Luck-Adjusted Pitching Statistics

by Will Larson

March 9, 2009

Attempting to forecast players' performances causes headaches, despite the many theories around on what contributes to ERA, WHIP and Wins. Linear prediction systems such as FIP Era are on the right track, but there is a deeper level of causal effects at work. Any forecasting system should be able to predict player performance based on as much luck-independent information as possible.

FIP Era is calculated as (HR*13+(BB+HBP-IBB)*3-K*2)/IP +3.2. This is a rough parametrization of the effects of BB, HBP, IBB, K and IP on ERA, but is not estimated in any statistical sense. Also, FIP Era ignores useful information on the type of contact pitchers tend to induce beyond HRs, reflected in BABIP, and the strand rate of runners the pitcher allows. Like FIP Era, I try to estimate ERA using only the statistics that the pitcher can presumably control and that are readily accessible, but I also add more information by using additional statistics. To estimate ERA, luck-adjusted statistics first need to be computed; in particular BABIP, LOB%, and HR/9.

There are three statistics that are related to luck more than anything, BABIP, LOB%, and the number of HRs a picther gives up over the course of a season. These statistics are considered here to be the only causes of randomness for pitchers. Once these are estimated, luck-adjusted WHIP, ERA, and Wins is estimated using similar methods. At each step of the way, interesting pieces of information are uncovered that reveal a clearer picture of what determines a pitchers success.

Below are the summary statistics for the dataset of pitchers being analyzed.

Variable	Obs	Mean	Std. Dev.	Min	Max
babip	126	.3034921	.0235735	.24	.37
lob	126	72.02778	4.664412	61.1	84.4
gb	126	43.74603	6.776355	22	64
fb	126	35.92619	6.214707	20.4	53.4
iffb	126	9.706349	2.97205	4	19
k9	126	6.568492	1.585725	3.35	11.01
bb9	126	3.027936	.9407021	1.35	5.98
hr9	126	1.01746	.3001784	.44	1.94
teamruns	126	4.590079	.3899877	3.86	5.46
era	126	4.239365	.9498427	2.07	6.66
whip	126	1.357857	.1688057	1	1.8
w32	126	12.02906	3.574368	4.363636	22.70968

The sample includes 126 pitchers in 2008 with more than 100 innings and no more than 5 relief appearances. Average BABIP is 0.303, average LOB% is 72%, 43% of all hit balls are ground balls, and so on. W32 is the number of wins normalized over 32 starts. Teamruns is the average number of runs per game the pitcher's team scores over the course of the regular season.

In every estimation, general specifications are attempted first, using all potentially relevant, luck-independent variables. The estimates reported are the final, parsimonious regressions and these can be considered the best estimates of a change in a right-hand-side variable on the left-hand-side variable.

First, I estimate a model for a pitcher's BABIP.

BABIP=A1+A2*GB%+A3*FB%+A4*IFFB%+EA

Source	SS	df	MS
Model	.017682096	3	.005894032
Residual	.0517814	122	.000424438
Total	.069463495	125	.000555708

Number of obs	=	126
F(3, 122)	=	13.89
Prob > F	=	0.0000
R-squared	=	**0.2546**
Adj R-squared	=	0.2362
Root MSE	=	.0206

babip	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
gb	**-.0043473**	.0008444	-5.15	0.000	-.0060188	-.0026757
fb	**-.0053375**	.0009242	-5.78	0.000	-.0071671	-.0035079
iffb	**-.0011018**	.0006985	-1.58	0.117	-.0024846	.0002809
_cons	**.6961184**	.0691862	10.06	0.000	.5591573	.8330794

Where EA is a stochastic error term. These results show that GB%, FB%, and IFFB% explain about 25% of the variation in BABIP from pitcher to pitcher and indicate that BABIP is fairly random. An increase of 1% in GB% decreases the estimated BABIP of a pitcher by -0.0043 (other estimates are interpreted similarly). Therefore, a difference of 5% in GB% will result in a change in BABIP for a league average pitcher from .303 to about .323. Using these parameter estimates, I estimate a luck-adjusted BABIP as

aBABIP=A1+A2*GB%+A3*FB%+A4*IFFB%
aBABIP=0.696-0.00110*GB%-0.005*FB%-0.00434*IFFB%

Next, I estimate a luck-adjusted LOB%.

LOB%=B1+B2*K9+B3*BABIP+EB

Source	SS	df	MS
Model	1275.67491	2	637.837453
Residual	1443.91758	123	11.7391673
Total	2719.59249	125	21.7567399

Number of obs	=	126
F(2, 123)	=	54.33
Prob > F	=	0.0000
R-squared	=	0.4691
Adj R-squared	=	0.4604
Root MSE	=	3.4262

lob	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
k9	1.419747	.1933051	7.34	0.000	1.037111	1.802382
babip	-98.2937	13.00313	-7.56	0.000	-124.0326	-72.55481
_cons	92.53354	4.129896	22.41	0.000	84.35867	100.7084

Perhaps surprisingly, 46% of the pitcher to pitcher variation in LOB% can be explained just by the pitchers K/9 rate and its BABIP. An additional K/9 reduces the LOB% of a pitcher by 1.4%. Also, and more importantly, a .001 increase in BABIP reduces LOB% by 0.98%. This means that bad luck not only gets runners on base, it helps them to score. Therefore, rather than estimate aLOB% with BABIP to calculate luck-adjusted LOB%, I use the aBABIP previously estimated to control for a lucky BABIP.

aLOB%=B1+B2*K9+B3*aBABIP

I can also estimate the number of home runs per 9 innings as a function of the fly ball % and the number of strike outs per 9 innings.

