

Quality, Statistics, and Probability in Sports

DAN RAND

ASQ 1216 – DECEMBER 11, 2009

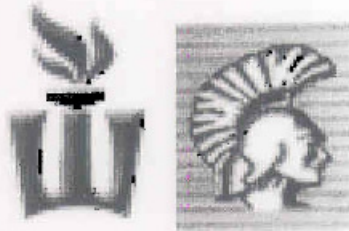
Quality – in sports card printing and manufacturing

Statistics - in sports research (?) and reporting

Probability – everywhere in sports

Use them all to improve your bowling and golf

Predicting Outcomes In Football



Alan Hartung
Winona State University
Stat 488 Project

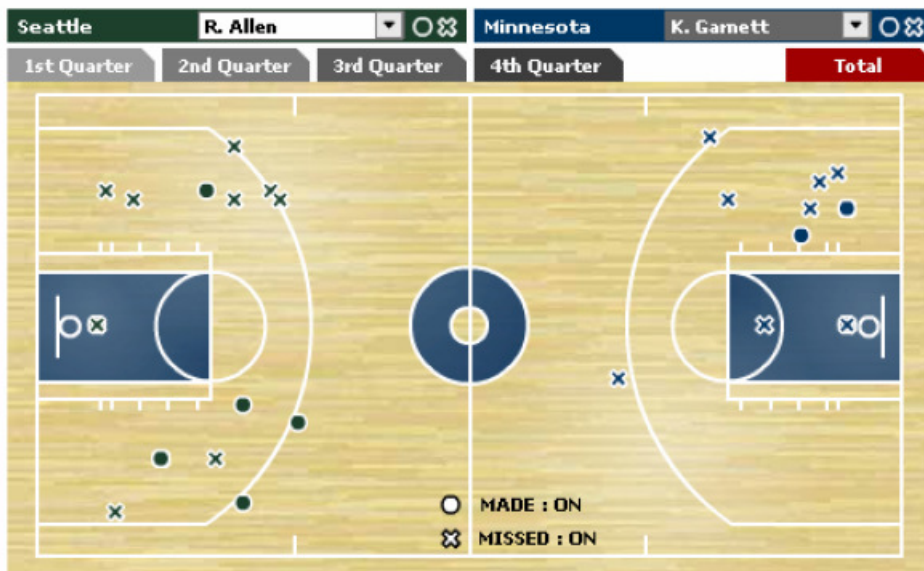
The main purpose of this work is to see if the prediction of offensive outcomes based on certain situations in the game. Some of these outcomes include, scoring, success of plays, and yardage gained. These predictions may help coaches make crucial decisions. The data collected for this work is from the Winona State football team 2002 season. Every offensive play for the entire season has been recorded. A total of 860 plays over 12 games. Three data sets will be considered in this work. The first consists of all plays. The second consists of only first down plays. The third consists of plays where a touchdown is scored. The investigation of these data sets allows us to isolate certain situations. This in turn will hopefully produce better models and make prediction better. Several variables were recorded: quarter, time, down, distance, yards to goal line, personal, formation, pass or run,

yards gained, outcome, success, and point differential.

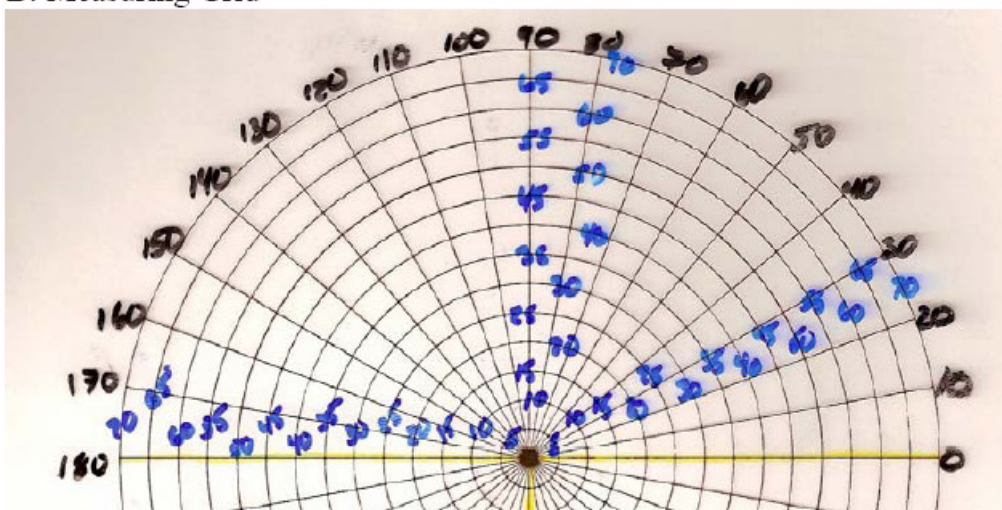
Jenni Hutchison - Sports Statistics

Abstract:

This is a discussion of the use of circular statistics on basketball shooting. Two NBA players were used to run the tests and compare their performances. This project looks at all the shots taken by two different NBA players for a given period of time, comparing the makes and misses. The goals were to find an average angle at which a player shoots the most often and has a high percentage of makes versus their average angle for missed shots.



