

1. Crazy Ideas
2. Run Scoring
3. Head-to-Head
4. Prospects

Some Good (and Bad) Ideas in Baseball Analysis

Chris Long,
clong@padres.com

October 13, 2009

1. Crazy Ideas
2. Run Scoring
3. Head-to-Head
4. Prospects

Table of Contents

1. Crazy Ideas
2. Run Scoring
 - Quasigeometric Distribution
 - Binned Weibull
3. Head-to-Head
 - log5
 - Bradley-Terry
4. Prospects
 - Amateur
 - Dustin Pedroia
 - Kellen Kulbacki
 - Bryce Brentz
 - Robert Stock
 - Gordon Beckham
 - Buster Posey

1. Crazy Ideas
2. Run Scoring
3. Head-to-Head
4. Prospects

If I were a GM ...

- Little people - already done (Eddie Gaedel)

1. Crazy Ideas
2. Run Scoring
3. Head-to-Head
4. Prospects

If I were a GM ...

- Little people - already done (Eddie Gaedel)
- Scrunching

1. Crazy Ideas
2. Run Scoring
3. Head-to-Head
4. Prospects

If I were a GM ...

- Little people - already done (Eddie Gaedel)
- Scrunching
- Barry Bonds in 2008 and 2009

1. Crazy Ideas
2. Run Scoring
3. Head-to-Head
4. Prospects

If I were a GM ...

- Little people - already done (Eddie Gaedel)
- Scrunching
- Barry Bonds in 2008 and 2009
- Creating knuckleballers

1. Crazy Ideas
2. Run Scoring
3. Head-to-Head
4. Prospects

If I were a GM ...

- Little people - already done (Eddie Gaedel)
- Scrunching
- Barry Bonds in 2008 and 2009
- Creating knuckleballers
- Female knuckleballers

1. Crazy Ideas
2. Run Scoring
3. Head-to-Head
4. Prospects

If I were a GM ...

- Little people - already done (Eddie Gaedel)
- Scrunching
- Barry Bonds in 2008 and 2009
- Creating knuckleballers
- Female knuckleballers
- Starter by committee

1. Crazy Ideas
2. Run Scoring
3. Head-to-Head
4. Prospects

If I were a GM ...

- Little people - already done (Eddie Gaedel)
- Scrunching
- Barry Bonds in 2008 and 2009
- Creating knuckleballers
- Female knuckleballers
- Starter by committee
- Prospects in India, Guatamala, Ecuador, etc.

1. Crazy Ideas
2. Run Scoring
3. Head-to-Head
4. Prospects

Quasigeometric Distribution

- Geometric except for $n = 0$, so two parameters

1. Crazy Ideas
2. Run Scoring
3. Head-to-Head
4. Prospects

Quasigeometric Distribution

- Geometric except for $n = 0$, so two parameters
- Woolner, Tango, Glass & Lowry, Gregory

Quasigeometric Distribution

- Geometric except for $n = 0$, so two parameters
- Woolner, Tango, Glass & Lowry, Gregory
- Q: Estimate parameters from Runs/Game

Quasigeometric Distribution

- Geometric except for $n = 0$, so two parameters
- Woolner, Tango, Glass & Lowry, Gregory
- Q: Estimate parameters from Runs/Game
- Q: Relationship to Pythagorean Expectation,

$$WF(R, A) \approx \frac{R^\gamma}{R^\gamma + A^\gamma}$$

Quasigeometric Distribution

- Geometric except for $n = 0$, so two parameters
- Woolner, Tango, Glass & Lowry, Gregory
- Q: Estimate parameters from Runs/Game
- Q: Relationship to Pythagorean Expectation,

$$WF(R, A) \approx \frac{R^\gamma}{R^\gamma + A^\gamma}$$

- Q: Incorporate park factors
- Q: Relationship to binned Weibull

1. Crazy Ideas
2. Run Scoring
3. Head-to-Head
4. Prospects

Quasigeometric Distribution
Binned Weibull

Binned Weibull

- Steven Miller, Brown University (now Williams College)

1. Crazy Ideas
2. Run Scoring
3. Head-to-Head
4. Prospects

Binned Weibull

- Steven Miller, Brown University (now Williams College)
- Sample from a continuous distribution, binning afterwards