HR9=C1+C2*FB%+C3*K9+EC

Source	SS	df	MS
Model	3.70082823	2	1.85041411
Residual	7.56255914	123	.061484221
Total	11.2633874	125	.090107099

Number of obs	126
F(2, 123)	30.10
Prob > F	0.0000
R-squared	0.3286
Adj R-squared	0.3177
Root MSE	.24796

hr9	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
k9	-.075395	.0146186	-5.16	0.000	-.1043317	-.0464584
fb	.0262808	.00373	7.05	0.000	.0188974	.0336641
_cons	.5685244	.1420498	4.00	0.000	.2873455	.8497034

A high strikeout rate and low FB% both reduce the frequency of surrendering home runs. Clearly, park factors also contribute to the frequency of HRs allowed, but these are omitted in this analysis. Luck-adjusted HR9 can be estimated as:

aHR9=C1+C2*FB%+C3*K9+EC

and the estimated number of home runs surrendered over the course of a season as

aHR=aHR9*IP/9

From aHR and aBABIP, estimated hits can also be calculated by

aH=aBABIP/BABIP*(H-HR)+aHR

and an estimated WHIP can be calculated as

aWHIP=(BB+aH)/IP

Next, ERA can be estimated by

ERA=D1+D2*LOB%+D3*K9+D4*BB9+D5*BABIP+D6*HR9+ED

Source	SS	df	MS
Model	108.36269	5	21.6725381
Residual	4.41246267	120	.036770522

Number of obs	126
F(5, 120)	589.40
Prob > F	0.0000
R-squared	0.9609

```
-------------+------------------------------           Adj R-squared =  0.9592
       Total |  112.775153   125  .902201224           Root MSE      =  .19176

------------------------------------------------------------------------------
         era |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         lob |  -.0944805   .0051727   -18.27   0.000    -.1047222   -.0842389
          k9 |  -.1243317   .0136964    -9.08   0.000    -.1514497   -.0972137
         bb9 |   .3139747   .0191164    16.42   0.000     .2761257    .3518238
       babip |   11.02971   .8819844    12.51   0.000     9.283445    12.77598
         hr9 |   1.324773   .0591788    22.39   0.000     1.207603    1.441943
       _cons |   6.215228    .544756    11.41   0.000     5.136649    7.293806
------------------------------------------------------------------------------
```

This equation predicts an astonishing 96% of the pitcher-to-pitcher variation in ERA. Put another way, values for LOB%, K9, BB9, BABIP, and HR9, ERA can be predicted with 96% accuracy. The randomness with ERA therefore lies within the components of these 5 variables. Again, looking at the "luck" variables, an increase in LOB% reduces ERA by 0.09 and an increase in babip of .001 increases ERA by .01. Luck-adjusted ERA is predicted using the same method as aBABIP, aLOB, and aHR9

aERA= D1+D2*aLOB%+D3*K9+D4*BB9+D5*aBABIP+D6*aHR9

Wins per 32 games started is estimated using the model

W32=E1+E2*ERA+E3*RF+EE

```
      Source |       SS       df       MS              Number of obs =     126
-------------+------------------------------           F(  2,   123) =   54.52
       Model |  750.475191     2  375.237596           Prob > F      =  0.0000
    Residual |  846.537934   123  6.88242222           R-squared     =  0.4699
-------------+------------------------------           Adj R-squared =  0.4613
       Total |  1597.01312   125   12.776105           Root MSE      =  2.6234

------------------------------------------------------------------------------
         w32 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         era |  -2.474227   .2475097   -10.00   0.000    -2.964157   -1.984297
    teamruns |   2.188119   .6028275     3.63   0.000     .9948587    3.381379
       _cons |   12.47458   2.907535     4.29   0.000     6.719287    18.22986
------------------------------------------------------------------------------
```

where RF, also called "teamruns" is the average runs the pitchers' team scores per game in all games it plays during the regular season for that year. A 1 point increase in ERA reduces the predicted number of games won by 2.47 over 32 starts. An increase of 1 average run per game for the pitchers team increases the number of wins by 2.1 over 32 starts.

and predicted as

aW32=E1+E2*aERA+E3*RF

and normalized luck-adjusted Wins are calculated, aW=aW32*Starts

Now that all of these new fancy luck-adjusted statistics have been created, who were the lucky and unlucky guys last year? How good is FIP Era?

FIP Era and aERA are naturally similar when the underlying luck-dependent statistics are close to the luck-independent statistics. However, in cases where a pitcher has luck-independent statistics that are farther away from the average for all pitchers, aERA and FIP Era diverge. By definition, since aERA is estimated using a minimum variance estimator, it is better at estimating within-season performance. In terms of forecasting, these two methods can be evaluated comparing actual to predicted statistics for years out-of-sample. This has not yet been done, but is necessary to establish which system has the better forecasting performance.