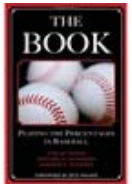
TANGO LICHTMAN DOLPHIN

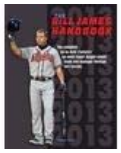
THE BOOK
Playing the Percentages in Baseball

THE BOOK ARTICLES BLOG ABOUT US



The Unwritten Book is Finally Written!
An in-depth analysis of: The sacrifice bunt, batter/pitcher matchups, the intentional base on balls, optimizing a batting lineup, hot and cold streaks, clutch performance, platooning strategies, and much more.
Read Excerpts & Customer Reviews

Buy The Book from Amazon



2013 Bill James Handbook

SEARCH
Advanced

InsideTheBook
Google Search

THE BOOK--Playing The Percentages In Baseball

2011/04/01 (Fri) @ 13:01:00

<< Back to main

Friday, April 01, 2011

How to figure out how much talent there is: example with NHL goalies

By Tangotiger, 01:01 PM

Someone supplied me with even strength shots faced and saves made for, well, I didn't ask. It has 55 goalies, with an average of 2665 even strength shots faced, and Kipper leading at 4784. That's probably about four seasons worth. Not important.

Step 1: Figure out how much one standard deviation is based on a

Latest...

COMMENTS

Feb 10 22:49
You say Goodbye... and I say Hello

Jan 25 14:36
Blog Beta Testers Needed

Jan 18 22:41

binomial distribution. Vokoun faced 4249 shots, and the league average save percentage was .920, so one SD is sqrt (.92*.08/4249) = .0042.

Step 2: Figure out how much away you are from the mean. Vokoun's save percentage was .931, and so was +.0108 from the mean.

Step 3: Figure out how many SD that is. .0108/.0042 = 2.59. That's his z-score.

Step 4: Do it for all the goalies. (Thomas is 2.57, Luongo is 2.53… Holmqvist is -3.11, Raycroft is -2.34).

Step 5: Find the standard deviation of all the z-scores. In this case, for these 55 goalies, it's 1.38.

Step 6: Rejoice if the number is substantially higher than 1.00. Happiness sets in at 1.10. You did good at 1.20. If you get 1.40, you've definitely found something.

Step 7: Figure out the average number of opportunities for each player. In this case, the average shots faced was 2665.

Step 8: Do this: 1 - 1/1.38^2 = 0.47. That's your r or r-squared. (Longer story later. Just call it r for now.) That 1.38 was from Step 6.

Step 9: Do this: (1-r)/r * 2665. We get 2969. The 2665 is from Step 7.

That's the key number. 2969. Let's call it 3000. That's how much you use to regress a goalie's performance. You add 3000 shots of league average performance. So Vokoun's 4249 shots at .931 save percentage gets added to 3000 shots at .920 save percentage for a best-estimate true talent level of .926. Holmqvist's .900 with 1809 shots becomes .912. So, the observed difference between the two goalies (.031 saves per shot) becomes a true difference of .014.

A couple of important points:
1. This tells us how many talent there is FOR THE SAMPLE of goalies. Put in less goalies, the talent level will be tighter, and the regression will be higher. Put in more goalies, the talent level will be wider, and the regression will be smaller.

2. It's best not to mix years without adjusting. As long as the seasons you are mixing up has a fairly static talent level and league averages, then don't worry.

NHL apologizes for being late, and will have players make it up for them

Jan 17 11:31
NHL, NHLPA MOU

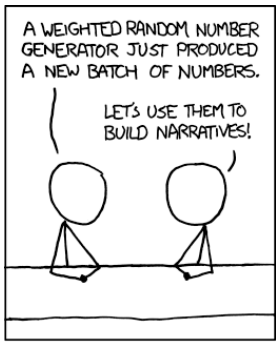
Jan 15 15:40
Looks like I picked a good day to suspend blogging

Jan 05 13:24
Are the best one-and-done players better than the worst first-ballot Hall of Famers?

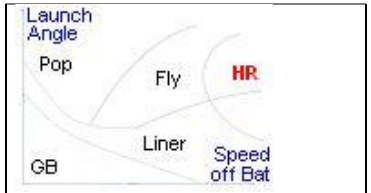
Jan 05 12:52
Poll: I read eBooks on…

Jan 05 12:06
Base scores

Jan 04 15:45
"The NHL is using this suit in an attempt to force the players to remain in a union"



• Sabermetrics • Talent Distribution



#1 Tangotiger (see all posts)

That “3000” is similar to BABIP for pitchers. Interestingly, a goalie and starting pitcher will each face about 20 even strength shots or balls in play per start.

