The State Championship Standard: How Statistics Can Get You There

Rob Beam
Head Volleyball Coach
Oakland University

1. Introduction

Wikipedia defines statistics as "a mathematical science pertaining to the collection, analysis, interpretation, explanation, and presentation of data." In describing their use, the website continues, "Statistics provide tools for prediction and forecasting." While many people will stare blankly at me or perhaps run in the opposite direction when I admit this, but this is exciting news to me! In fact, it is so exciting I think that I can help each of you move closer to a state championship based on understanding statistics.

In this session I will show you how to manipulate the information contained in a basic box score to make it more meaningful. Along the way we will discover the factors or components that most strongly correlate with winning and losing. We will also debunk long perpetuated volleyball myths. And, finally, we will take this information and use it to shape 17 different aspects of our program including our drill designs, match evaluations, player interactions, and more.

2. A Statistical Caveat: The best we can do is correlation

    a. Definition of Correlation: "A relationship in which two or more things are mutual or complementary. In statistics, the degree to which two or more variables are related and change together." (Encarta Dictionary)

    b. Definition of Causality: "Is the process of making something happen. It denotes a necessary relationship between one event, cause, and another event, effect." (Wikipedia)

    c. The unfortunate truth is that we cannot create causality. Meaning, if we simply kill more balls, although that would be a good start as we will learn later, it will not cause a win. At best an increase in kills/set would correlate strongly with winning. While this is a small distinction, it is a necessary one.

3. Box Score Basics

- K = Kill (successful attack)
- E = Attack Error (unforced or stuffed attack)
- TA = Total Attempts (all attack attempts combined)
- SA = Service Ace (a serve that leads directly to a point)
- SE = Service Error (a serve that leads directly to an error)
- RE = Receive Error (a serve reception that leads directly to a point)
- Total Team Blocks (total number of BS (block solo) and BA (block assists)
- BE = Block Error (Net or centerline violation during a block motion)
- GEN BHE = General Ball Handling Error

4. The Stats Inside the Stats:

    a. Positive Points = K+SA+TTB
    b. Negative Points = E+SE+RE+BE+BHE
    c. + Points/Set = (K+TTB+A)/# of Sets
    d. – Ponts/Set = (E+SE+RE+BE+BHE)/# of Sets
    e. Differential (Dif) = (K+TTB+A)-(E+SE+RE+BE+BHE)
    f. Dif/Set = (K+TTB+A)-(E+SE+RE+BE+BHE)/# of Sets
    g. Hitting Efficiency (Hit Eff) = (K-E)/TA
    h. Kill % = K/TA
    i. Error % = E/TA
    j. SA:SE Ratio = SA/SE (1.00 = the number of SA = the number of SE, >1.00 = more SA than SE, < 1.00 = more SE than SA)
    k. SA:RE Ratio = SA/RE (1.00 = the number of SA = the number of RE, >1.00 = more SA than RE, < 1.00 = more RE than SA)

5. The State Championship Standard

What if I told you that based on what we have already learned we have the information necessary to win a State Championship? While it might be hard to believe, it is all there. In order to accurately prepare for this presentation I evaluated 80 High School Matches at the State Semifinal and Finals levels in Michigan, Ohio, and Illinois. In Chart A you will find the results of this data collection. The whole spreadsheet is included as well and is sorted by Dif./Set.

In the left column you will see 8 critical statistics. There are then four additional columns grouped by either "Correlation to Winning" or "Correlation to Losing." In the former we find the information that correlates with a team winning a match. This is easily understood as "if you do this, you tend to win." In the latter, we find the number that correlates with losing. Meaning, if you are achieving at or below these levels, you are likely to lose.

The "Absolute" number is the threshold number where every team that achieved that level of performance won. While this is interesting information, it would be very difficult to train a team at these levels. (Even the undefeated 2008 Penn State National Championship team barely reached the +11.3 Dif/Set mark.) In two areas you will notice that there are question marks. That means that in that statistic there was no threshold for performance that correlated with an outcome.

The "Strong" number is the more practical number. This is the level of performance at which there was a high correlation with the corresponding outcome. These "Strong" numbers are where I believe we should focus our attention.

Chart A

	Correlation to Winning		Correlation to Losing	
	Strong	Absolute	Strong	Absolute
Kill % (K/TA)	33.30%	42.90%	30.10%	27.70%
Error % (E/TA)	14.20%	10.40%	17.00%	22.40%
Hit Eff. ((K-E)/TA)	0.172	0.281	0.150	0.130
+Pts/Set (K+SA+TTB)	15.50	21.67	14.00	13.00
-Pts/Set (E+SE+RE+BE+BHE)	8.00	5.00	10.25	?
Dif/Set ((K+SA+TTB)/(E+SE+RE+BE+BHE))	+7.00	+11.3	+4.67	+2.50
SA:RE (SA/RE)	+1.50	+3.0	+.75	+.33
SA:SE (SA/SE)	+1.50	?	+.43	+.2

6. What Statistics Matter Most?

Now that we have the standards at which we want to perform, it is important to know which statistics correlate the strongest with winning and losing matches. Chart B lists the 3 statistics that correlate best with winning. Chart C lists the 3 statistics that correlate best with losing. Chart D lists in order the statistics and their cumulative strength or importance to winning.

