Bayesian Inference: Welcoming Subjectivity in an Objective Framework

Zachariah Zanger M.S.

University of San Francisco Baseball Quantitative Studies Team

Introduction

In Judea Pearl's exceptional book on recent breakthroughs in Cause and Effect Inference, *The Book of Why: The New Science of Cause and Effect* he writes about the English statistician, philosopher, and minister Thomas Bayes:

> "A Presbyterian minister who lived from 1702 to 1761, the Reverend Thomas Bayes appears to have been a mathematics geek. In an article published after his death, Bayes tackled a problem that was the perfect match for him, pitting math against theology. To set the context, in 1748, the Scottish philosopher David Hume had written an essay titled "On Miracles," in which he argued that eyewitness testimony could never prove that a miracle had happened. The miracle Hume had in mind was, of course, the resurrection of Christ. Hume's main point was that inherently fallible evidence cannot overrule a proposition with the force of natural law, such as "Dead People stay dead." For Bayes, this assertion provoked a natural, one might say Holmesian question: How much evidence would it take to convince us that something we consider improbable has actually happened? When does a hypothesis cross the line from impossibility to improbability, and even to probability or virtual certainty?

Bayes was interested in establishing a statistical framework that takes prior hypotheses of probability for a specific situation and updates them with evidence. Bayes developed the framework, to update the prior hypothesis of one's perceived probability in the resurrection of Christ given eyewitness evidence, that is one of the most powerful paradigms in modern day for evaluating probability. Bayesian inference offers end users freedom and structure to incorporate subjectivity.

After months of speaking with front office executives and analytics personnel in Major League Baseball, I can attest that Bayesian inference is the most widely used statistical framework that teams evaluate probabilities both at the player and team level. I think that is due in part to modern computation tools enabling users (like Baseball teams) to scale efficiently. The math for Bayesian Inference is quite involved, so I am not going to write about it. I will attach equations and code I used for this analysis at the end to show I did not bring this stuff out of thin air.

I hope to paint a picture of the interpretation of output. I am going to use 2 USF Baseball examples comprised mostly of words to explain the intuition, one at the player level and one at the team level.

Nick Yovetich K-Rate

Nick Yovetich struck out in 35% of plate appearances in 2018. In 2019, he struck out in 25% of plate appearances. It does not take a Baseball genius to think that Nick can make an exponentially greater impact if his strikeout rate stays down (given power and run tools remain the same). On the same note, coaches face decisions regarding keeping Nick in the lineup if strikeouts increase. On one hand, we do not want to give up upside and know that taking him out of the lineup could disrupt his rhythm as a player. On the other hand, we need to win games (goal of ~65% wins) and viable replacement players could provide more production. Bayesian inference could offer more clarity on this situation.

Suppose we were 60% confident that Nick's K-Rate would hover around 25%. After 50 plate appearances, Nick struck out 15 times and was teetering in dangerous territory of falling above 30% K-Rate. An obvious confounding variable regarding the decision to keep Nick in the lineup would be what happened in the other 35 plate appearances, but Bayes can help us out with adjusting our prior beliefs to a posterior understanding for the strikeouts.

Intuition Check

Prior: We are 60% confident Nick Yovetich is a 25% K-Rate hitter

Evidence: Nick Yovetich strikes out in 15 of first 50 Plate Appearances (30%)

