Abstract Bracketology 2

Barry Wright, III

July 3, 2008

1 Basic Multiround Theory

To begin our analysis of larger games, we examine some general properties of a tournament involving n rounds. For now, we disregard the possibility of any play-in games, assuming a strict single-elimination format. In this case, it is clear that there are 2^n teams participating, and $2^n - 1$ games to be played. By definition of the standard rules, the number of points awarded for a correct prediction in round j is equal to 2^{j-1}, yielding the standard $1, 2, 4, 8, ...$ scoring format.

First, we consider the possible "brackets" which can be played in such a tournament. Given a particular round, there are 2^{n-j} games to be predicted, each with two valid choices. (Note that in many cases, the two valid choices may both be wrong, given incorrect predictions earlier in the tournament, an artifact of the fact that the entire tournament is predicted before play begins, rather than a round by round format.) We make the following definitions;

Definition: A *j-round profile* is a set of predictions for the games in round j. A *tournament profile*, commonly referred to as a "bracket," is a collection of j-round profiles, for $1 \le j \le n$. These j-round profiles must be *consistent*, such that the predicted winners of round $j - 1$ are the valid choices for round j.

Now since we have 2^{n-j} games in round j, and two choices per game, we have $2^{2^{n-j}}$ j-round profiles. The number of possible tournament profiles is then;

$$\#Tournament\ profiles = \prod_{j=1}^{n} 2^{2^{n-j}} \quad (1)$$

$$= 2^{\sum_{j=1}^{n} 2^{n-j}} \quad (2)$$

$$= 2^{2^n - 1} \quad (3)$$

For the standard NCAA tournament, this is 2^{63} possible brackets, and including the play-in game gives $2^{64} \approx 1.84 \times 10^{19}$ total tournament profiles. While a very large number of these profiles are never played (such as ones with play-in participants being the overall winner), it is apparent that having identical brackets is extremely rare, and thus selecting a completely correct bracket is nearly impossible.

Let us now consider the relative importance of correctly selecting the overall winner of a tournament, as well as the last 2 and last 4 participants. We will first study this for a general n round tournament, and then apply our results to determine some important quantities for the standard 6-round tournament. One important quantity is the total number of points available in a tournament. Consider a round j. The total number of points should be the number of games multiplied by the number of points per correct game. This yields;

$$2^{n-j} \times 2^{j-1} = 2^{n-1} \tag{4}$$

We note that this is actually independent of the round j, so that the total number of points available in a tournament is;

$$\sum_{j=1}^{n} 2^{n-1} = n2^{n-1} \tag{5}$$

Now, although each round can score an equal number of points, the ability to score points in later rounds is dependent upon success in previous rounds, so we should not consider the rounds to be of equal importance unless given a stronger mathematical reason to do so. This fact becomes apparent when we calculated the number of points guaranteed by correctly choosing the overall winner. Correctly choosing the winner of the final game is worth 2^{n-1} points by itself. However, having the opportunity to make the correct choice implies that correct choices were made in each previous round (by the consistency requirement). Thus, the total number of points guaranteed by a correct choice of overall winner is given by;

$$\sum_{j=1}^{n} = 2^n - 1 \tag{6}$$

This allows us to calculate the maximum advantage generated by correctly choosing the overall winner, which we denote MAX^1 (the fraction of points lost by selecting the overall winner to lose in the first round);

$$MAX^1 = \frac{2^n - 1}{n2^{n-1}} = \frac{2^n}{n2^{n-1}} - \frac{1}{n2^{n-1}} \tag{7}$$

$$= \frac{2}{n} - \frac{1}{n2^{n-1}} \tag{8}$$

We see that MAX^1 quickly ($n \geq 4$) approaches the very simple quantity $\frac{2}{n}$, and that as n approaches infinity, both this and the exact quantity converge to zero. [GRAPHICAL INTERPRETATION OF THIS] Considering the standard 6-round NCAA tournament, we see that MAX^1 is approximately $\frac{1}{3}$, with exact value $\frac{63}{192} = .328125$, a significant fraction of the total possible score.

Now this quantity is often not applicable, because it is rare that people select the overall winner to lose in the first round. Instead, we consider the minimum nonzero advantage of correctly choosing the overall winner, which we denote MIN_1^1 (the subscript will be important when considering the effects of correctly selecting the last 2 and last 4 participants);

$$MIN_1^1 = \frac{2^{n-1}}{n2^{n-1}} = \frac{1}{n} \tag{9}$$

This quantity corresponds to the advantage gained over someone who chooses the overall winner to lose in the final round, and we see that it is equal to one round's points, which makes intuitive sense.

We can extend a similar analysis to consider the advantages gained by correctly choosing the last 2 and last 4 participants. For tournaments with $n \geq 2$, consider the fraction of points guaranteed by correctly selecting the final two participants (no assumptions are made about the selection of the final game). Again, this represents the maximum advantage for correctly choosing the final two participants, which we denote MAX^2;

$$MAX^2 = \frac{2(2^{n-1}-1)}{n2^{n-1}} = \frac{2(2^{n-1})}{n2^{n-1}} - \frac{2}{n2^{n-1}} \tag{10}$$

In general, we can derive the maximum advantage for correctly choosing the final 2^k participants;

$$MAX^{2^k} = \frac{2^k(2^{n-k}-1)}{n2^{n-1}} = \frac{2}{n} - \frac{2^k}{n2^{n-1}} \tag{11}$$

Note that when $k = n$, the value goes to zero, which makes sense as we get no advantage for correctly writing down the teams which are participating in the tournament. Now, these values differ from the maximum possible disadvantage in incorrectly selecting the final 2^k participants. This is because incorrect selection of all of the final 2^k participants implies incorrect selection for the final $2^{k-1}, 2^{k-2}, ..., 1$ participants. There was no effect in the case of MAX^1, as there are no rounds after the final round. Taking this into account, we can calculated the maximum disadvantage, associated to incorrectly selecting the final 2^k participants to lose in the first round, which we denote $MAXD^{2^k}$:

$$MAXD^{2^k} = \frac{2^n - 1 + (2^{n-1} - 1) + 2(2^{n-2} - 1) + ... + 2^{k-1}(2^{n-k-1} - 1)}{n2^{n-1}} \tag{12}$$

$$= \frac{2^n - 1 + \sum_{j=0}^{k-1} 2^j (2^{n-1-j} - 1)}{n2^{n-1}} \tag{13}$$

$$= \frac{2^n - 1 + \sum_{j=0}^{k-1} 2^{n-1} - \sum_{j=0}^{k-1} 2^j}{n2^{n-1}} \tag{14}$$

$$= \frac{2^n - 1 + k(2^{n-1}) - (2^k - 1)}{n2^{n-1}} \tag{15}$$

$$= \frac{2 + k}{n} - \frac{2^k}{n2^{n-1}} \tag{16}$$