The Math of Moneyball
. . . with applications to airline safety

David Czerwinski

SJSU College of Business

Presented at the SJSU Math Colloquium
April 1, 2009

A brief proof of the equality P=NP

Squaring the Circle: We've finally done it

Fermat's Lesser Known "First Theorem"

Leibniz and Newton: Secret Lovers?

Moneyball

- Follows the 2002 season of the Oakland A's

Moneyball

- Follows the 2002 season of the Oakland A's
- Is the story of how the A's use mathematical modeling to gain an edge

Moneyball

- Follows the 2002 season of the Oakland A's
- Is the story of how the A's use mathematical modeling to gain an edge
- By the author of *Liar's Poker*

The objective

- What is the objective of a major-league baseball team?

The objective

- What is the objective of a major-league baseball team?
- For the 2002 Oakland A's, the objective is to make the playoffs

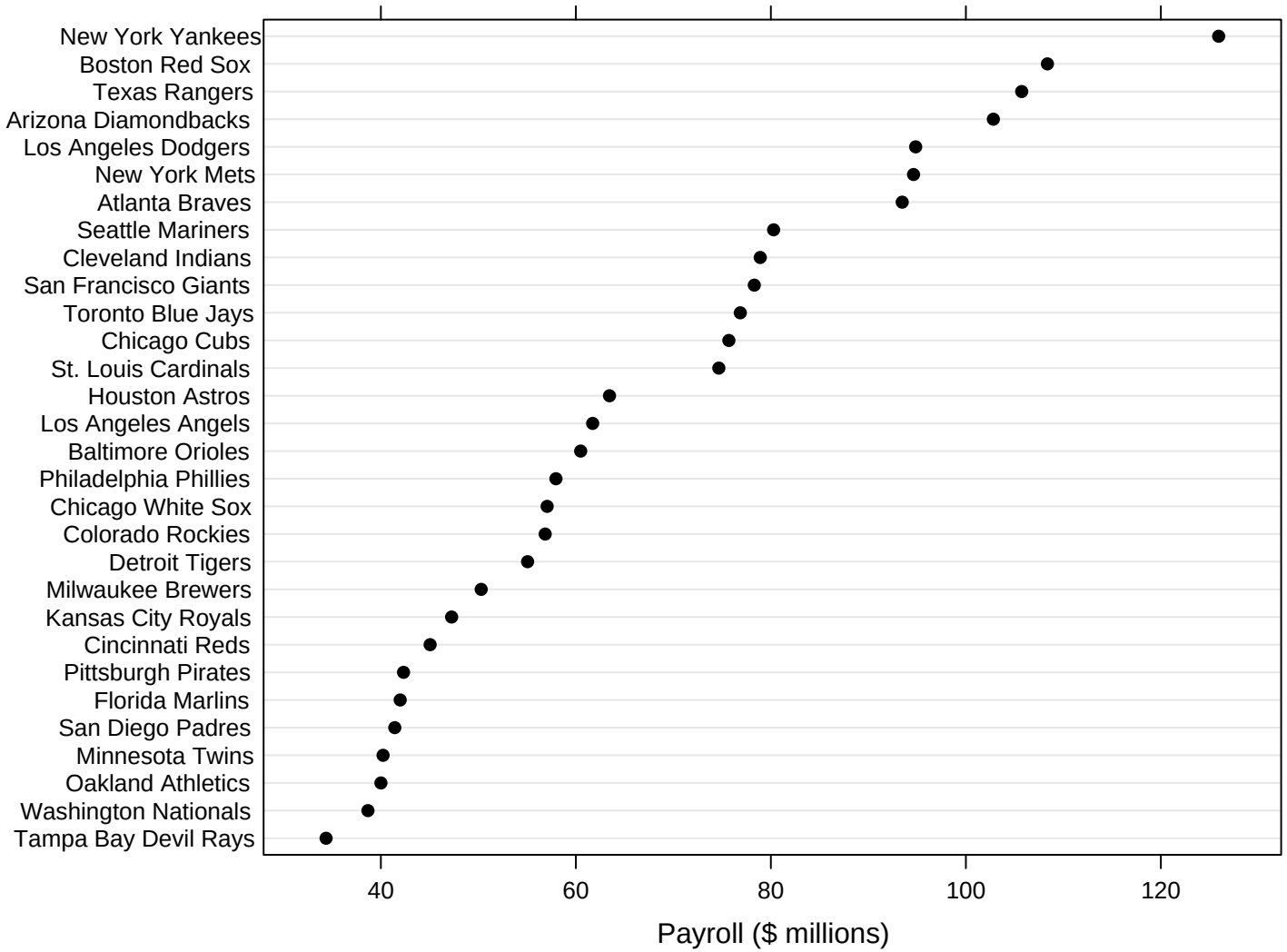
The constraints

- What are the constraints of a major-league baseball team?

The constraints

- What are the constraints of a major-league baseball team?
- The overriding constraint for the 2002 Oakland A's is their budget

2002 MLB payrolls



Some baseball basics

- Hits: Single, double, triple, homerun
- Batting average (AVG)

$$\text{Hits}/\text{At-Bats}$$

- Runs batted in (RBI)
- On-base percentage (OBP)

$$(\text{Hits} + \text{Walks})/(\text{At-Bats} + \text{Walks})$$

- Slugging percentage (SLG)

$$(1 \times \text{Singles} + 2 \times \text{Doubles} + 3 \times \text{Triples} + 4 \times \text{Homeruns})/\text{At-Bats}$$

The main characters

Billy Beane

- Former player
- HS Education
- Joined the A's in 1990 and was promoted to GM in 1997

Paul Depodesta

- Harvard Econ major
- Assistant to Billy Beane

The amateur draft

- Teams take turns picking players
- Mainly high school and college players
- After being drafted, a player usually spends time in the minor leagues
- How do you predict which players will be successful in the majors?

