**Using an RE24 Variant to Determine the Value of Intentional Walks**

Barry Bonds has been intentionally walked more than twice as much as any other player in the history of baseball. Even if you just account for his age 37, 38, and 39 seasons he would rank 6th all time, right between Vlad and Junior. You can find however many absurdities you want with this season, the best season by any player of all time. Here’s one: on May 28th that year, the Diamondbacks walked him with the bases loaded in the bottom of the ninth.

What kind of decision making might a manager use to decide to walk Bonds? There are some obvious situations that seem obvious like, up 2 bottom of the ninth, 2 outs, runner on 2nd, Bonds at bat. You’re definitely walking him there. But how might a manager make these decisions in more ambiguous situations? How does a manager come to the conclusion that they should walk in a run rather than let Bonds bat?

In this article I develop a variant on the run expectancy matrix that shows up in the RE24 computation. This variant, which is achieved via something called Monte Carlo simulation, lets a manager know directly ‘walking Bonds here has the other team score fewer runs on average’ or ‘walking Bonds here decreases the probability of the Giants scoring 2 runs this inning’.

**The 24-State Run Expectancy Matrix**

The base/out run expectancy matrix tells us how many more run a team is expected to score in an inning given who is on base and how many outs there are. Take a look (data from tangotiger.net)

Here is how I like to think of things. Every at-bat transitions from one state to another. You strike out with a guy on first, your at-bat caused the transition from ‘runner on 1st, 1 out’ to ‘runner on 1st, 2 outs’. Because you caused a decrease in the expected number of runs your team will score in the inning from 0.509 to 0.224. In RE24, you then get credit for *negative* 0.285 runs.

Putting everything in terms of purely `expected runs’ lets us compare apples to oranges, so to speak. We can use the table to help decision making in-game. For example, we can (and did! (hyperlink)) discuss the value gained by stealing bases and compute the steal percentage you need for the attempt to be ‘worth it’. You could look at the value gained or lost by sacrifice bunts.

Unfortunately, what you can’t do is something like ‘evaluate the current situation and tell me if I should intentionally walk Barry Bonds here’. This is because the run-expectancy matrix is league-average, something Barry certainly is not. No matter what, adding a runner to first base always *increases* a team’s expected runs according to the table.

To determine the value of intentional walks, you need a way to compute run expectancy for individual players. Then we can maybe say ‘The Giants run expectancy with 1 out, a runner on 1st, and Bonds at-bat is larger than the expectancy with 1 out, runners on 1st and 2nd, and Edgardo Alfonzo at-bat’. Instead of 24 base/out states, we’ll propose computing the expected inning runs for 216 base/out/player states for a given lineup.

**A Side Story**

I grew up a Cincinnati Reds fan, though unfortunately was only one year old the last time we won a playoff series. I wasn’t around for the big red machine or for the nasty boys. I don’t remember Griffey Sr., the closest I got to his son was going to the same high school. My fandom was fostered by my Dad telling me stories from a long time ago.

My Dad loved telling me about Pete Rose, about his hustle, his drive, everything. The prime example was from an otherwise unmemorable game. It was May 14, 1972, the first of a double header. The Giants decided to intentionally walk Pete Rose and if this happened in 2020, they would have succeeded. However, Pete Rose had other intentions, took a swing at the 3-0 delivery, and reached on an extremely understandable error. Then, in the most Pete Rose turn of events possible, he was thrown out at second after trying to extend his single into a double.

**Simulating at Bats the Monte Carlo Way**

Baseball has the good fortune of being extremely easy to simulate. If I want to simulate on of Barry Bond’s 2004 at bats, I essentially roll a weighted dice. This particular dice has a 37.6% chance of coming up ‘Walk’, a 9.7% chance to come up ‘single’, etc. Then, I can simulate entire innings by giving each player in the order their specific dice tuned to their specific stats and rolling them consecutively.

What I am doing above is called *Monte Carlo Simulation*. Simply put, if something is too hard to compute the average of, just simulate it hundreds of thousands of times and compute the average of your simulations. It turns out, the average of your simulations will be extremely close to the average you want to compute. This technique is so reliable that applied mathematicians use it to decide what stocks to buy, to predict the weather, and to decide how much to charge for interest. In a previous article we used it to simulate different batting orders to show that the choice of batting order doesn’t really matter (hyperlink).