B: Measuring Grid



Tests of Significance - we start another aspect of statistical inference, the **Hypothesis Test**.

Free throw example – my brother claims he can make 80% of his free throws, but in trying to prove it he makes 8 of 20, do you believe his claim?

Hypothesis tests are all about setting rules to decide if evidence supports a claim, and giving us some numbers derived from sample data to help decide the close calls.

"An outcome that would rarely happen if a claim (a null hypothesis) were true is good evidence that the claim is not true."

What is the probability that I will have the best poker hand in Texas Hold'em after the river card is turned up,

if I have Ace-Queen,

my only remaining opponent has Jack-Jack,

and the 4 community cards (the flop and the turn) are 4, 7, 8, King, all 4 of different suits?

Belichick's 4th and 2

The New York Times - 5th Down Blog

(Brian Burke, a former Navy pilot who has taken up the less dangerous hobby of N.F.L. statistical analysis, operates [Advanced NFL Stats](#), a blog about football, math and human behavior.)

The Patriots had a better chance to win the game by going for it on 4th and 2 from their own 28 than they would have if they'd punted:

With 2:00 left and the Colts with only one timeout, a successful conversion wins the game for all practical purposes. A 4th-and-2 conversion would be successful 60% of the time. Historically, in a situation with 2:00 left and needing a TD to either win or tie, teams get the TD 53% of the time from that field position. The total win probability for the 4th-down conversion attempt would therefore be:


















$$(0.60 * 1) + (0.40 * (1-0.53)) = 0.79 \text{ WP}$$

A punt from the 28 typically nets 38 yards, starting the Colts at their 34. Teams historically get the TD 30% of the time in that situation. So the punt gives the Pats about a 0.70 WP.

Statistically, the better decision would be to go for it, and by a good amount. You'd have to expect the Colts had a better than 30 percent chance of scoring from their 34, and an accordingly higher chance to score from the Pats' 28. But any adjustment in their likelihood of scoring from either field position increases the advantage of going for it. You can play with the numbers any way you like (**I did**), but it's pretty hard to come up with a realistic combination of numbers that makes punting the better option. At best, you could make it a wash.

P(NE wins by going for it)	P(NE wins by punt)	P(Indy scores from NE 28)	P(Indy scores from Indy 34)
0.788	0.7	0.53	0.3
0.76	0.6	0.6	0.4
0.72	0.6	0.7	0.4
0.68	0.6	0.8	0.4
0.64	0.6	0.9	0.4
0.64	0.7	0.9	0.3

From [Advanced NFL Stats](#),

AFC Percent Probability Playoff Seeding							
Team	1st	2nd	3rd	4th	5th	6th	Total
 IND	99	1	0	0	0	0	100
 SD	0	51	18	9	12	6	96
 NE	0	18	34	46	0	0	98
 CIN	0	16	32	35	2	8	93
 DEN	0	12	7	3	32	26	80
 PIT	0	2	8	3	30	21	63
 BAL	0	0	1	3	16	18	39
 NYJ	0	0	0	1	0	3	4
NFC Percent Probability Playoff Seeding							
Team	1st	2nd	3rd	4th	5th	6th	Total
 NO	86	14	0	0	0	0	100
 MIN	14	81	4	1	0	0	100
 DAL	0	1	36	13	9	22	81
 ARI	0	2	22	60	1	2	86
 PHI	0	2	29	3	19	24	76
 NYG	0	0	8	8	9	20	45
 GB	0	0	0	0	60	25	85
 SF	0	0	1	16	1	3	21
 ATL	0	0	0	0	1	4	6

The guy behind MNF stats- “**Steve Hirdt**, who is in his 28th year supplying statistical girding for "Monday Night Football," still loves the challenge of giving facts on the fly during telecasts of the games” the executive vice president of the Elias Sports Bureau

"As the game ended, before Mike Tirico threw it to our sideline reporter for an interview with **Drew Brees**, we were able to put Brees' final numbers on the air and determine that there have been **only four games in the history of the NFL in which a quarterback had completed 75% of his passes, thrown for over 350 yards, with five touchdowns and no interceptions.** All three of the others were by **Tom Brady.**"

Elias Sports Bureau

STATS LLC

ESPN fans fielding pct

Title: Better Golf Through Better Data Analysis

Copyright: 1994, ASQC

Author: *Whitaker, Jack*

Organization: *Process Management International, Minneapolis, MN*

Subject: Measurement and control, Root cause analysis (RCA), Individuals charts, Data analysis, Control charts, Continuous improvement (CI), Check sheets;

