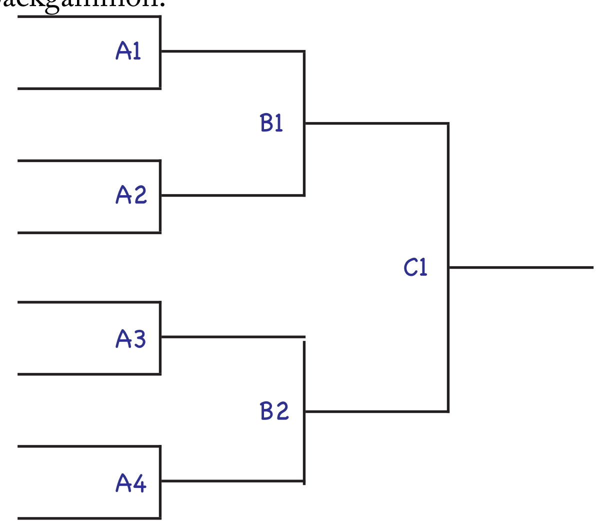
Skill and luck

One of the most important things to keep in mind while designing a tournament is the degree to which luck influences the outcome of individual matches. The greater the influence of luck, the less the outcome of any one match has to say about which player is really better.

Luck need not be the result of some explicit randomizing device, like shuffled cards or thrown dice. Baseball, a game with no cards or dice is, notwithstanding, a very high-luck game. There is a lot of luck in baseball because the outcome of a play often depends critically on physical differences so small that they are, for the most part, beyond skillful control. The difference between a home run and an out can be a matter of a millimeter or so on the path of a swung bat. This is not to say that baseball is not also a game of skill—it most certainly is. But it is a high-luck game because elements that are beyond skillful control are often more determinative of the outcome than those that are.

Throughout this book, I'll frequently talk in terms of high-luck events and high-skill events, reporting experimental results from simulated tourneys to show how expectations differ between the two. Perhaps the best way to understand the difference is this: a high-skill event is something like tennis, and a low-skill event is something like backgammon.

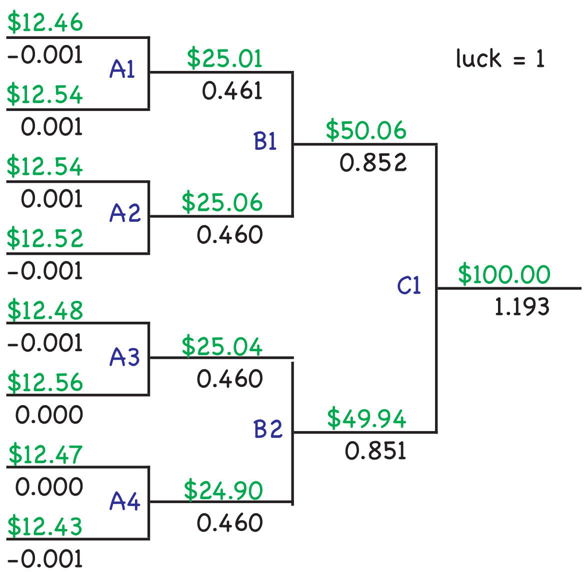


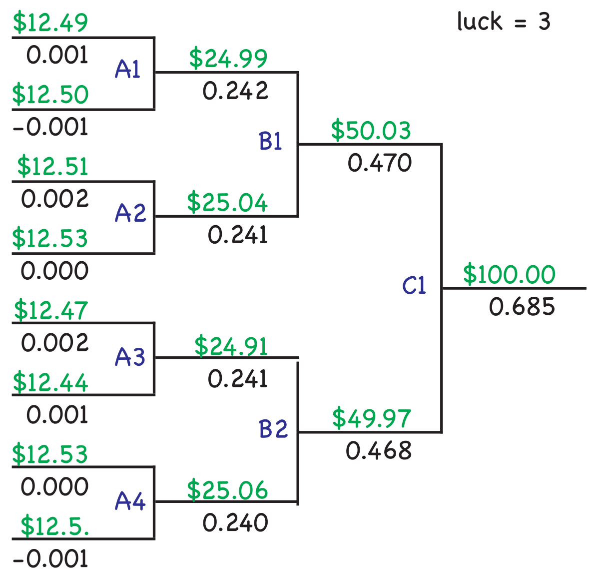
The details of my tournament simulator and the match model it uses are hidden away in an appendix, where only the geekiest readers will see them. For everyone else, let me offer a few statistics from an exceedingly simple kind of tournament to help give you a feel for the difference between high-luck events and high-skill events. At the right is a bracket for an eight-player, single elimination tournament. The tourney consists of seven matches in three rounds: four in round A, two in round B, and one in round C.