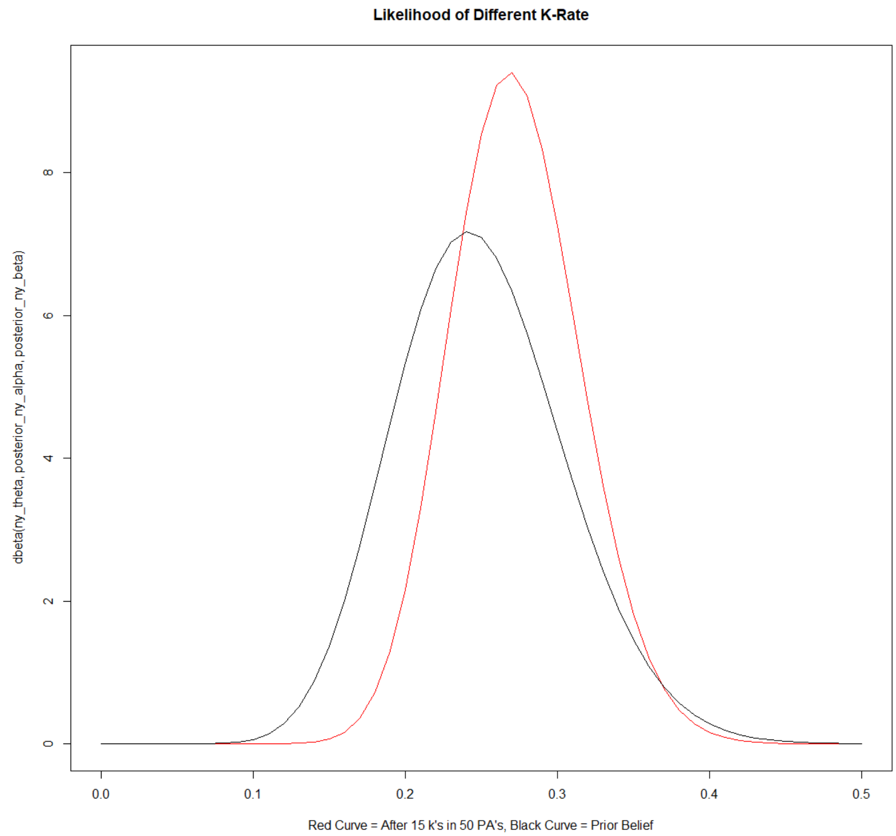
Posterior: ???

Do we completely abandon our 60% confidence that Nick Yovetich is a true 25% K-Rate hitter and say he is now a 30% K-Rate hitter? We do not. Bayesian Inference assesses the probability that if Nick were a true 25%, how likely is it that he would strike out in 15 of 50 plate appearances. Our subjective 60% prior confidence is also nested right into the calculation.

Bayes tells us not to freak out. Bayesian Inference tells us that Nick's new most realistic K-Rate, after establishing a prior and observing evidence, should be adjusted to 27.7%. We can do even better though. We can assess the likelihood that Nick hits different K-Rates by the end of the year after observing the 30% in the first 50 plate appearances with 60% confidence that he is a 25% K-Rate hitter.

True K-Rate	Likelihood after 15/50 K's
Less than or equal to .35	.96
Less than or equal to .30	.75
Less than or equal to .29	.67
Less than or equal to .28	.58
Less than or equal to .27	.48
Less than or equal to .26	.39
Less than or equal to .25	.30
Less than or equal to .20	.03

The prior updates to a new posterior with the black curve representing our prior belief and red curve representing our posterior belief. Our posterior's most likely value is the new 27.7%.



<u>Getting to 36 Wins (65%)</u>

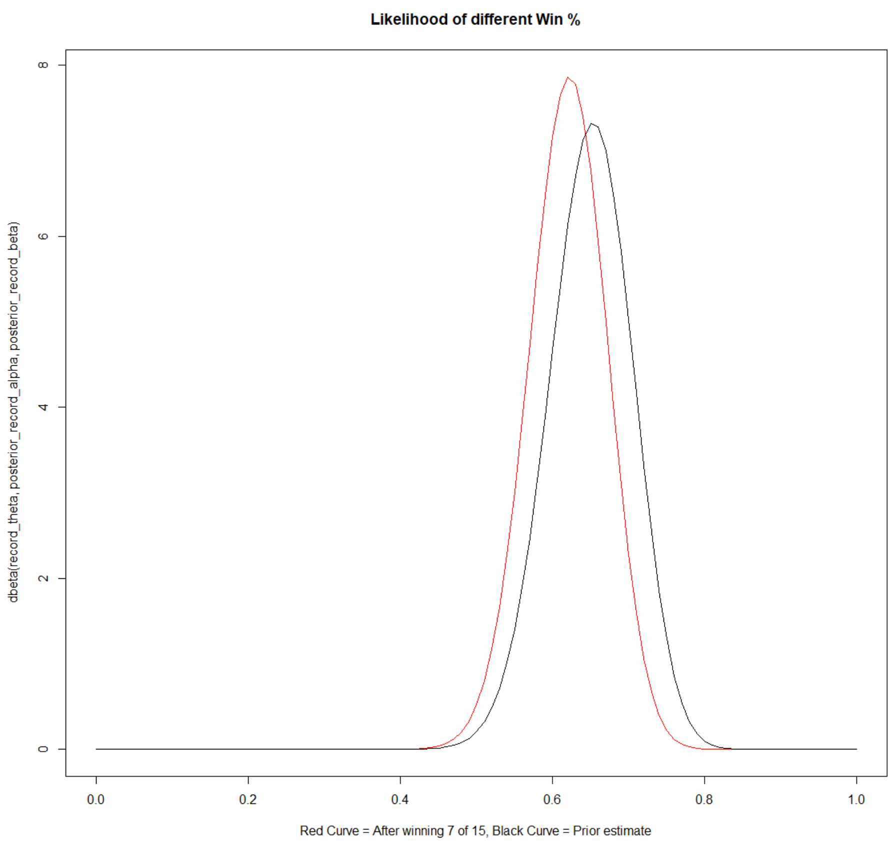
During the leadoff dinner, Coach stated winning 36 games is a goal for this season. The goal corresponds to winning 65% of games. We could use Bayesian inference to monitor the likelihood of achieving the goal in real time during the season. I do understand one fault in doing this with Bayesian Inference is that not all opponents are created equal, and that is not modeled in (as of now).