Series: *Quality Progress*, Vol. 27, No. 2, February 1994, pp. 53-58

Abstract: Analysis of data to identify causes of performance problems can lead to continuous process improvement. Through an extended example, the author demonstrates how careful data analysis of golf scores can help golfers increase their performance. As in many data analyses, it is critical to collect useful data. Baseline data should include at least 36 holes, all on the same or similar courses, collected over no more than 30 days. Expand the typical golf score card and gather data like the following for each shot: stroke type, club used, and qualitative shot results (acceptable, unacceptable, and gimme). Data analysis involves check sheets, control charts, and Pareto diagrams. Examination of the charts should suggest causes that affect performance. It is best that the golfer focus improvements on one cause at a time. In the example, it was most appropriate to work on approach/chip shots and then to improve driving accuracy. This method allows those with a handicap of 15 or less to improve by three to five strokes per round over one season.

A Personal Quality Improvement Program for Golfers

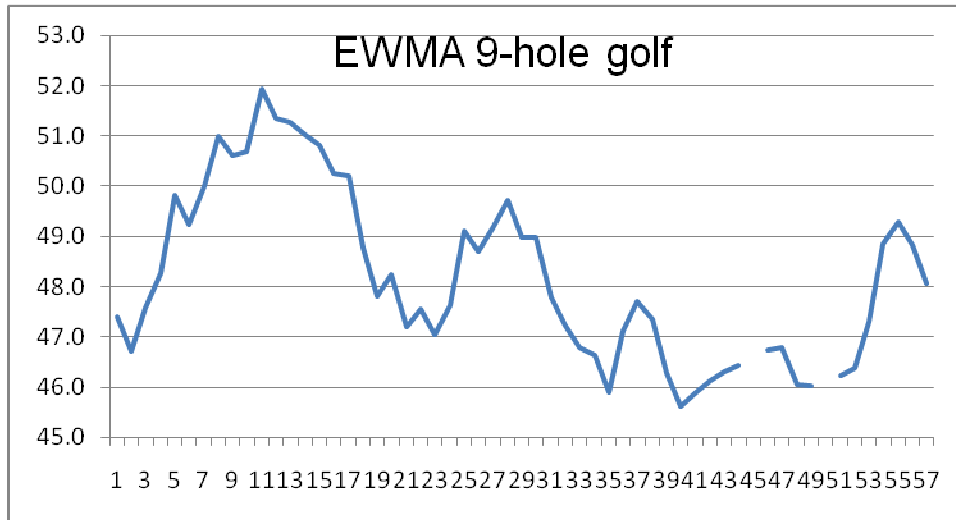
by Karlin, Ernest W.; Hanewinkel, Ernie, 1998-07; Quality Progress; 31(7)

Figure 1. Course-Specific Quality Improvement Data Card for Single Round of Golf

Player: _____ Course: <i>GCC West</i> Tees: <i>Blue</i> Date: _____																		Tee Shot Quality A. Good distance in fairway or on green B. Good distance (in rough or par 3 bunker) C. Poor distance (but playable) D. Shot out of bounds or in water hazard					
Hole	1	2	3	4	5	6	7	8	9	Out	10	11	12	13	14	15	16	17	18	In	Tot	Hcp	Net
Yardage	374	322	490	332	352	157	518	165	367	3077	364	336	543	342	173	336	354	180	531	3159	6236		
Par	4	4	5	4	4	3	5	3	4	36	4	4	5	4	3	4	4	3	5	36	72		
Handicap	4	14	8	10	6	18	2	16	12		9	11	1	13	17	7	5	15	3				
Gross score																							
Net score (ESQ adj.)																							
Tee-shot quality																				A/	B/	C/	D/
Strokes from 100 yards																							% of net
Putts on green																							% of net
Sand saves																							
Pitch/chip saves																							
>3																				Major problems:			
3																							
2																							
1																							
0																							
minus par																							
-1																							
-2																							

©KARLIN Associates, Inc., Delray Beach, FL, 1997

Net score: Gross score adjusted for equitable stroke control
Tee shot quality: A, B, C, or D per definitions in top right corner of card
Strokes from 100 yards: Total strokes from 100 yards into the cup
Putts on green: Total number of putts after ball is on green
Sand saves: Two strokes or less from green-side bunker into the cup
Pitch/chip saves: Two strokes or less from off the green into the cup
Net score minus par: Plot value of net score minus par for the hole



score	In 100 yds	greens in reg	score	In 100 yds	greens in reg
51	+5	1	47	+2	1
52	+2	0	47	+1	0
46	+2	1	47	+1	0
49	+2	0	48	+4	1
43	-2	3	47	0	1
45	-3	2	43	-1	2
45	-2	2	46	-1	2
46	0	0	47	+1	1
43	-3	2	47	+4	2
52	+5 S4	1	51	+6	1
50	+3	0	55	+8	0
46	0	1	51	+6	0
42	-3	0	47	+2	0
43	-2	2	45	-1	1

Bowling control chart

