TECHNOLOGY SEMINAR - 05

Sheldon L. Epstein
Bachelor of Science – Electrical Engineering, Massachusetts Institute of Technology

In the last Technology Seminar (TS04), we described how all science begins with a question in a format of: How? What? When? Why? Where? A scientist's or engineer's proposed answer to such a question is called a *hypothesis* (http://en.wikipedia.org/wiki/Hypothesis). Here, we will focus on using a branch of mathematics called *probability theory* (http://en.wikipedia.org/wiki/Probability_theory) to test a hypothesis.

All of us use hypotheses to make every-day decisions. For example, the whole concept of *trademarks* (http://en.wikipedia.org/wiki/Trademarks) or *brand names* (http://en.wikipedia.org/wiki/Brand_names) is based on a hypothesis that consumers will form hypotheses that trademarked or branded products and services have superior qualities over products bearing other trademarks or brands or store-branded or unbranded products and services. Customers rarely subject their hypotheses to scientific tests so marketers may be correct.

An easy introduction to hypothesis testing and probability theory is the study of *gaming*, which is legal, or *gambling* (http://en.wikipedia.org/wiki/Gambling), which is illegal, that is the playing of a *game of chance* (http://en.wikipedia.org/wiki/Games_of_chance) having an uncertain outcome for the specific purpose of winning money or something of value. This is not to say that either *gaming* or *gambling* is necessarily honest or dishonest.

Two *games of chance* with small numbers of outcomes that are easily analyzed include *coin toss* (http://en.wikipedia.org/wiki/Coin_toss) and *dice* (http://en.wikipedia.org/wiki/Dice_game). For *coin toss*, the outcomes are either *head* or *tail*. For *dice* the outcome for a single *die* is a number from the set [1,2,3,4,5,6].

When playing either game, a starting hypothesis is that the coin or the die or dice are *fair*. By *fair*, we mean that a) each of the outcomes is equally likely to occur over a large number of trials and b) the outcome of every trial is *independent* of each and all preceding trials. In fact, this may not be the case.

Necessary conditions for a *fair* coin or die are:

- The material used to make the coin or die must be uniform; that is, *homogeneous* (http://en.wikipedia.org/wiki/Homogeneous) ;
- The *centroid* (http://en.wikipedia.org/wiki/Centroid) of the coin or die must be at the same location as its *center of mass* (http://en.wikipedia.org/wiki/Center_of_mass) ; and
- Every point on the surface of the coin or die must be *symmetric* (http://en.wikipedia.org/wiki/Symmetric) about the *centroid* with another surface point.

While these conditions are necessary, they may not be sufficient. For example, a coin or die containing even a minor amount of *magnetic* (http://en.wikipedia.org/wiki/Magnetic) or *dielectric* (http://en.wikipedia.org/wiki/Dielectric) material inserted in an otherwise homogeneous material could – at least theoretically – be influenced by external magnetic or electric forces.

Suppose that you are invited to wager in a game of coin toss. The gambler who invites you to bet wants you to adopt a hypothesis that the coin is *fair*. How do you test the gambler's hypothesis?

One way to test the gambler's hypothesis is to test the coin by flipping it several times – for example, four-times. You can construct the following table where *head = 1* and *tail = 0*:

TRIAL	***OUTCOME***
0	0000
1	0001
2	0010
3	0011
4	0100
5	0101
6	0110
7	0111
8	1000
9	1001
A	1010
B	1011
C	1100
D	1101
E	1110
F	1111

Note that there are only 16 possible outcomes and they are labeled *0 … F* in *hexadecimal* numbers (http://en.wikipedia.org/wiki/Hexadecimal) . [As an exercise, expand this example for dice.]

Since there are only 16 possible outcomes, then the *probability* that any individual outcome could occur for a *fair* coin is P(X) = 1/16 = 0.0625. If the coin is *fair*, then the outcome *A* is just as likely as the outcome *5* or the outcome *F*. Thus, this simple test of the gambler's suggested hypothesis is insufficient to determine whether the gambler's coin is *fair*.

You could construct a second table, based on the first table, that could be used to generate probabilities of patterns. The following table summarizes the frequency of occurrences of *1s* and *0s*:

# of 1s	***TRIALS***	***P(X)***
0	0	1/16 = 0.0625
1	1,2,4,8	4/16 = 0.2500
2	3,5,6,9,A,C	6/16 = 0.3750
3	7,B,D,E	4/16 = 0.2500
4	F	1/16 = 0.0625

Based on values of *P(X)* in the second table, it would appear that a trial having any of the outcomes [*3,5,6,9,A,C*] would not invalidate or reject the hypothesis that the coin was *fair*. Nevertheless, that would not necessarily lead to a conclusion that the coin was *fair*.

Suppose that our gambler was particularly crafty and that he constructed a coin in which the *center of mass* moved from the *centroid* in a particular way when the coin bounced so that the result of a last toss was opposite that of an immediately preceding toss. Then, outcomes [*A.C*] would predominate and the gambler's suggested hypothesis would not be rejected by a single four-toss trial.

Clearly, such a deception would likely be uncovered if you ran a second trial, a third trial and so on. However, gamblers are too clever to use a ruse that could readily be detected by repeated trials. Besides, it is not necessary for them to use such a ruse in order to earn substantial profits.

Consider the casino game of roulette that is described in great detail at (http://en.wikipedia.org/wiki/Roulette). In America, the game comprises a wheel having 38 slots, numbered [*1 …36, 0, 00*]. If you successfully bet on a number, then your payout is 36X the amount you wagered. If you bet $1, then you would recover your $1 plus $35 in winnings. But, what would happen if you bet $1 on every number? The gambler would receive $38; but, only pay out $36 – leaving a profit of $2 or 2/38 = ~5.3¢ for each $1 bet.

Players do not have to go to a casino to gamble at unfavorable odds. The *Numbers Game,* also known as the *Policy Game*, (http://en.wikipedia.org/wiki/Policy_game) is heavily advertised in poor neighborhoods by gamblers known as bookies. The basis of their game is to have players select a number in the range [000 … 999] that will appear in a newspaper or broadcast. The odds of a player winning are P(x) = 1/1000 or 0.001; however, payoffs are limited to about 500:1 to 800:1. Some gamblers bias their games by excluding popular numbers, such as *777* or telephone exchange numbers, that could result in heavy losses in the event they are selected. Even when gamblers do not exclude numbers, number selection is not necessarily fair as gamblers know in advance which numbers have the heaviest bets. Methods by which gamblers *bias* (http://en.wikipedia.org/wiki/Bias) number selection are time-tested, very subtle and beyond the scope of this seminar.

For the reason described in the previous two paragraphs, it is not necessary for a gambler to win a large sum of money from a single bet to become very rich. For coin-toss and dice, the gambler only needs to *bias* coins or dice by a small amount to win a large amount of money over a long series of trials. The methods by which this is done are also beyond the scope of this seminar.

Despite the fact that the public knows (or should know) that gamblers are in business to make money and that they structure their games to disadvantage players, the players seem to have an insatiable desire to beat the gamblers at their own games. In addition to making fortunes for the gamblers, tales of such activities are ready fodder for movie and book plots – as well as avid imaginations of students – especially mathematics students.

One scenario worth your spare-time reading is a story of some M.I.T. students who undertook to beat the Las Vegas *blackjack* card dealers – and succeeded – for a time. You can start by going to http://en.wikipedia.org/wiki/Card_counting . Enjoy!

This Technology Seminar note is at http://www.k9ape.com/publicservice/PSM/TS05.pdf. The INTERNET version contains active URL links for your convenience.