The old-fashioned way

Scouting

The fallibility of observation

- Consider a .300 hitter and a .275 hitter

The fallibility of observation

- Consider a .300 hitter and a .275 hitter
- The difference: one hit every two weeks

The fallibility of observation

- Consider a .300 hitter and a .275 hitter
- The difference: one hit every two weeks
- Even if you were to watch 15 games, there's a 40% chance that the .275 hitter would have more hits

The fallibility of observation

- Consider a .300 hitter and a .275 hitter
- The difference: one hit every two weeks
- Even if you were to watch 15 games, there's a 40% chance that the .275 hitter would have more hits
- "The difference between a good hitter and an average hitter is simply not visible." (Bill James, 1977)

Data

- Instead of relying on observation, you can look at the data
- But what data should you look at?
- There's lots of data... but lots of it is meaningless
- The classic baseball statistics for a batter is the batting average
- But how is batting average related to the A's objective?

Model 1

- The first model relates making the playoffs to the number of games the team wins.
- 95 wins = playoffs

Model 2

- The second model relates wins to the number of runs scored and allowed by the team.
- Runs scored - runs allowed > 135

Model 3

- Finally, the third model relates runs scored to individual statistics.
- (They also have a model for runs allowed.)
- Runs scored are predicted by OBP and SLG

The Mathematical Edge

- These models give the A's an edge because they are able to more accurately value players
- They find that OBP and SLG are undervalued by the market
- Other attributes, such as speed and fielding, are overvalued by the market

Model predictions

- DePodesta's prediction for 2002 is that the A's:

	Prediction	Actual
Score:	800-820 runs	
Allow:	650-670 runs	
Win:	93-97 games	

Model predictions

- DePodesta's prediction for 2002 is that the A's:

	Prediction	Actual
Score:	800-820 runs	800 runs
Allow:	650-670 runs	
Win:	93-97 games	

Model predictions

- DePodesta's prediction for 2002 is that the A's:

	Prediction	Actual
Score:	800-820 runs	800 runs
Allow:	650-670 runs	653 runs
Win:	93-97 games	

Model predictions

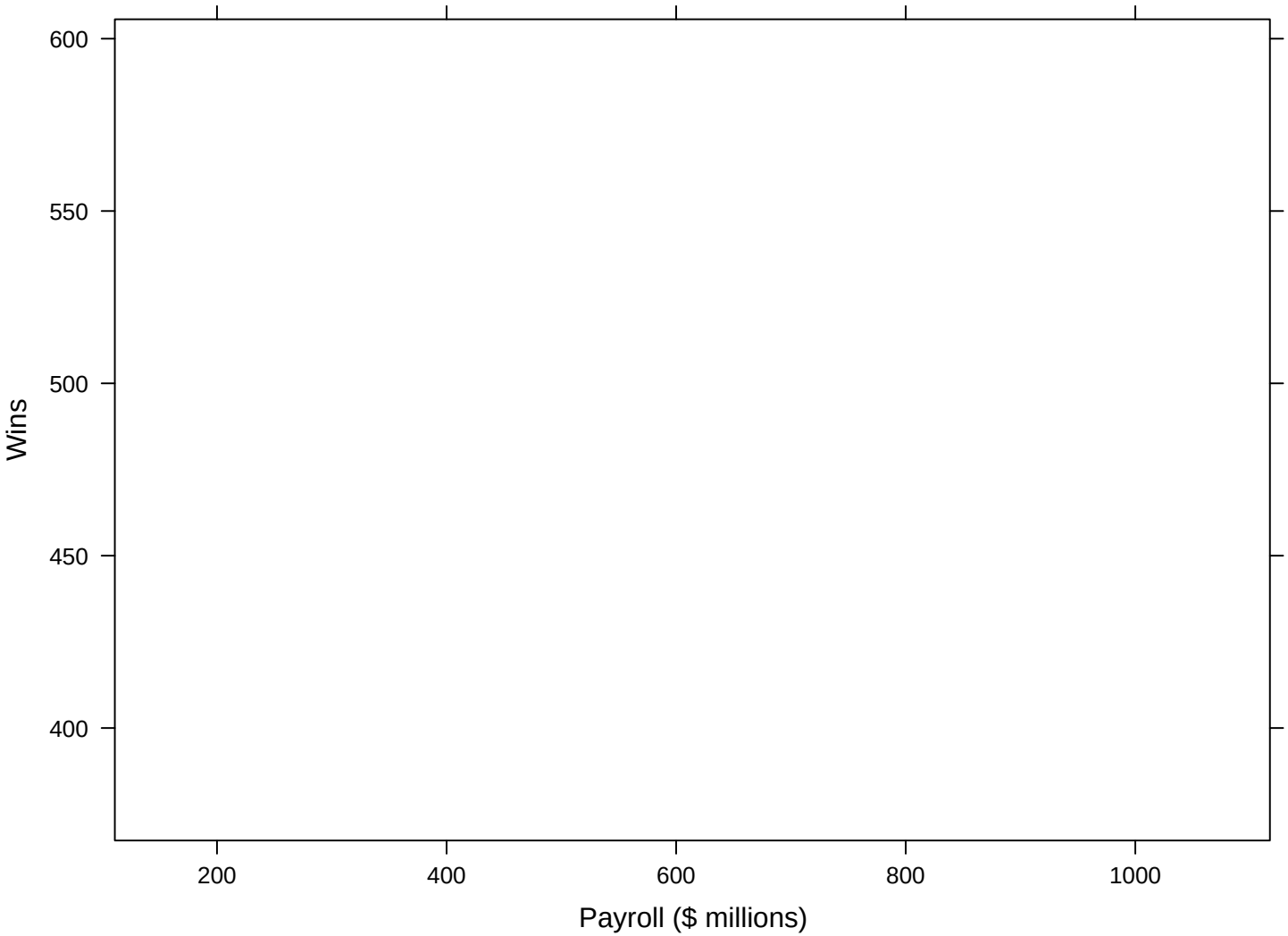
- DePodesta's prediction for 2002 is that the A's:

	Prediction	Actual
Score:	800-820 runs	800 runs
Allow:	650-670 runs	653 runs
Win:	93-97 games	103 games

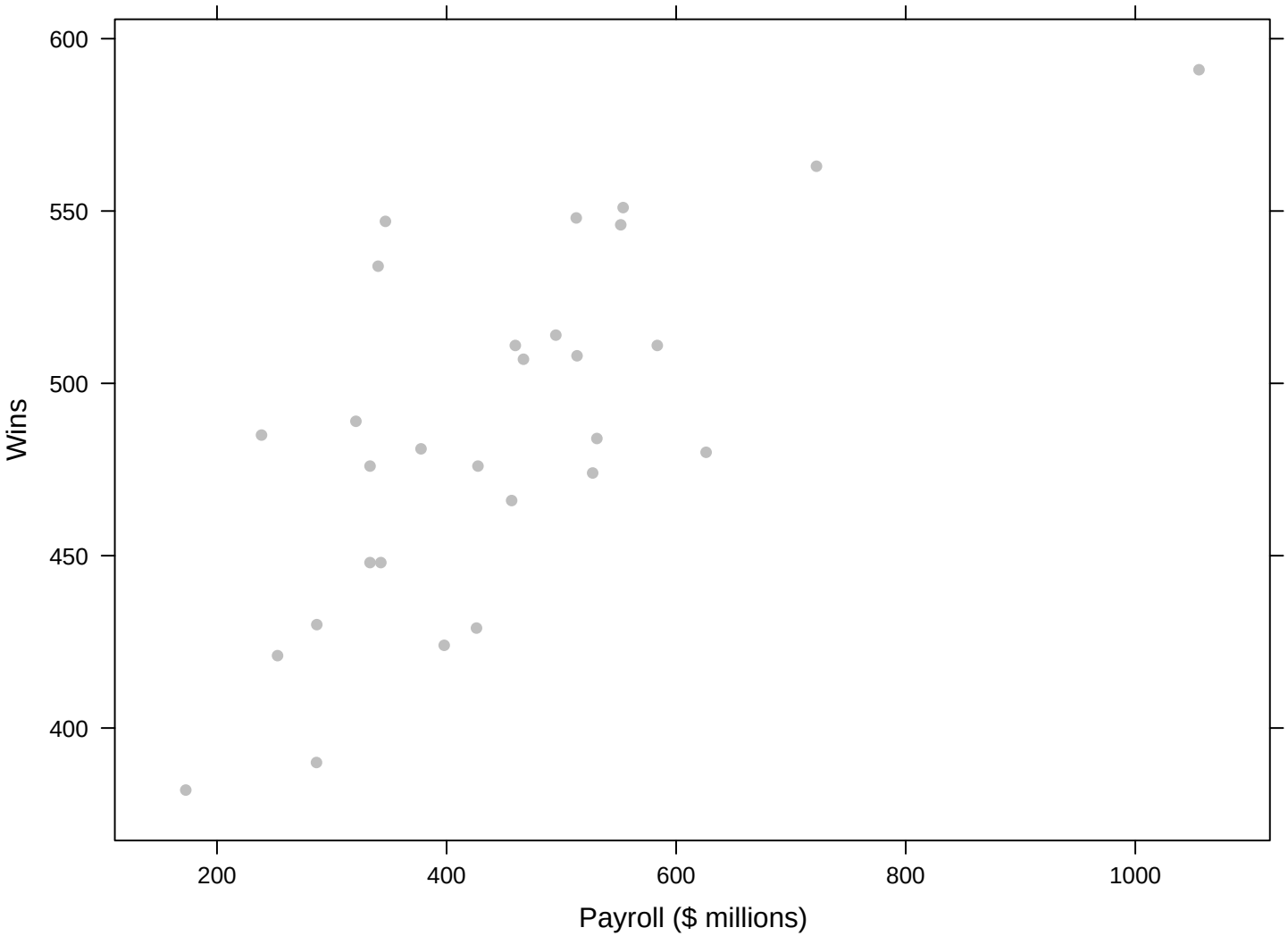
2002-2007

Season	Record	Playoffs
1999	87-75	
2000	91-70	First round
2001	102-60	First round
2002	103-59	First round
2003	96-66	First round
2004	91-71	
2005	88-74	
2006	93-69	Second round
2007	76-86	