Binned Weibull

- Steven Miller, Brown University (now Williams College)
- Sample from a continuous distribution, binning afterwards
- Three-parameter Weibull,

$$f(x; \alpha, \beta, \gamma) = \frac{\gamma}{\alpha} \left(\frac{x - \beta}{\alpha} \right)^{\gamma-1} e^{-((x-\beta)/\alpha)^\gamma}$$

Binned Weibull

- Steven Miller, Brown University (now Williams College)
- Sample from a continuous distribution, binning afterwards
- Three-parameter Weibull,

$$f(x; \alpha, \beta, \gamma) = \frac{\gamma}{\alpha} \left(\frac{x - \beta}{\alpha} \right)^{\gamma-1} e^{-((x-\beta)/\alpha)^\gamma}$$

- Here $R = \alpha\Gamma(1 + \gamma^{-1}) + \beta$, β controls the binning translation

Binned Weibull

- Steven Miller, Brown University (now Williams College)
- Sample from a continuous distribution, binning afterwards
- Three-parameter Weibull,

$$f(x; \alpha, \beta, \gamma) = \frac{\gamma}{\alpha} \left(\frac{x - \beta}{\alpha} \right)^{\gamma-1} e^{-((x-\beta)/\alpha)^\gamma}$$

- Here $R = \alpha\Gamma(1 + \gamma^{-1}) + \beta$, β controls the binning translation
- Result:

$$WF(R, A, \beta, \gamma) = \frac{(R - \beta)^\gamma}{(R - \beta)^\gamma + (A - \beta)^\gamma}$$

Binned Weibull

- Steven Miller, Brown University (now Williams College)
- Sample from a continuous distribution, binning afterwards
- Three-parameter Weibull,

$$f(x; \alpha, \beta, \gamma) = \frac{\gamma}{\alpha} \left(\frac{x - \beta}{\alpha} \right)^{\gamma-1} e^{-((x-\beta)/\alpha)^\gamma}$$

- Here $R = \alpha\Gamma(1 + \gamma^{-1}) + \beta$, β controls the binning translation
- Result:

$$WF(R, A, \beta, \gamma) = \frac{(R - \beta)^\gamma}{(R - \beta)^\gamma + (A - \beta)^\gamma}$$

- Miller takes $\beta = -0.5$ and recovers the estimates $\gamma = 1.79$, commonly $\gamma = 1.81$

1. Crazy Ideas
2. Run Scoring
3. **Head-to-Head**
4. Prospects

log5
Bradley-Terry

log5

- Bill James (1981); now with the Boston Red Sox

log5

- Bill James (1981); now with the Boston Red Sox
- Let p_i, q_i be success/failure fractions for players $i = 1, 2$, p be the success fraction for 1 over 2, then

$$p \approx \frac{p_1 q_2}{p_1 q_2 + p_2 q_1}$$

log5

- Bill James (1981); now with the Boston Red Sox
- Let p_i, q_i be success/failure fractions for players $i = 1, 2$, p be the success fraction for 1 over 2, then

$$p \approx \frac{p_1 q_2}{p_1 q_2 + p_2 q_1}$$

- Some algebra yields

$$\frac{p}{1-p} \approx \frac{p_1}{1-p_1} \frac{1-p_2}{p_2}$$

1. Crazy Ideas
2. Run Scoring
3. **Head-to-Head**
4. Prospects

log5
Bradley-Terry

Bradley-Terry

- Bradley & Terry (1952); Luce (1959)

Bradley-Terry

- Bradley & Terry (1952); Luce (1959)
- Player ratings $A, B > 0$, and

$$p = \frac{A}{A + B}$$

Bradley-Terry

- Bradley & Terry (1952); Luce (1959)
- Player ratings $A, B > 0$, and

$$p = \frac{A}{A + B}$$

- Result: Let L be the league average rating, and assume $p_1 = A/(A + L)$, $p_2 = B/(B + L)$, then the BT model implies

$$\frac{p}{1 - p} \approx \frac{p_1}{1 - p_1} \frac{1 - p_2}{p_2}$$

Bradley-Terry

- Bradley & Terry (1952); Luce (1959)
- Player ratings $A, B > 0$, and

$$p = \frac{A}{A + B}$$

- Result: Let L be the league average rating, and assume $p_1 = A/(A + L)$, $p_2 = B/(B + L)$, then the BT model implies

$$\frac{p}{1 - p} \approx \frac{p_1}{1 - p_1} \frac{1 - p_2}{p_2}$$

- Recover log5; note, Elo rating system also uses BT model (with logistic link)