**The 216 State Run Expectancy Matrix**

Suppose we want to figure out the expected runs the Giants will score with Barry Bonds at bat, 1 out, and a runner on 2nd. What we do is simulate an inning starting in exactly this setting then proceed through the batting order rolling dice until we reach three outs. Then we count how many runs were scored, do it again a couple thousand times, then compute the average. In the following experiments we used the 2004 stats

|  |  |  |  |
| --- | --- | --- | --- |
| Base State | No Outs | One Out | Two Outs |
| Empty | 0.464 | 0.308 | 0.182 |
| Runner on 1st | 0.880 | 0.638 | 0.395 |
| Runner on 2nd | 0.994 | 0.761 | 0.485 |
| Runners on 1st and 2nd | 1.427 | 1.109 | 0.743 |
| Runner on 3rd | 1.118 | 0.882 | 0.549 |
| Runners on 1st and 3rd | 1.501 | 1.206 | 0.787 |
| Runners on 2nd and 3rd | 1.659 | 1.303 | 0.829 |
| Bases Loaded | 2.25 | 1.761 | 1.206 |

Note: Because we wanted to test the effects of walking/not walking Barry Bonds, we computed Barry Bonds percentages for single, double, etc. by discounting all the intentional walks. This way, we get a true measure of how much damage he will do without being walked.

Now, the next guy in the order is Alfonzo, so we do the same thing for him and compute his particular run expectancy matrix.

|  |  |  |  |
| --- | --- | --- | --- |
| Base State | No Outs | One Out | Two Outs |
| Empty | 0.299 | 0.179 | 0.083 |
| Runner on 1st | 0.605 | 0.389 | 0.186 |
| Runner on 2nd | 0.752 | 0.534 | 0.313 |
| Runners on 1st and 2nd | 1.038 | 0.784 | 0.420 |
| Runner on 3rd | 0.905 | 0.657 | 0.394 |
| Runners on 1st and 3rd | 1.192 | 0.871 | 0.519 |
| Runners on 2nd and 3rd | 1.362 | 1.012 | 0.625 |
| Bases Loaded | 1.719 | 1.253 | 0.796 |

**When Should Bonds be Walked?**

Intentionally walking Barry Bonds switches from using *his* run-expectancy matrix to using Alfonzo’s. How can we decide whether or not to walk Barry Bonds? Let’s pick a specific example. Suppose the current game state is ‘Two Outs, runner on 2nd, and Barry Bonds at-bat’. On average, we expect the Giants to score 0.485 runs in this situation. If we walk Bonds intentionally, we transition to the ‘Two Outs, runners on 1st and 3nd, and Edgardo Alfonzo at-bat’. Notice that this sitatuon has an expected runs value of 0.420. Walking Barry Bonds here actually saves us an average of 0.065 runs. Here, it would be a good idea to walk Bonds.

The following table shows, for each base and out situation, the amount of runs saved by walking Barry Bonds to get to Alfonzo. Notice that positive values are in green and represent run savings: these are good times to walk him.

|  |  |  |  |
| --- | --- | --- | --- |
| Base State | No Outs | One Out | Two Outs |
| Empty | -0.141 | -0.081 | -0.004 |
| Runner on 1st | -0.158 | -0.146 | -0.025 |
| Runner on 2nd | -0.044 | -0.023 | 0.065 |
| Runners on 1st and 2nd | -0.293 | -0.14 | -0.053 |
| Runner on 3rd | -0.07 | 0.01 | 0.029 |
| Runners on 1st and 3rd | -0.22 | -0.04 | -0.009 |
| Runners on 2nd and 3rd | -0.06 | 0.05 | 0.033 |
| Bases Loaded | -0.48 | -0.49 | -0.581 |