For example, say the club gets off to a slow start in the first 15 games before league play begins and goes 7-8. If we had 60% confidence before the season that we could win 65% of our games, how likely are we to achieve our goal if we get off to a slow start? According to Bayes, we better not get off to a slow start. If we got off to a 7-8 start, given our prior 60% confidence in winning 65% of games, our new most realistic win rate would drop to 61%.

The schedule of course could dictate the ebbs and flows of the season. But another beauty of Bayesian inference is that posterior beliefs today are the prior beliefs of tomorrow and we are always updating beliefs, never staying married to a belief. Here are the likelihood values:

True Win Rate	Likelihood
Greater than or equal to .40	.99
Greater than or equal to .45	.99
Greater than or equal to .50	.98
Greater than or equal to .55	.91
Greater than or equal to .60	.65
Greater than or equal to .65	.27

Per Bayesian Inference, would the 65% wins goal be in jeopardy if we got off to a slow start? Yes, but the likelihood that the season snowballing at the point of a 7-8 record is not feasible and we should not panic. Below is the visual that displays the new 61% as the most likely outcome in red.



Conclusion

What makes Bayesian Inference unique is that *subjectivity acts as an input to the objective framework*. The shape of the curves above are influenced by our priors and those are subjective. How much the data/evidence moves our priors is influenced by how subjectively confident we are in our priors. Running up to, and throughout the season, I am going to sharpen the application of Bayesian Inference that I think is scalable. Getting your head around the math involved is only half the battle, if that. The bigger part is moving the ideas through a computer efficiently. I am interested in both monitoring outcomes at the player performance and team performance levels. I will be in touch with everyone regarding players and team level outcomes I am interested in monitoring. If you have anything you would like tracked with the process outlined above, please do not hesitate to call me and we will talk.

Bayes Theorem Interpretations

Bayes theorem serves as the foundation of all Bayesian Inference application, which extends to very complex methods. Some (not all) interpretations of Bayes theorem are as follows:

$$P(Hypothesis|Data) = \frac{P(Hypothesis) * P(Data|Hypothesis)}{P(Data)}$$

OR

$$P(\theta|D) = \frac{P(D|\theta) * P(\theta)}{\sum P(D|\theta) * P(\theta)}$$

OR

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(B|A) * P(A)}{\left(P(B|A) * P(A)\right) + (P(B|A^C) * P(A^C))}$$

Graphical Approach with Beta Binomial Distribution Conjugate Prior

Prior Beta Binomial Distribution ~ $Beta(\alpha, \beta)$

Mean of Beta Prior = $\frac{\alpha}{(\alpha+\beta)}$

Standard Deviation of Beta Prior = $\frac{\alpha*\beta}{(\alpha+ \beta)^2*(\alpha+ \beta+1)}$

Mean of Beta Posterior = $\frac{\alpha+ \sum Y_i}{\alpha+ \beta+N}$