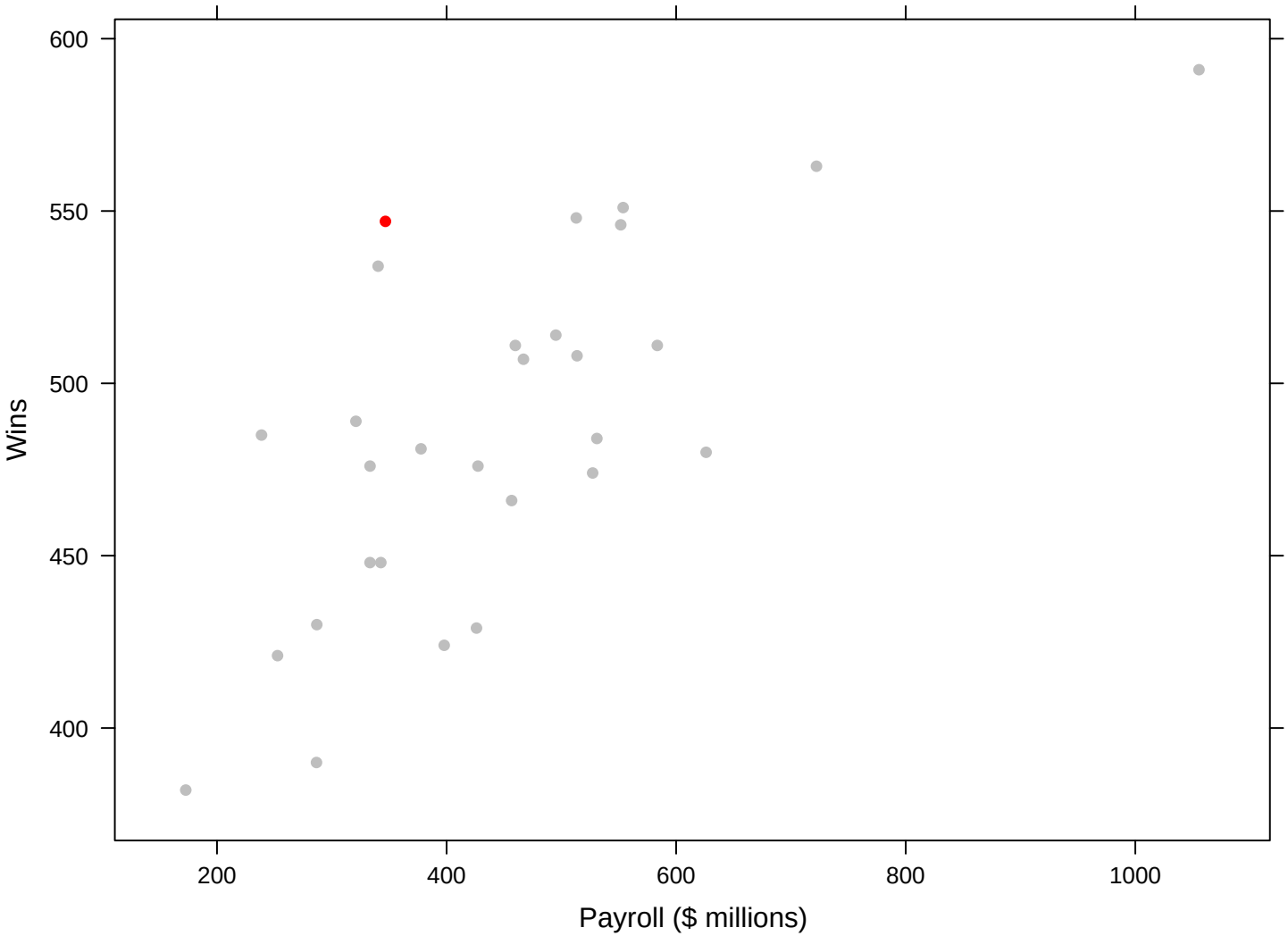
Payroll vs. Wins (2002-2007)