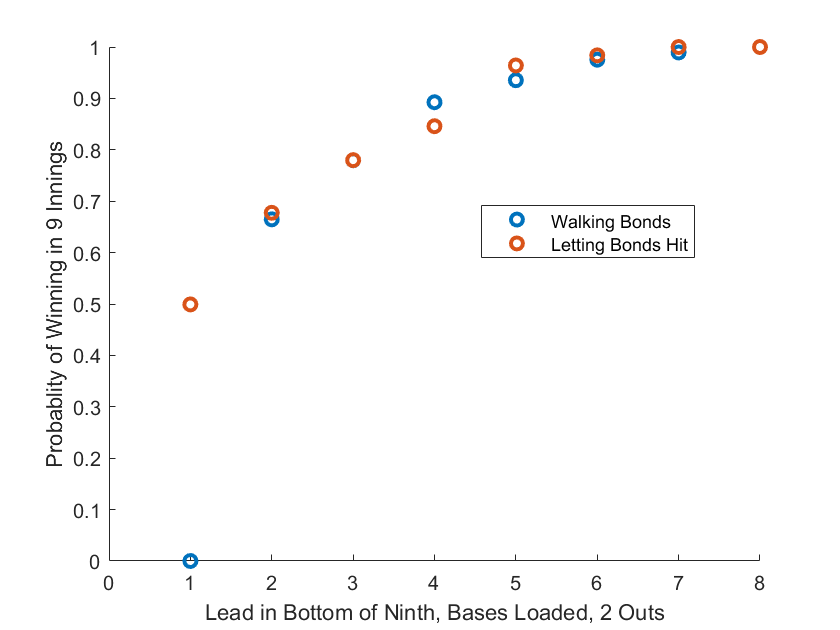
This analysis only tells you how many runs on average you save by walking Bonds. It makes sense that walking a guy with the bases loaded is not a good idea in a vacuum. However, if we want to understand perhaps *why* the Diamondbacks walked him with the bases loaded, we need to look at the situation a little more closely.

**Situational Run Expectancy**

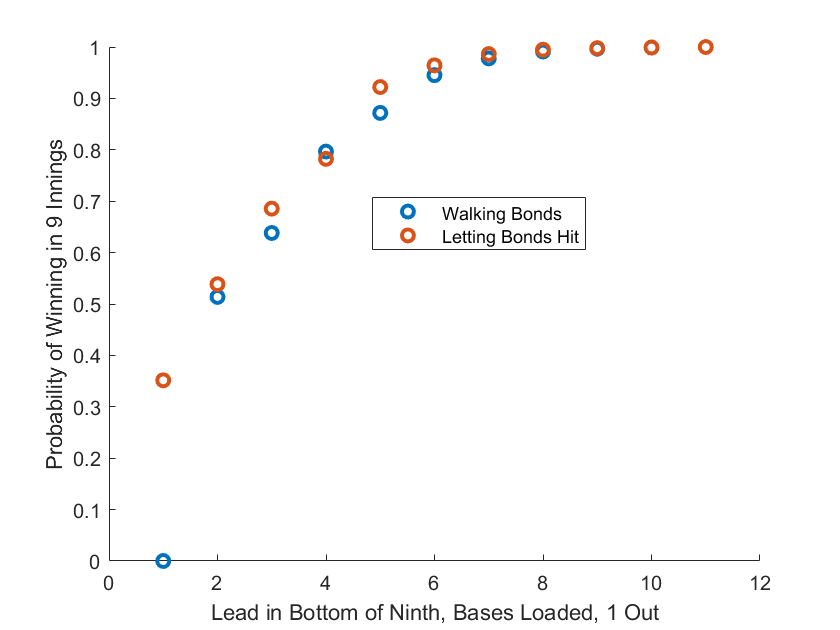
The above suggests that you should never even consider walking Barry Bonds with the bases loaded. Perhaps, though, there are situations when you may want to. The idea behind an intentional walk late in the game is not necessarily to scrape together every ‘tenth of a run’ advantage you can find. Sometimes you just want to prevent the big play.

Can throwing an intentional walk with the bases loaded actually increase your chances of winning? Here’s what we did to check. Suppose you are playing the giants and Bonds is coming up to bat with the bases loaded and 2 outs. Should you walk him or let him hit?

It turns out, you should only walk him if you are up by exactly four runs in this situation. The chart below shows your team’s win probability for a given lead. The blue dots are for walking Barry Bonds, the red letting him hit. The dot which is higher is the one that is a ‘better choice’ for your team to do.



So it doesn’t really seem like the intentional walk is a good idea in this case. What if there are fewer outs. Does that change things? Let’s re-run the numbers with one out and see what it says.



The data for the 0-out case is quite similar. The conclusion is this: It pretty much only makes sense to walk Barry Bonds with the bases loaded if you are up precisely four runs in the bottom of the ninth. Each other case, the intentional walk would decrease your team’s chance of winning.

**Conclusions**

While this article was largely about the decision making process when it comes to intentional walks, the true focus was how to use the run expectancy matrix (and a variant) for decision making in-game. This information would be invaluable to a manager trying to choose the lesser of two evils in a no-win situation. This technique could help managers defend their decisions. It could give them new ideas they might never have had. The point is that sabermetrics aren’t limited to describing player’s abilities. Team’s should be using statistical analysis as much as possible throughout all aspects of their games.