Coke or Pepsi?

KEYWORDS:
Teaching;
Binomial distribution
Hypothesis testing;
Experiment.

Maita Levine and Raymond H. Rolwing
University of Cincinnati, USA.

Summary
A binomial experiment dealing with a volunteer's ability to distinguish Coca-Cola from Pepsi-Cola is used to introduce the concepts of hypothesis testing and Type I and Type II errors.

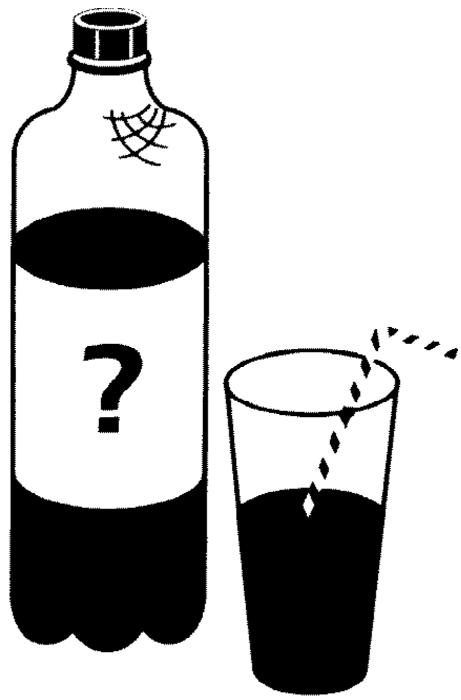
THIS ARTICLE describes an introduction to hypothesis tests which the authors have used many times in an elementary statistics course. The experiment and its analysis have proved to be intellectually stimulating as well as entertaining. The exercise contributed toward overcoming students' perception of statistical techniques as dull "number crunching", and, more important, it provides an effective vehicle for clarifying the basic statistical concepts. After spending a couple of class periods discussing the binomial random variable, we begin the topic of hypothesis testing by asking for a volunteer who can distinguish Coca-Cola from Pepsi-Cola. (There are usually quite a few confident volunteers.) And, the volunteer has always been willing to bet a small amount of money on his/her ability to tell Coke from Pepsi Now, with the money from both student and teacher on the table, we pose the following questions:

1. With what probability can you tell the difference between Coke and Pepsi?
2. What criterion shall we use to determine who will win?
3. Using this criterion:
   a) With what probability will you lose even though your claim is correct? (You burned your tongue this morning; the glasses aren't clean; you get nervous or have lost your concentration.)
   b) With what probability will you win if, in fact, you're guessing and you really cn't tell the difference between Coke and Pepsi? (You have a string of good luck).
4. How small must the probabilities in question 3 be so that each of us will be willing to take the risk?

The prospect of carrying out an experiment in front of the class is sufficiently intimidating to persuade the volunteer that he/she can't distinguish Coke from Pepsi 100% of the time. Usually, the student will settle on a probability between 0.7 and 0.8 as realistic. For our illustrative example, we'll assume 0.7.

Now, we must design the experiment and determine the criterion for success. Remember that we are dealing with a binomial experiment; we are interested in computing the probability of x successes in n independent trials where the probability of success on each trial is p.

Generally, with some suggestions from the class, the volunteer will agree to taste 15 glasses of soft drinks and announce whether each one is Coke or Pepsi. With a 70% probability of being correct, the expected number of correct responses is 10.5. What criterion for success will be fair for both people placing a bet? Ideally, of course, the probabilities discussed in question 3, above, should be fairly small and should be equal or, at least, almost equal.



In an effort to establish a criterion which will be likely to result in success, the volunteer may first suggest that at least 10 correct choices constitutes a win. The student's probability of winning, then, is P (at least 10 successes in 15 trials given that the probability of success on each trial is 0.7) for which a table of binomial probabilities gives 0.722. Now, the teacher is interested in computing the probability of losing his/her money if the student is guessing, i.e., P (at least 10 successes in 15 trials given that the probability of success on each trial is 0.5). This latter probability is 0.151. We have now established that the student will lose even if his/her claim is correct with probability I - 0.722 = 0.278. And, the teacher will lose with probability 0.151 if the student is merely guessing. Under these circumstances are they willing to pay?

With considerable discussion and additional computation, the following conclusions can be drawn:

1. By increasing the required number of correct responses, the volunteer's probability of establishing his/her claim - even though the claim is correct - decreases. That is, his/her probability of losing increases, in spite of having a 0.7 probability of distinguishing Coke from Pepsi.
2. On the other hand, by increasing the required number of correct responses, the teacher's probability of losing, if the student is guessing, decreases. So, changing the criterion will always increase the probability of one type of error and decrease the probability of another type of error.
3. Under the assumption that the volunteer's taste buds continue to function successfully, increasing the sample size will enable the experimenter to maintain small probabilities for both types of errors. For example, if the volunteer is willing to taste 20 glasses of soft drinks (no one is likely to agree to more than 20 trials!), and the criterion for winning is at least 12 successes, the probability that the volunteer, who **can** differentiate between Coke and Pepsi with probability 0.7, will lose is 0.113; the probability that the teacher loses if the student is guessing is 0.252. If the teacher is a good sport, this criterion is reasonable, with only a small amount of money at stake.

Following the discussion relating to this particular problem, one can introduce the formal concepts of hypothesis testing, Type I and Type II errors, and the usual notation associated with these statistical procedures.

The next class period, of course, is devoted to carrying out the experiment. Bring several cans of Pepsi and Coke, and 21 plastic cups. With the volunteer out of the room, use 20 tosses of a coin — heads for Coke, tails for Pepsi — to determine the contents of each of 20 glasses, lined up in a row. Finally, fill the remaining glass with water so that the volunteer can clear his/her taste buds between sips of soft drink.

The authors have won only twice in many Coke-Pepsi experiments. In one instance, the student had, in fact, burned his tongue drinking coffee the morning of the experiment. Apparently, many students can, indeed, tell the difference between Coca-Cola and Pepsi-Cola. Classes have enjoyed the involvement in a "real world" experiment and they have had little difficulty understanding the relevant statistical concepts.