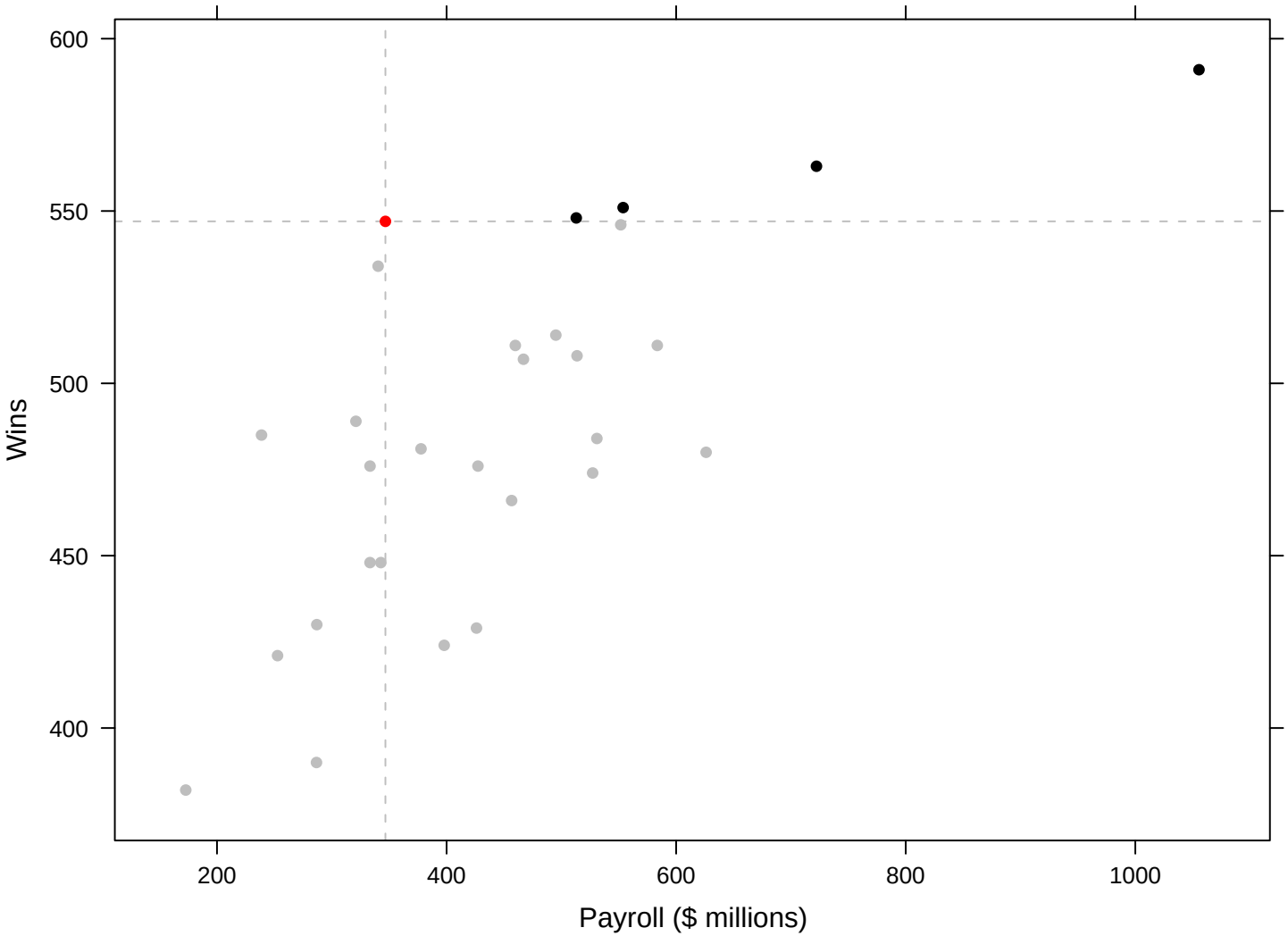
Payroll vs. Wins (2002-2007)



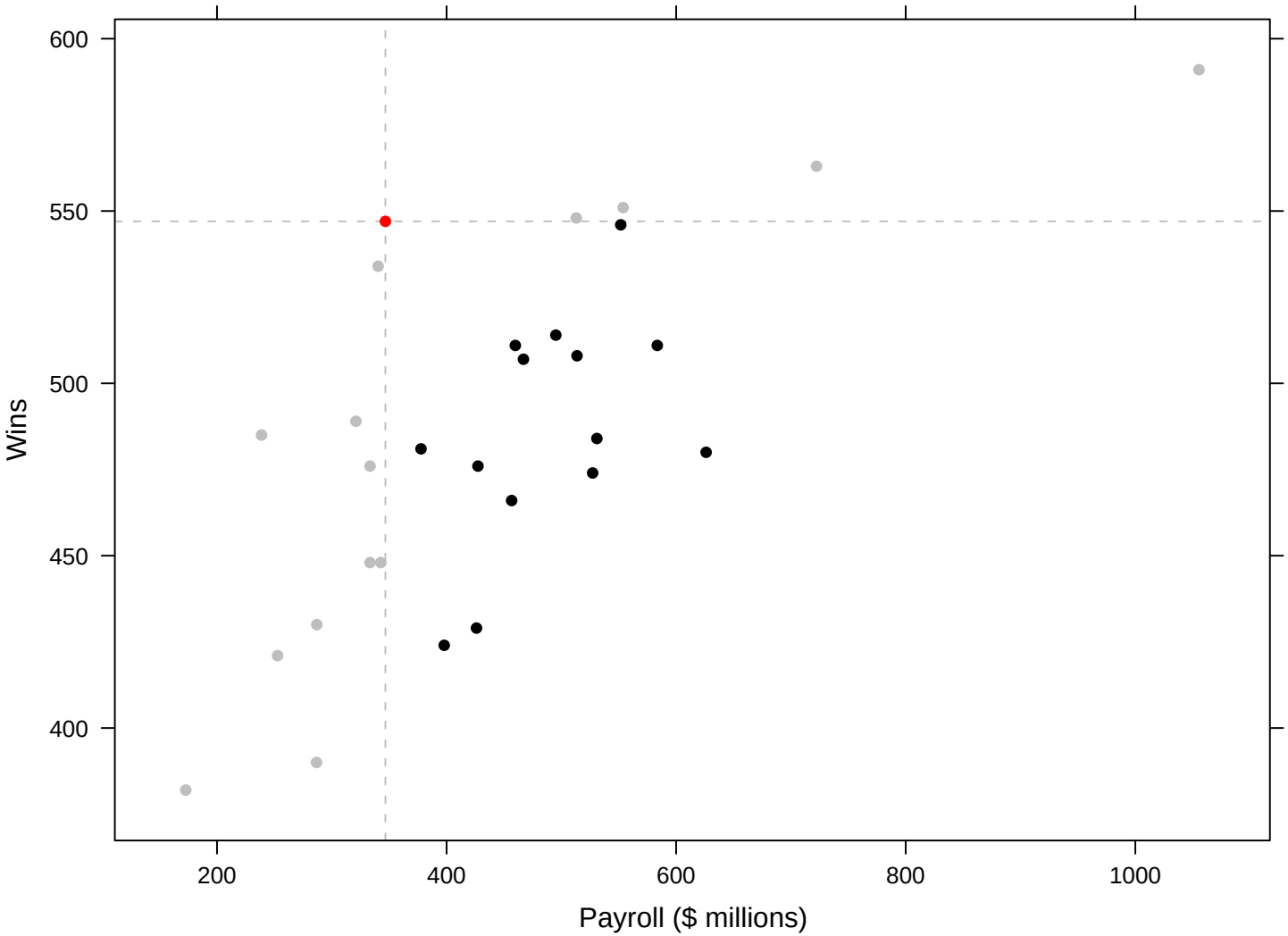
Payroll vs. Wins (2002-2007)



