

Cricket Winning Probability Calculation Model - Second Innings

Given the target to be chased, runs made so far, balls bowled and wickets fallen, this model calculates the probability of runs being chased by the batting team using data from past matches. The work involves a dynamic programming framework similar to the score prediction model but is a little more complicated.

Let $P(r, b, w)$ be the probability of winning (i.e. runs chased) when r runs need to be made, b balls have been bowled and w wickets have fallen. Further, let $p(b, w)$ be the probability of losing a wicket on the ball no. $b+1$ when w wickets have already fallen. Valid runs mostly scored in a ball in a game of cricket are 0, 1, 2, 3, 4, 6. For each of the runs in the set, we use $q(r, b, w)$ to denote the probability that r runs are scored in the ball no. $b+1$ when w wickets have fallen. We can then write:

$$\begin{aligned} P(r, b, w) = & P(r, b+1, w+1) * q(0, b, w) * p(b, w) + P(r, b+1, w) * q(0, b, w) * (1 - p(b, w)) + \\ & P(r-1, b+1, w) * q(1, b, w) + P(r-2, b+1, w) * q(2, b, w) + P(r-3, b+1, w) * q(3, b, w) + \\ & P(r-4, b+1, w) * q(4, b, w) + P(r-6, b+1, w) * q(6, b, w) \end{aligned} \quad (1)$$

The basic idea is to make the model work backwards and impose a boundary condition. There are also certain assumptions involved to make the calculation easy such as extra runs and invalid balls are not taken into account. Also, no runs are made on a ball on which a wicket is lost.

Boundary Conditions

$$\begin{aligned} P(r, 300, w) &= 0 \text{ for } r > 0 \\ P(r, b, 10) &= 0 \text{ for } r > 0 \\ P(0, b, w) &= 1 \end{aligned}$$

The above three conditions are quite straightforward. If no additional runs need to be made that is if $r = 0$, you always win. And if there are no balls left or no wickets left, you always lose. Using these conditions, the model can be solved backwards to calculate the entire P matrix.

Pranav Sodhani
Email: sodhanipranav@gmail.com