

Index

68-95-99.7 rule, 57

ability

career, 49

model, 187

peak, 49

ability and performance, 156

absorbing state, 213

adjusted ERA statistic, 53

All-Star Baseball, 5, 116

Altuve, Jose, 103

association

negative, 76

positive, 75, 82

attendance, ballpark, 23, 42

balk, 216

Baseball Archive, 11

baseball attendance, 102

baseball card, 9

Baseball Reference, 10, 20

batted ball data, 71

batter's runs average, 81

batting average, 7, 59

true, 164

Bayesian thinking, 160

bell-shaped, 182

Belt, Brandon, 230, 235

Bernoulli trials, 143

bias model, 185

Big League Baseball, 114

binomial distribution, 139

Bogg, Wade, 165

Bonds, Barry, 54, 149

boxplot, 48

parallel, 56

boxscore, 235

Brantley, Michael, 173, 188

Brenly, Bob, 224

Brett, George, 59

Bumgarner, Madison, 126

Cabrera, Melky, 140

Cabrera, Miguel, 80

career trajectory, 49

Carew, Rod, 59

Carter, Joe, 50

Cobb, Ty, 132

consistent hitter, 191

correlation, 79

Counsell, Craig, 224

Crawford, Brandon, 235

Cruz, Nelson, 103

data distribution, 15

average, 15

shape, 15

spread, 15

Davis, Chris, 112, 207

Davis, Wade, 237

derived baseball statistic, 13

designated hitter, 16, 27

dice

roll of two, 115

die

roll of 10 sided, 191

roll of 20 sided, 120

roll of one, 114

to simulate hitting data, 191

DiMaggio, Joe, 28

Donaldson, Josh, 172

dotplot, 18, 26

Cleveland-style, 25

parallel, 27, 56

earned run average, 8

- empirical rule, 57
- equally likely outcomes, 115
- ERA
 - adjusted, 53
 - statistic, 8, 51
- expected number
 - of hits, 141
 - of visits to a state, 214, 219
 - runs in an inning, 221
- Fangraphs, 11
- fastball speeds, 70
- Feller, Bob, 32
- five-number summary, 17, 46, 56
- Ford, Whitey, 51
- fundamental matrix, 221
- geometric distribution, 151
- Glavine, Tom, 66
- Goldschmidt, Paul, 173
- goodness of fit, 82
- Gould, Stephen Jay, 64
- Greinke, Zack, 43, 70
- Gwynn, Tony, 59, 64, 134, 165
- Hall of Fame pitchers, 38
- Henderson, Rickey, 1, 28, 61, 94, 124, 148, 168, 193, 225
- histogram, 24
- home run rate, 55, 76
- Howard, Ryan, 187
- James, Bill, 1, 8
- Jeter, Derek, 17
- Johnson, Randy, 20, 134
- Johnson, Walter, 101
- Keeler, Willie, 64
- Kershaw, Clayton, 70
- Koufax, Sandy, 31
- lapta, 233
- least-squares fit, 82, 87, 90, 92
- length of baseball game, 36
- leverage, 177
- Maddux, Greg, 66, 120
- Mantle, Mickey, 132, 157
- many spinners model, 183
- Maris, Roger, 9, 51, 54, 157
- Markov Chain, 211
- Martinez, Victor, 173
- McGwire, Mark, 54, 120
- mean, 57
- mean square error, 83
- median, 17, 46
- model
 - ability, 187
 - bias, 185
 - many spinners, 183
 - no effect, 185
 - one spinner, 182
- Morse, Michael, 235
- moving average, 188
- multiple linear regression, 86
- multiplication rule, 123
- negative association, 94
- negative binomial distribution, 143
- no effect model, 185
- nonlinear equation, 89
- normal curve, 183
- official at-bat, 236
- on-base percentage, 8, 46
 - true, 168
- on-base profile, 139
- one spinner model, 182
- OPS statistic, 9, 46, 81
- outcomes
 - of one die roll, 114
 - of two die rolls, 115
- outliers, 48
- Pagan, Angel, 125
- peak ability, 49
- Pearson residual, 144
- Pedroia, Dustin, 169
- Pence, Hunter, 230, 235
- Perez, Juan, 235
- performance, 156
- Pesky, Johnny, 8
- pitch count, 177
- pitches, number of, 101
- PITCHf/x system, 11
- plate appearance, 236
- Poisson distribution, 149
- positive association, 75
- probability
 - of set of outcomes, 114
 - of sum of two dice, 121
 - relative frequency interpretation, 112
 - subjective interpretation, 162
- probability interval, 164
- Pujols, Albert, 44, 45, 66, 71, 107
- Pythagorean method, 88

- quartile, 17, 46
 - lower, 17
 - upper, 17
- Raines, Tim, 4, 61
- Ramirez, Manny, 45, 66
- regression effect, 91
- relative standing, 15
- residual, 82, 91
 - Pearson, 144
- Retrosheet, 11
- Reynolds, Mark, 125
- right skewed, 23
- Roberts, Robin, 50
- Rodriguez, Alex, 162
- root mean square error, 83
- rounders, 233
- run expectancy matrix, 221
- run potential matrix, 224
- run value, of batting event, 222
- runs created, 8, 81
- runs of hot and cold performance, 190
- Ruth, Babe, 30, 54, 116

- sabermetrics, 8
- SABR, 8
- sacrifice bunt, 26
- salaries, 40
- sample space, 114
- Sandoval, Pablo, 235
- Sandoval, Pedro, 237
- scatterplot, 4, 19, 74, 82
 - labeled, 75
- scatterplot matrix, 79
- Schmidt, Mike, 117
- simulation
 - of binomial distribution, 141
 - of data from consistent hitter, 192
 - of dice rolls, 158
 - of hitting data, 167
 - of situational hitting, 186
 - one spinner model, 183
 - run scoring, 152
 - to learn about hitting ability, 160
- situational
 - ability, 180
 - statistics, 176
- skewed right, 23
- slugging percentage, 8, 46, 57
- smoothing curve, 19
- Society of American Baseball Research, 8, 80
- sophomore slump, 91
- spinner
 - calculation of probability regions, 117
 - to model hitting, 116, 181
- spinner model, 157
- standard deviation, 57
- standard normal density, 164
- standardized score, 59
- states of Markov Chain, 212
- statistics, meaning of, 2
- stemplot, 14, 21, 22, 57, 59
 - back to back, 15, 23, 46, 51
- step
 - for determining outliers, 48
- stolen base rate, 99
- Strat-O-Matic Baseball, 119
- streaky
 - ability, 188
 - performance, 188
- strikeout rate, 21, 126
- Stuart, Dick, 8
- subjective probability, 162
- Suzuki, Ichiro, 71, 124, 188, 202, 206
- symmetric, 24, 57

- time of game, 101
- time series plot, 18, 22, 53
- total average, 81
- transition probability matrix, 212
- tree diagram, 122
- triple rate, 76
- Trout, Mike, 80, 158, 176
- two-way table, 171

- value
 - of batting event, 222
 - of home run, 223
 - of sacrifice bunt, 224

- weights of batting events, 86
- Whiz Kids, 50
- Williams, Ted, 59
- win/loss ratio, 88
- winning percentage, team, 98
- Womack, Tony, 224
- World Series
 - 1950, 50
 - 1980, 50
 - 1993, 50
 - 2001, 224
 - 2008, 50
 - 2014, 24, 234
- z-score, 60