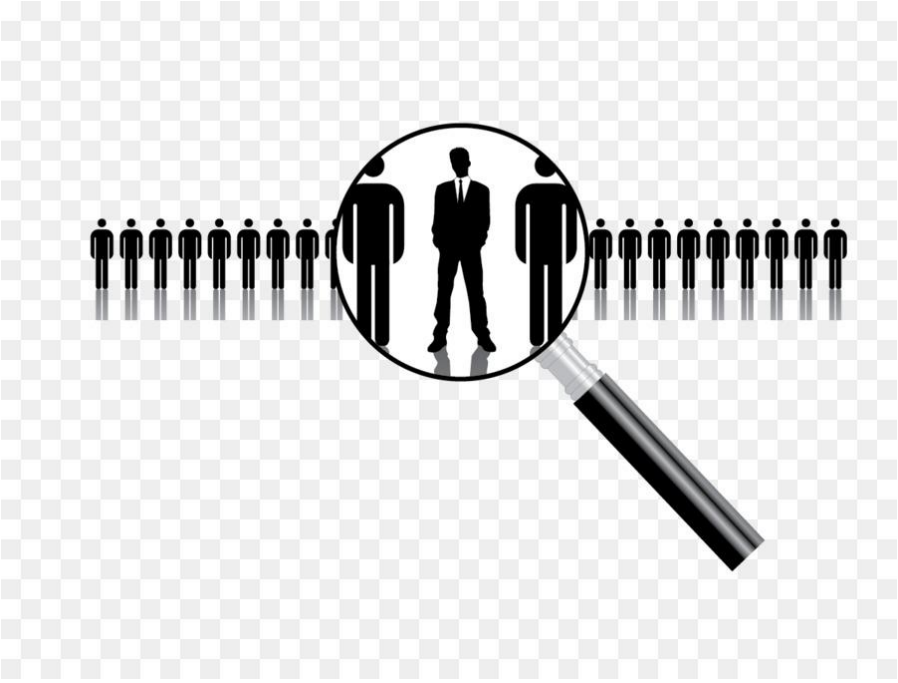
The technique, however, does not allow us to find out the actual people who evade income tax, but it does allow us to compute how many people in an area or a system do so.

The technique employs the concept of probability, the work being started by first selecting a few people for our experiment: the greater the number of people, the better would be the estimate, a thousand people being an adequately large figure. We call these

people the target group. Next, we need a fair coin, implying that when flung, there is equal probability of observing the head or the tail. In other words, according the law of large number in statistics, if the coin is flipped a thousand times, both head and tail are observed approximately five hundred times. We will approach the people of our target group with such a coin.

Let us suppose we have approached Mr. X and ask him whether he evades income tax, but we should not embarrass him. How is that possible? That is perfectly possible: we give him the fair coin, requesting him to flip it, with the instruction to reply in the affirmative (to the question whether he evades income tax) if the head is observed, or to tell the truth if the tail is observed. This ensures his self-respect since we do not learn whether he actually evades income tax, since he can reply in the affirmative for two reasons: either he has observed a head, or he has observed the tail and he actually evades income tax. Hence, we do not learn in which group he actually is. We must make him understand this point; otherwise he would be too embarrassed to participate in the experiment.

Let us suppose, we have experimented with 1000 and individuals and have observed 600 answers as 'Yes' to the question, 'Do you evade income tax?' We have to keep in mind that this figure does not imply 600 out 1000 people evade income tax: as we have already stated, there are some individuals among them who replied in the affirmative just because they observed the head, and some other people replied in the affirmative since they observed the tail and told the truth.



So how many people do evade income tax? Recall that the coin we have used for our experiment was a fair one, with the head and tail equally likely to be observed. Since we experimented with 1000 people, there were 1000 flips with the coin. Since the head and tail are equally likely, we can safely assume that, out of 1000 flips, approximately 500 heads were observed. So out of 1000 people, almost 500 people replied in the affirmative since they observed the head as they were asked. We will never know whether they actually evade income tax. But what about the remaining 100 people (recall that we have assumed to have received 600 answers as 'yes')? They said 'yes' since they observed the tail and they actually evade income tax. Since the coin is fair, there will be approximately 500 tails, providing us with the estimate that 100 out of 500 people i.e. 20 out of 100 people evade income tax. We cannot comment about the remaining 500 people.

Thus, we have attained our estimate: the proportion of people evading income tax is 0.2 i.e. 20 out of 100 people are associated with this corruption.

Things to keep in mind

- We have to inform the people of the target group that we will not know their true answer i.e. whether they actually evade income tax.
- Even though we work with a fair coin, there will not exactly equal number of heads and tails, but the fluctuation will be minimized if the number of people experimented is quite large.
- We can use the technique to estimate of any other corruption.
- Working with 1000 experimental units, our effective sample size is 500. If we instead use an unfair coin having the tail probability of, say 0.7, then our effective will rise to 700, providing us with a better estimate.