The difference is that a goalie will have about twice as many starts as a pitcher, and so, we are aware of his save skills twice as fast as we are aware of a pitcher’s BIP skill.

#2 Sunny Mehta (see all posts)

I'm sure this method can be a useful model, perhaps for skills with a lower regression point and larger sample of players (e.g., maybe walk rate for batters). Imo it has a few question marks here.

A "K" (or "X") of 3000 is quite high relative to a lot of other sport skills we model around here, and your sample of 55 players is a good bit smaller.

1) Is it fair to simple exclude the rest of the population? Or will you have a better model by including them?

2) You are calculating the binomial variance by assuming the goalies all faced the same number of shots. But they didn't. What would the luck distribution look like if you didn't make that assumption?

3) You are assuming the talent distribution is normally distributed. Is that fair?

4) While your method finds "something" going on beyond binomial variance, is it fair to call it "goaltender skill"? Scorer bias is a real problem wrt shots in the NHL, and there's also the possibility of team defense or opposition offense having some influence. How much of what you've found is puck-stopping skill versus scorer bias/team effects/quality of competition/misc? Particularly with a K value that high, it's not like we're staring in the face of a blatantly obvious/evident skill.

5) What happens if we use this model to try and predict a future season for these goalies? How does the model fare? How does it fare for goalies who switch teams? How does it fare using road data only?

#3 Tangotiger (see all posts)

1: The conclusion I made is based only on the population at hand. If I had the data for 1000 goalies in 2010, then I would find a definite skill and the regression point would be much smaller.

The smaller I make my population, the less spread in talent I will find. This is true of anything. If I only looked at the top 55 pitchers or hitters in MLB in 2010, and I looked at the spread of their results in 2011, it will be much tighter than if I looked at the top 155, 255, or 355 players from the previous season.

2: No, I'm not doing that.

3: I think that is a pretty fair thing to presume because I don't think you can ever get THE actual 55 best goalies. What we are (implicitly) selecting is 55 semi-random goalies from a group of the 100 best goalies.

And, I don't think it matters much anyway.

4: Excellent point. Team defense and scorer bias could definitely play a role here. While five years of data would mitigate the issue of team defense (it's not like teammates are static), the scorer bias would be more real.

If I had the data by rink, I would have used that. If someone wants to supply me that, then fine.

5: It'll fare as good as the underlying model's data assumptions.

Since the model did not account for team-switchers, you can't extrapolate the model to look only at the team-switchers.

#4 Sunny Mehta (see all posts)

Thanks for the reply, Tangotiger. One of the things I love about this blog is that you make all of your methodologies "open source" and open to debate.

re: "2: No, I'm not doing that."

Aren't you finding the variance of the luck distribution by doing (.920*(1-.920)/2665)? I.e. you are figuring out the average number of shots faced and the average save percentage for these goalies and calculating the binomial variance? That's what you did in the original thread on this subject, as well as the similar thread on batted ball FIP.

And I'm saying you'll get a slightly different number if you give them all the average save percentage but give them their ACTUAL shots faced.

#5 Tangotiger (see all posts)

Sunny, I said this in Step 1:

> Vokoun faced 4249 shots, and the league average save percentage was .920, so one SD is sqrt (.92*.08/4249) = .0042.

So, I am calculating the z-score properly:, for each goalie, I use his actual shots faced.

The 2665 only comes into play to decide how relevant the SD of these z-scores (1.38) means.

Otherwise, how would you propose I proceed exactly (please provide instructions)?

#6 rempart (see all posts)

does this method only when work a percentage is involved like the 92% above?

#7 Tangotiger (see all posts)

It has to be p / (p+q), where p=success and q=failure.

#8 rempart (see all posts)

thought so. thanks alot.

#9 Sunny Mehta (see all posts)

"Otherwise, how would you propose I proceed exactly (please provide instructions)? “

Hi Tango,

Sorry about the long delay in response. Basically, the easiest way is to run a sim.

You want to know what the distribution would look like if there were no skill whatsoever, so you give every goalie the same “save percentage coin” and have them flip it the number of times equal to their actual shots faced, and then you measure the spread of that distribution (you don’t have to use

only sd here, you can use mad).

Rinse and repeat several thousand times and you have a good idea of what the luck distribution looks like. Then you can compare it to the observed distribution and take a guess at what the skill distribution is (you don't have to just subtract variances).

I don't know what language you code in, but for example, in R the "rbinom" function outputs random numbers from the binomial distribution (you choose the N, n, and p).