Now let me introduce the idea of an analyzed bracket. This is a bracket that's been annotated with the results of a simulation experiment. For each match in the bracket, the analyzed bracket shows the average winnings for the players who reach that line—this is the number in green, above the line. For these brackets, I'm assuming that the total prize fund was $100, and that it all went to the winner. $100 is a convenient number because it makes it possible to interpret the green numbers as percentages.

Below the line, in black, is the average skill of the players reaching that line. In real tournaments, of course, you never know precisely how good any player it. But in simulated tournaments, skill is simply a number. The skill numbers used here are Z scores, generated at random from what is assumed to be a full normal (Gaussian) distribution. The average skill level of an entrant is zero.

Here are two analyzed brackets for the eight-player single-elimination tournament. They differ only in the balance of skill and luck:





The bracket on the left is for a high-skill competition. This corresponds to the parameter luck = 1, which means that in the simulator, skill and luck have an equal role in determining the outcome of an individual match. The right-hand bracket is for a high-luck competition, in which the luck = 3, which means that skill accounts for only 25% of the outcome of a match.

The numbers above and below the lines are averages from a simulation run of 1,000,000 trials. The green numbers show the expectation of a player on that line. They don't differ meaningfully between the two brackets. In each, the players in the A round earn about one-eighth of the $100 prize fund, or $12.50. Players that win a round and ascend to the B lines earn about $25, and the finalists in the C round earn about $50. The winner takes the entire prize fund of $100.

Note that even after a million trials the prize money figures vary a little from line to line within a round. With enough trials, the numbers should converge to exactly $12.50, $25, and $50, but that would take a very large number of trials indeed. The fact that individual trials each award the whole $100 to the winner and $0 to everybody else makes the averages a bit unstable.

Skill progression

The key difference between the two brackets is shown by the black numbers below the lines. These numbers are the average Z scores representing the mean skill level of the players occupying that line. These numbers vary less for lines within a round than the prize money numbers, but they still vary a bit even after a million trials.

In both cases, the average skill for round A is zero, which represents the average skill level for players in general. To get to the B round, a player has to win a match, and a player that's won a match is, on average, better than a player who hasn't. For the high-skill tourney, the B round player has an average skill level of about 0.460, while for the high-luck tourney, that skill level is only about 0.240. Players that win twice are, on average, better still, with scores of about 0.852 and 0.469 for the high-skill and high-luck tourneys, respectively. The average score of the winner of the tourney is 1.193 and 0.685.

In both cases, the average skill of the best player is the same, about 1.423. But the best player doesn't always win. In the high-skill contest, the best player wins the tourney about 59% of the time, but in the high-luck contest, the best player prevails only about 32% of the time.

A sequence of average skill scores from one round to the next is called a skill progression. The skill progression for a tourney where luck = 0 (thus, the better player always wins) is this one:

Luck = 0: A - 0.000 → B - 0.564 → C - 1.029 → W - 1.423

which represents he maximum possible skill progression. As above, we have smaller skill progressions for increasing luck values, and no skill progression at all for a game with no skill at all:

Luck = 1: A - 0.000 → B - 0.469 → C - 0.852 → W - 1.193

Luck = 3: A - 0.000 → B - 0.240 → C - 0.469 → W - 0.685

Luck = ∞: A - 0.000 → B - 0.000 → C - 0.000 → W - 0.000

The skill progression really doesn't matter very much for the simple eight-player, single-elimination tourney. It does affect the fairness of a tourney (specifically, fairness (C)), and we'll discuss that presently. But there's nothing about it that would suggest that the bracket be drawn any differently.

As we'll see, though, where there's more than one bracket, the skill progression is the guide that will help us manage the way that brackets interact with each other. This means that we will (or at least should) draw brackets differently depending on the luck factor of the underlying game. The best bracket for, say, backgammon will not always be the same as the best bracket for tennis.