A Statistical Approach to the Problem of Division

C. Andy Tsao
Department of Applied Mathematics
National Dong Hwa University
Hualien 974, Taiwan
Email: chtsao@cc.ndhu.edu.tw

Yu-Ling Tseng
Department of Applied Mathematics
National Dong Hwa University
Hualien 974, Taiwan
Email: yltseng@cc.ndhu.edu.tw

Abstract

The problem of division is investigated via a statistical approach. An observabalist, operationalist viewpoint has been taken to construct a de Finetti type estimator. This estimator is then compared with other contenders such as the classical estimators, frequentist estimators, Bayes estimators, etc. A brief discussion of the unconventional notion of risk will be given.

The Problem

The problem of division is one of the most important problems in the emergence of probability. In 1494 Luca Pacioli cast the problem in this form:

A team plays ball in such a way that a total of 60 points is required to win the game, and each goal counts 10 points. The stakes are 10 ducats. By some accident they cannot finish the game, and one side has 50 points and the other 20. One wants to know what share of the prize money belongs to either side.

The readers are referred to Hacking (1985) for the historic importance in the emergence and development of probability of this problem. It has been long considered "solved" from a probabilistic viewpoint. The classical solution for the game described above is to divide the stake to 15:1 and the larger share goes to the one in the lead assuming each player has equal chance of winning.

A Statistical Approach

We do not find the classical solution satisfactory because it does not utilize the current winning scores sufficiently. The main idea behind this project is to consider the problem of division as a statistical decision problem which makes use of the available observation (i.e., the scores). Precisely, consider two players: player A and player B and the game is determined when player \( A, B \) score \( N_A \) and \( N_B \) respectively. For simplicity, we assume that \( N = N_A = N_B \). The current status of the game is denoted by \( s = (n_A, n_B) \) and let \( n = n_A + n_B \). The problem of division, denoted by \( (n_A, n_B, N) \), is to find fair share of each player given \( s = (n_A, n_B) \) and \( 0 < n_A, n_B < N \). From the statistical viewpoint, the following questions arise naturally

- What is the odds of winning of each player conditioning on the current status or estimated odds based on the data?

What would be the outcome predicted by the data?

What is the range of odds based on the insufficient knowledge (non-determined game)?

The different foci will presumably lead to different solutions. And these solutions can be different from the classical solutions. In our opinion, there may be no unique good solution to this problem and the correctness or rather the optimality of the solution depends on the interpretation of the problem and the criteria of concern.

**Results**

Here we consider the problem of division as a statistical decision problem. The outcome of the \( i \)-th match is denoted as \( X_i, i = 1, \cdots, n \) which takes value 1 if player A wins; 0 if otherwise. Note that \( n_A = \sum_{i=1}^{n} X_i \). Also, let \( a = N-n_A, b = N-n_B, m = a+b-1 \) and \( X(1..n) = (X_1, \cdots, X_n) \). Given a \((n_A, n_B, N)\) problem of division, it is desired to predict/estimate the possible value of

\[
(1) \quad w(X(n + 1..n + m)) = \sum_{i=a}^{m} 1_{[S_{n+m} - S_n = i]},
\]

where \( 1_E \) is the indicator function of an event \( E \) and \( S_n = \sum_{i=1}^{n} X_n \). Using the current score \((n_A, n_B)\), this problem can be viewed as a point estimation problem of finding suitable estimator \( \hat{w}(x(1..n)) \) for \( w(X(n + 1..n + m)) \) for \( w(X(n + 1..n + m)) \).

The operational, subjectivist approach to statistical inference has given a sound theoretical ground by de Finetti’s representation theorem, see for example, de Finetti (1990), Bernado and Smith (1993). We employ this approach and derive a solution to the problem of division. The solution is then compared with other contenders such as the classical solutions, frequentist plug-in estimator, Bayes estimators, etc. Brief discussion will be given on the notion of risk function because it is not as straightforward as in the usual decision theoretical problems.

**REFERENCES**