So if N is the number of goalies, n is a vector of the goalies' shots faced, and p is the static save percentage you are hypothesizing. The code in R would be something like:

```
results=replicate(10000, var(rbinom(N,n,p)/n))
summary(results)

#[you can replace var with whatever measure of spread you want.]
```

#10 Tangotiger (see all posts)

But why would I need to use a sim for this:

> You want to know what the distribution would look like if there were no skill whatsoever, so you give every goalie the same "save percentage coin" and have them flip it the number of times equal to their actual shots faced, and then you measure the spread of that distribution (you don't have to use only sd here, you can use mad).

This is exactly what I already did with the empirical data. I calculated each goalie's z-score doing exactly that. I took .920 as the skill-free save percentage (skill-free, meaning skill-free among the 50 goalies in the sample), I calculated the standard deviation using each goalies actual shots faced, and then calculated how many standard deviations his performance was from the mean.

I calculated all those z-scores, and then took a standard deviation of those z-scores. If there was no skill at all, the SD of the Z would come in at 1.000. It didn't. Nowhere close to 1.000.

There's something there. You can argue scorer bias, team bias, or that it's actually the goalie, or a combination of the three. Without a doubt, we would find that the biggest impact will be from the goalie. You simply will

not find scorer or team bias have the effect we are seeing.

So, I don't understand the objection to what I did, other than the potential biases.

#11 Sunny Mehta (see all posts)

tango,

that's not the part we're disagreeing about. the initial z-score part is fine because there you are essentially just figuring out "what percent bigger or smaller is the observed distribution than the luck distribution." (in this case the sd of observed was 1.38 times wider than the luck.)

but when it comes time to put an actual number on the sd of the luck distribution (i.e., ok we know the observed is 1.38 times wider than the luck, but WHAT IS the sd the luck?), you are doing it by giving every goalie the same number of shots faced. and i'm saying that's not quite accurate.

does that make sense?

#12 Tangotiger (see all posts)

Right, I am doing that, but in a practical sense, it won't matter, since all the goalies where between 1500 and 4500 shots, with an average of 2660 (or whatever the actual numbers were).

Whether I take the average, or 1/avg(1/shots), or more generally as (avg(shots^x))^(1/x), where x is plus or minus 0.5, 1, or 2, I'm going to end up to around the same thing as the straight mean.

Had I looked at all goalies with at least 1 shot faced, then you'd have a much better point. In this particular case, the point isn't a concern.

#13 Tangotiger (see all posts)

This is the various mean I get, using various values of "x" in this equation: (avg(shots^x))^(1/x)

2868 2

2665 1
2558 0.5
2352 -0.5
2257 -1
2092 -2

That sets one standard deviation at between .0051 to .0059, depending on which of the above schemes is better to use.

#14 rempart (see all posts)

While watching the Cornell-Princeton lacrosse game Saturday, it became fairly obvious that the Princeton goalie was having the better game. It also appearsas if lacrosse goalies are at a much greater disadavtage compared to their hockey counterparts. I got to wondering, after having seen this article, what the true talent level of an NCAA lacrosse goalie. And, it also appeared that a lot of luck is involved in saving the ball.

I found the stats on the ncaa lacrosse site and performed what I thought Tango had done for NHL goalies.

The average save percentage of the 55 goalies is .532, they faced an average of 225 shots.

1. std dev of .033

2. std dev of z-score 1.28, lower than 1.38 for NHL goalies

3. an r=.39, I added 350 shots at .532 sv%

4.Getting back to the original Princeton v Cornell

Pri Tyler Fioritto has faced 210 Shots with .605 Sv%, his true talent level drops to .559.

Cor AJ Fiore has seen 172 shots and has a .506 sv%. his true talent is .523.

5. The observed vs the the true talent diff drops from (.605-.506) or.099 to .036

#15 Tangotiger (see all posts)

Going back to the NHL one:

the random variation we'd expect was one SD = .0053.

That is: sqrt(.92*.08/2665).

With the observed spread being 1.38 times wider, the observed SD was .0073.

That gives us an implied true of .005 (square root of .0073^2 - .0053^2).

***

In the lacrosse case, at 225 shots faced with a .532 save percentage, one SD is .033. With an observed 1.28, that gives us an observed SD of .043. Or a true spread of one SD = .027.

So, the observed spread in NCAA goalies is pretty wide (relative to NHL goalies). But, that's hardly surprising.

The BABIP skill is spread far wider in high school and college than it is in MLB.

This is true of pretty much all skills.