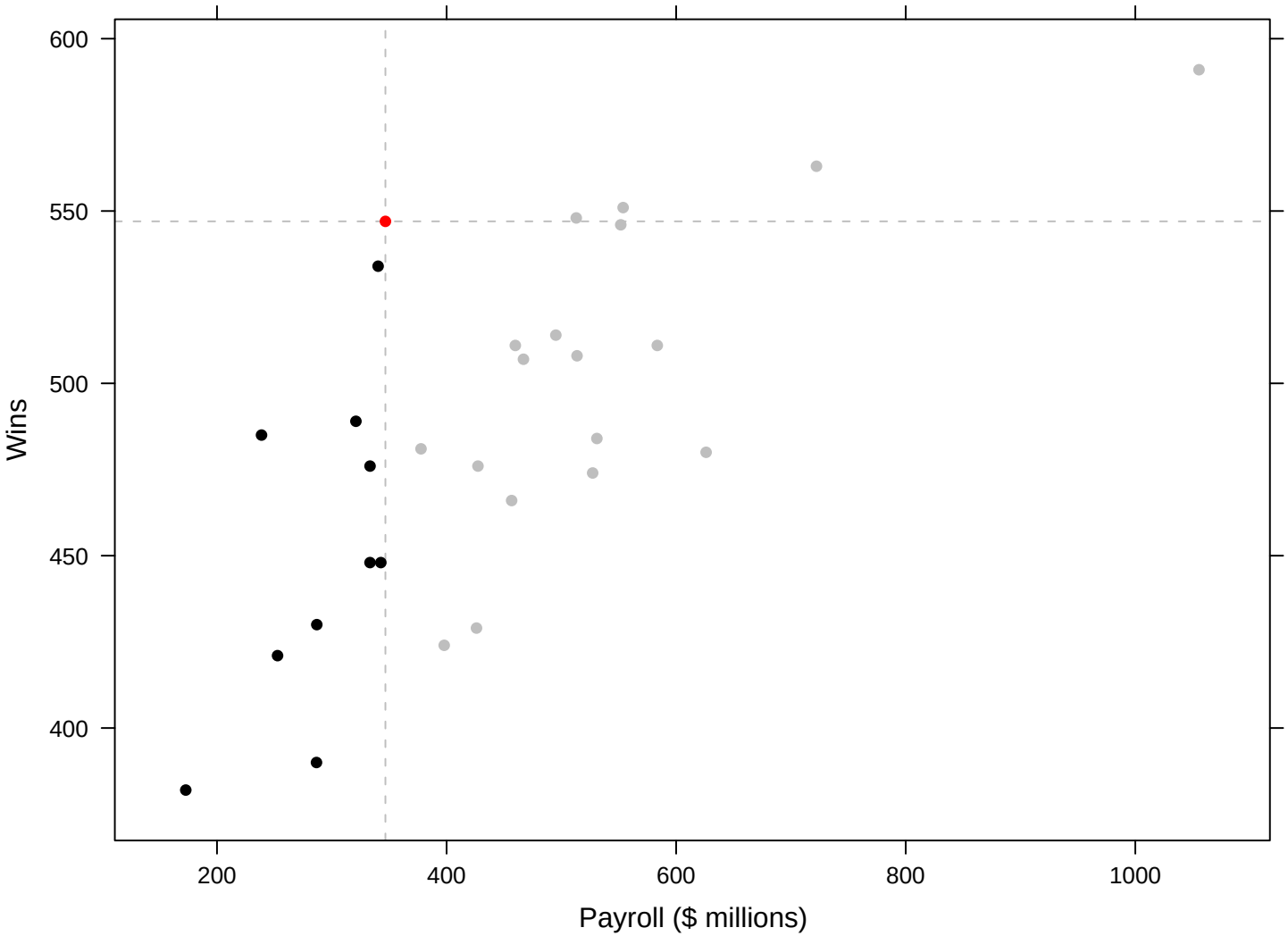
Payroll vs. Wins (2002-2007)