Bradley-Terry

- Bradley & Terry (1952); Luce (1959)
- Player ratings $A, B > 0$, and

$$p = \frac{A}{A + B}$$

- Result: Let L be the league average rating, and assume $p_1 = A/(A + L)$, $p_2 = B/(B + L)$, then the BT model implies

$$\frac{p}{1 - p} \approx \frac{p_1}{1 - p_1} \frac{1 - p_2}{p_2}$$

- Recover log5; note, Elo rating system also uses BT model (with logistic link)
- Q: Extend to multinomial outcomes

1. Crazy Ideas
2. Run Scoring
3. Head-to-Head
4. **Prospects**

Amateur

Amateur

- One of the most important areas in baseball, especially for teams with low payrolls (Padres)

1. Crazy Ideas
2. Run Scoring
3. Head-to-Head
4. **Prospects**

Amateur

- One of the most important areas in baseball, especially for teams with low payrolls (Padres)
- Age, position, defense, athleticism, signability, park factors, strength of schedule, . . .

1. Crazy Ideas
2. Run Scoring
3. Head-to-Head
4. **Prospects**

Amateur

- One of the most important areas in baseball, especially for teams with low payrolls (Padres)
- Age, position, defense, athleticism, signability, park factors, strength of schedule, . . .
- Scouts!

1. Crazy Ideas
2. Run Scoring
3. Head-to-Head
4. Prospects

Amateur

Dustin Pedroia

- Dustin Pedroia: 2nd round (2004) by Boston Red Sox

- JR year (Arizona State)

BA	OBP	SLG	OPS
393	502	611	1113

- 9 HR, 48 BB, 15 SO

1. Crazy Ideas
2. Run Scoring
3. Head-to-Head
4. Prospects

Amateur

Dustin Pedroia

- Dustin Pedroia: 2nd round (2004) by Boston Red Sox

- JR year (Arizona State)

BA	OBP	SLG	OPS
393	502	611	1113

- 9 HR, 48 BB, 15 SO
- Result: 2008 AL MVP!

1. Crazy Ideas
2. Run Scoring
3. Head-to-Head
4. **Prospects**

Amateur

Kellen Kulbacki

- Kellen Kulbacki: 1st round (2007) by San Diego Padres

- SO year (James Madison)

BA	OBP	SLG	OPS
464	568	943	1511

- 24 HR, 32 BB, 30 SO, 18 HBP

Kellen Kulbacki

- Kellen Kulbacki: 1st round (2007) by San Diego Padres
- SO year (James Madison)
- 24 HR, 32 BB, 30 SO, 18 HBP
- Result: Injuries, but possibly the best hitter in the minors from June-September 2008

BA	OBP	SLG	OPS
464	568	943	1511

Kellen Kulbacki

- Kellen Kulbacki: 1st round (2007) by San Diego Padres
 - SO year (James Madison)
- | BA | OBP | SLG | OPS |
|-----------|------------|------------|------------|
| 464 | 568 | 943 | 1511 |
- 24 HR, 32 BB, 30 SO, 18 HBP
 - Result: Injuries, but possibly the best hitter in the minors from June-September 2008
 - Comparison: position, size, defense, strength of schedule

1. Crazy Ideas
2. Run Scoring
3. Head-to-Head
4. **Prospects**

Amateur

Bryce Brentz

- Bryce Brentz: 1st round (2010) by ?

- SO year (Middle Tennessee)

BA	OBP	SLG	OPS
465	541	930	1471

- 28 HR, 31 BB, 32 SO (but only 7 HBP)

1. Crazy Ideas
2. Run Scoring
3. Head-to-Head
4. **Prospects**

Amateur

Bryce Brentz

- Bryce Brentz: 1st round (2010) by ?

- SO year (Middle Tennessee)

BA	OBP	SLG	OPS
465	541	930	1471

- 28 HR, 31 BB, 32 SO (but only 7 HBP)
- Strength of schedule . . .

1. Crazy Ideas
2. Run Scoring
3. Head-to-Head
4. **Prospects**

Amateur

Bryce Brentz

- Bryce Brentz: 1st round (2010) by ?