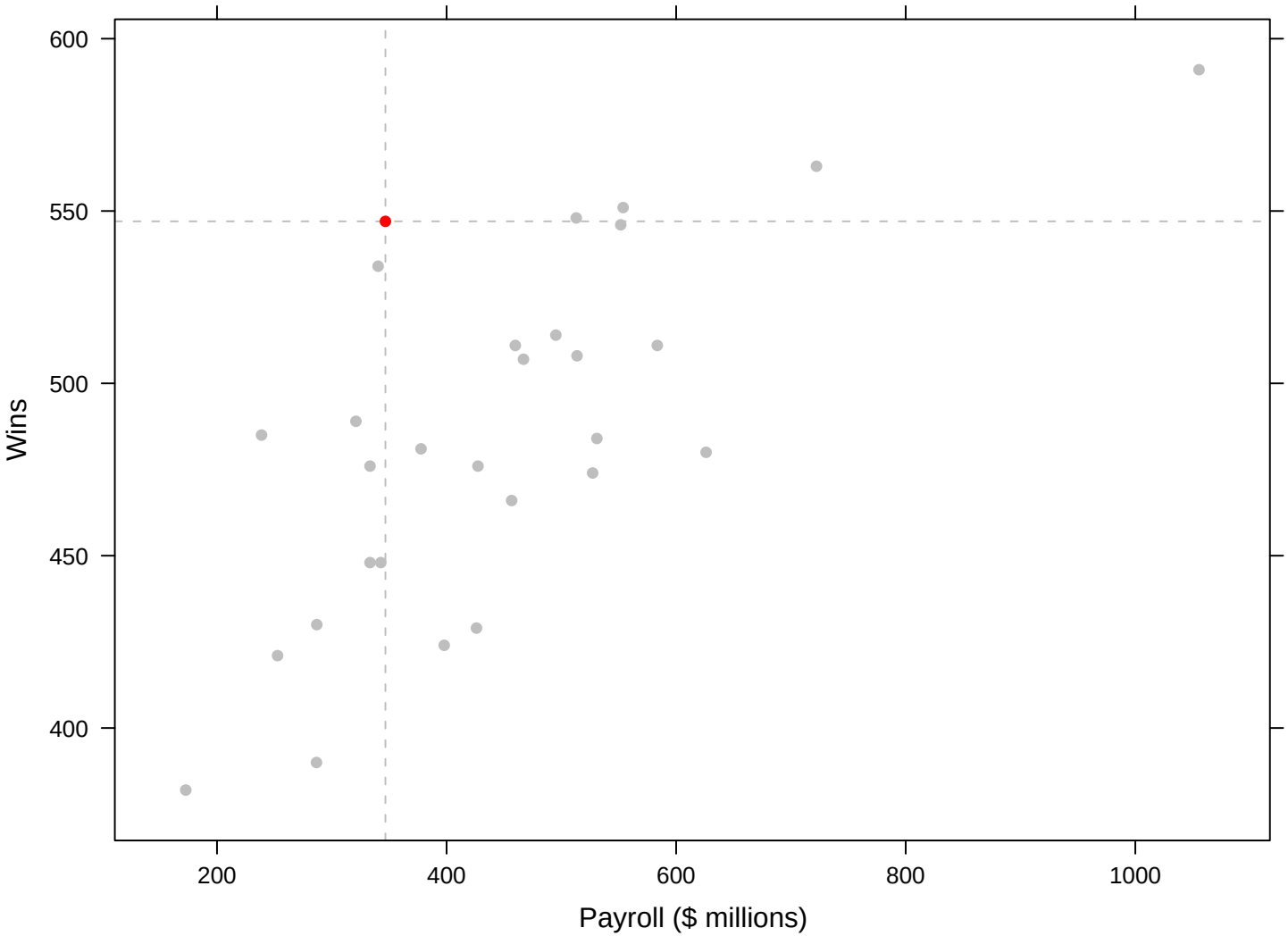
Payroll vs. Wins (2002-2007)



Payroll vs. Wins (2002-2007)



Payroll vs. Wins (2002-2007)



Quantitative methods today

- Every major-league baseball team now has a statistics group
- Perhaps 75% of teams actually incorporate quantitative methods into their decisions
- Quantitative methods are at the core of the decision-making process for roughly 25% of the teams
- On-base percentage is no longer undervalued
- A handful of basketball teams have begun to use quantitative methods
- At least one football team is using quantitative methods

Other Types of Decisions

Gametime Decisions

- Stealing base
- Changing pitchers
- Intentional walk

Off the field decisions

- Investing in a new stadium
- Television contracts
- Ticket prices

Further resources

- “The Numbers Game” by Alan Schwarz
- Baseball Prospectus
- tangotiger.net
- Society for American Baseball Research (sabr.org)

Airline Safety

- Over the period 1983 - 2002 there were 90 million Domestic US flights

Airline Safety

- Over the period 1983 - 2002 there were 90 million Domestic US flights
- The mortality risk on these flights was 1 in 11 million

Airline Safety

- Over the period 1983 - 2002 there were 90 million Domestic US flights
- The mortality risk on these flights was 1 in 11 million
- Choosing a flight at random each day, you could expect to fly for 29,000 years before succumbing to a fatal accident

Airline Safety

- Over the period 1983 - 2002 there were 90 million Domestic US flights
- The mortality risk on these flights was 1 in 11 million
- Choosing a flight at random each day, you could expect to fly for 29,000 years before succumbing to a fatal accident
- But are all of the airlines equally safe or are some safer than others?

Let's start by looking at the data

Year	Airline	FCEs	Year	Airline	FCEs
1985	Delta	0.83	1991	US Airways	0.24
	Midwest Express	1.00		United	1.00
1987	Northwest	0.99	1992	US Airways	0.53
	Continental	0.32	1994	US Airways	0.71
1988	Delta	0.12		US Airways	1.00
1989	United	0.40	1996	ValuJet	1.00
	US Airways	0.04		Delta	0.01
1990	Northwest	0.18	1999	American	0.07

- FCE - Full-Crash Equivalent

Distribution under Equal Safety Hypothesis

- Let p_i be the fraction of flights performed by airline i
- Let a_i be airline i's FCE score
- For example:
  - $P(a_i = 0) = (1 - p_i)^{16}$
  - $P(a_i = 8.44) = p_i^{16}$
- If exactly two events out of the 16 killed 1% of the passengers on board, then
  - $P(a_i = 0.01) = 2p_i(1 - p_i)^{15}$
- Rounded to the nearest hundredth, there are 825 distinct values that a_i can take on