- SO year (Middle Tennessee)

BA	OBP	SLG	OPS
465	541	930	1471

- 28 HR, 31 BB, 32 SO (but only 7 HBP)
- Strength of schedule . . .
- Team USA

1. Crazy Ideas
2. Run Scoring
3. Head-to-Head
4. **Prospects**

Amateur

Robert Stock

- Robert Stock: 2nd round (2009) by St. Louis Cardinals

- JR year (USC)

BA	OBP	SLG	OPS
226	354	453	807

- 6 HR, 23 BB, 24 SO

1. Crazy Ideas
2. Run Scoring
3. Head-to-Head
4. **Prospects**

Amateur

Robert Stock

- Robert Stock: 2nd round (2009) by St. Louis Cardinals

- JR year (USC)

BA	OBP	SLG	OPS
226	354	453	807

- 6 HR, 23 BB, 24 SO
- Terrible?

Robert Stock

- Robert Stock: 2nd round (2009) by St. Louis Cardinals

- JR year (USC)

BA	OBP	SLG	OPS
226	354	453	807

- 6 HR, 23 BB, 24 SO
- Terrible?
- Catcher
- Also a pitcher (closer)
- Only 19!
- Showcase superstar in high school

Robert Stock

- Robert Stock: 2nd round (2009) by St. Louis Cardinals

- JR year (USC)

BA	OBP	SLG	OPS
226	354	453	807

- 6 HR, 23 BB, 24 SO
- Terrible?
- Catcher
- Also a pitcher (closer)
- Only 19!
- Showcase superstar in high school
- “Robert Stock is like a mystery novel, you have to put the clues together.”

Robert Stock

- Robert Stock: 2nd round (2009) by St. Louis Cardinals

- JR year (USC)

BA	OBP	SLG	OPS
226	354	453	807

- 6 HR, 23 BB, 24 SO
- Terrible?
- Catcher
- Also a pitcher (closer)
- Only 19!
- Showcase superstar in high school
- “Robert Stock is like a mystery novel, you have to put the clues together.”
- Result: Very good debut in rookie ball

1. Crazy Ideas
2. Run Scoring
3. Head-to-Head
4. Prospects

Amateur

Gordon Beckham

- Gordon Beckham: 1st round, 8th overall (2008) by Chicago White Sox

- JR year (Georgia)

BA	OBP	SLG	OPS
411	519	804	1323

- 28 HR, 54 BB, 30 SO

1. Crazy Ideas
2. Run Scoring
3. Head-to-Head
4. Prospects

Amateur

Gordon Beckham

- Gordon Beckham: 1st round, 8th overall (2008) by Chicago White Sox

- JR year (Georgia)

BA	OBP	SLG	OPS
411	519	804	1323

- 28 HR, 54 BB, 30 SO
- Shortstop

Gordon Beckham

- Gordon Beckham: 1st round, 8th overall (2008) by Chicago White Sox

- JR year (Georgia)

BA	OBP	SLG	OPS
411	519	804	1323

- 28 HR, 54 BB, 30 SO
- Shortstop
- Signed bat
- Result: Already a major-leaguer

1. Crazy Ideas
2. Run Scoring
3. Head-to-Head
4. Prospects

Amateur

Buster Posey

- Buster Posey: 1st round, 5th overall (2008) by SF Giants

- JR year (Florida State)

BA	OBP	SLG	OPS
463	566	879	1445

- 26 HR, 57 BB, 29 SO

Buster Posey

- Buster Posey: 1st round, 5th overall (2008) by SF Giants

- JR year (Florida State)

BA	OBP	SLG	OPS
463	566	879	1445

- 26 HR, 57 BB, 29 SO
- Catcher

Buster Posey

- Buster Posey: 1st round, 5th overall (2008) by SF Giants

- JR year (Florida State)

BA	OBP	SLG	OPS
463	566	879	1445

- 26 HR, 57 BB, 29 SO
- Catcher
- Jersey
- Result: Already a major-leaguer

Buster Posey

- Buster Posey: 1st round, 5th overall (2008) by SF Giants

- JR year (Florida State)

BA	OBP	SLG	OPS
463	566	879	1445

- 26 HR, 57 BB, 29 SO
- Catcher
- Jersey
- Result: Already a major-leaguer
- Tim Beckham, 1st round, 1st overall, \$6.1m
- Eric Hosmer, 1st round, 3rd overall, \$6.0m
- Kyle Skipworth, 1st round, 6th overall, \$2.3m