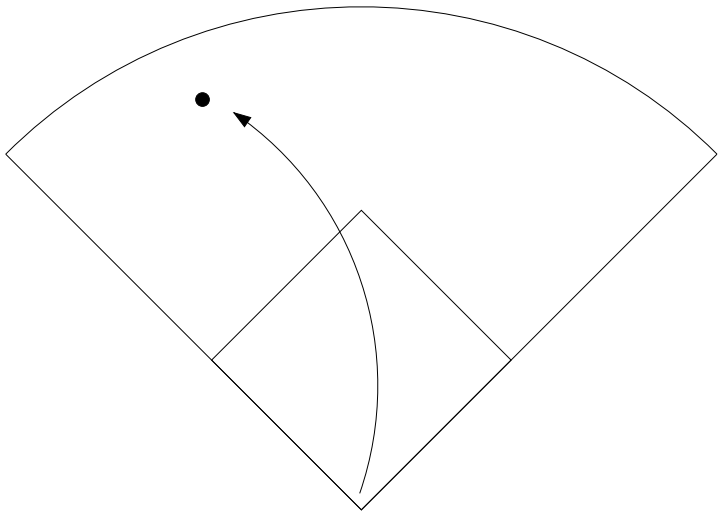
Table under Equal Safety Hypothesis

Airline	Expected FCE's	Actual FCE's	Percentile
American	1.23	0.07	11%
Continental	0.69	0.32	45%
Delta	1.46	0.96	30%
Northwest	0.82	1.17	73%
Southwest	0.96	0.00	7%
United	1.19	1.40	64%
US Airways	1.20	2.52	91%
New Entrants	0.88	2.00	90%

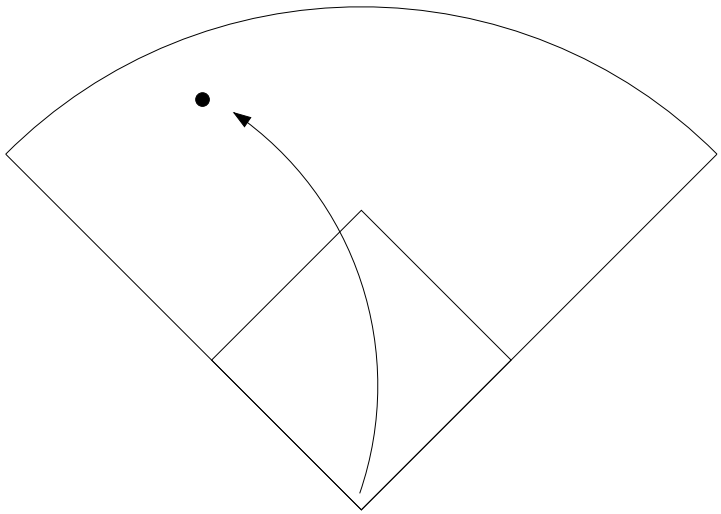
Test for Statistical Significance

- Let π_i be the observed percentile of the ith airline
- Let $D = \sum_{i=1}^{8} (\pi_i - 50)^2$
- Under the equal safety hypothesis, the 95th percentile of D is 10,250
- The computed value from our table is 7,754
- The equal safety hypothesis survives the test

Enter *Moneyball*



Enter *Moneyball*



- 20% of the time - out
- 40% of the time - single
- 20% of the time - double

A Restatement of the Equal Safety Hypothesis

- Let $P(E)$ be the probability of a life-threatening emergency
- And let $P(D)$ be the probability of a passenger's death
- Then,

$$P(D) = P(E)P(D|E)$$

- Now we'll look at the larger set of data concerning $P(E)$, using the approach suggested by the book *Moneyball*
- Equal Exposure Hypothesis: On all the airlines being compared, a passenger faces the same potential for death due to life-threatening emergencies

Fatal Accidents by Category

Description	Count	Total FCE's	Average FCE's Per Event
In flight collision with object	13	0.99	0.076
In flight collision with terrain	3	0.71	0.237
In flight encounter with weather	98	0.83	0.008
Loss of control in flight	6	0.98	0.163
Loss of control on ground	18	0.11	0.006
Malfunction	218	3.40	0.016
On ground collision	89	0.42	0.005
Fire/Explosion	22	1.00	0.045

Table under Equal Exposure Hypothesis

Airline	Expected Score	Actual Score	Percentile
American	1.23	1.59	91%
Continental	0.69	0.93	88%
Delta	1.46	1.52	59%
Northwest	0.82	0.84	56%
Southwest	0.96	0.34	0.1%
United	1.19	1.01	26%
US Airways	1.20	1.15	44%
New Entrants	0.88	1.05	79%

Table under Equal Exposure Hypothesis

Airline	Expected Score	Actual Score	Percentile
American	1.23	1.59	91%
Continental	0.69	0.93	88%
Delta	1.46	1.52	59%
Northwest	0.82	0.84	56%
Southwest	0.96	0.34	0.1%
United	1.19	1.01	26%
US Airways	1.20	1.15	44%
New Entrants	0.88	1.05	79%

- The test statistic D is 7,296. The equal exposure hypothesis survives the test

Conclusions

- We've tested the equal safety hypothesis in several ways
- By carefully expanding the data set, we were able to perform a more powerful test
- On the whole, there is not evidence that there are differences in safety between the airlines