Based on the analysis in Chart D, it turns out that the Dif./Set is the most important number that you can know. It is also interesting to see that the +point side of this ratio finished second overall and is far more significant than the –points side. The third most important number is Hitting Efficiency. And surprisingly, the fourth most important measure of performance is a team's Ace to Receiving Error Ratio.

Chart B

Winning Correlative Value Ranking	
1. Dif./Set	96.55%
2. SA:RE	87.50%
3. Hit Eff.	84.21%
4. +Pts./Set	83.34%
5. -Pts./Set	83.34%
6. Kill %	81.25%
7. SA:SE	68.00%
8. Error %	67.86%

Data based on "Strong Correlation to Winning" Numbers. The percent was determined by dividing the number of wins by the total number of matches present for a given statistical measure.

Chart C

Losing Correlative Value Ranking	
1. +Pts./Set	96.42%
2. Dif./Set	90.32%
3. Hit Eff.	87.80%
4. Kill %	87.50%
5. SA:RE	80.00%
6. SA:SE	80.00%
7. Error %	78.26%
8. -Pts./Set	70.00%

Data based on "Strong Correlation to Losing" Numbers. The percent was determined by dividing the number of wins by the total number of matches present for a given statistical measure.

Chart D

Cumulative Correlative Value Ranking
1. Dif./Set
2. +Pts./Set
3. Hit Eff.
4. SA:RE
5. Kill %
6. -Pts./Set
7. SA:SE
8. Error %

Ranking was determined by adding the two ranked finishes of each statistical measure, with the lowest cumulative number being the best.

7. **Interpretations and Implications: How do I apply this knowledge?**

   a. Basic Losing Factor Analysis: It is of utmost important to understand what factor or factors statistically correlate with a losing outcome. Here is a real example to illustrate the utility of the numbers. In 2008, we averaged a +4.63/set in Dif/Set. When I looked closer at our point production it was clear that we needed to score more points in the attacking component of that number. Further examination revealed a need to increase our Hit. Eff. by increasing the Kill % side of the equation. This led me to look at our set distribution to see if our highest Kill% attackers were receiving enough balls. It turns out that less than 20% of our offense was set to our top 2 attackers, both middle hitters, in terms of Kill %. In our spring season we worked very hard to correct this and we were able to increase our middle distribution to over 38% of our offense. This led to a direct increase in Dif/Set as we were able to achieve a +7.56/set because of this new focus.

   b. Basic Winning Factor Analysis: If you are fortunate enough to lead a team that is winning it is helpful to understand why you are winning. By running these numbers you might find the areas of strength or advantage where you are beating your opponents. For instance, in our wins last season we averaged hitting over .241 with 38.5% kills. By using this number I can set a standard in practice for my starting group that will challenge them to rise to the occasion in training.

   c. The Serve-Pass Game: The first myth that we are able to debunk is that an ace to service error ratio is important. This particular myth is a favorite of TV broadcasters and I cringe every time I hear it. It turns out that it is not at all. In fact, it is so unimportant that the team with the best ace to service error ratio lost the match in my study! What really matters is an ace to receiving error ratio. Our teams must be trained so that they produce more points from the service line than they give up in the receive phase. This means you can literally lob the ball in if you are the best passing team in the state. But, if you make a lot of receiving errors, then a tough, point scoring service game is a necessity if you want to win. And, if you are really gunning for a conference or state championship the combination of high ace numbers and low receive errors is essential.

   d. Point Scoring Mentality: The second myth that we can expose as fraudulent is the one that suggests that, "the team that makes the fewest errors wins." This is simply not true. In every way the data suggests that it is the team that produces the most points that will win. While it is true that limiting negative points is helpful, it is much better to focus on kills, blocks, and aces. It is also true that attack error % was either last or second to last in each comparison. It simply doesn't correlate to winning.

   e. Feedback Control: Because we understand the factors that correlate most strongly with winning we can choose the feedback that we give our athletes more accurately. For instance, if I have an outside hitter that is struggling with her hitting efficiency, I am much more likely to focus on her finding a way to kill the ball. The reason is simple, the statistics show that limiting errors will not as strongly correlate to winning a match. I want to prepare this athlete to be on the floor and make a contribution so I will teach her a new shot or find a set tempo that works better for her in terms of Kill%.

   f. Match Evaluations: Before our staff meets with the team after a match we always run the numbers. When you have a zero-sum situation as we do in volleyball it is easy to exaggerate a win or a loss. While winning is good, that outcome does not necessarily indicate a strong performance. By way of example we won a match last year while hitting .118. Conversely, all losses do not necessarily indicate a bad performance. For example we lost a match with a Dif/Set of +8.33 and a Hit. Eff. of .182. I should praise my team for a great performance in this case.

   g. Individual Meetings: We can use this data on an individual level as well. We track K%, E%, Hit. Eff., RE%, and several other numbers for each player in practice and in competition. When we set goals at the beginning of a season or training cycle we reference these numbers. During the middle of the season we look at trends over time and see what effect our training is having on a particular player's performance. In a meeting with a player that is struggling to crack the line-up we can have a rational, objective discussion about playing time. We can identify the areas easily where they are falling short of our goal numbers and set up

a training plan for them to make increases. I find this is even helpful in selling the importance of strength and conditioning. If an athlete is behind her teammates in Kill% one of my stock answers is "get stronger." In this case, if the athlete is committed to being on the court, they will attack the weights like never before. While it took almost two years for every player to find comfort in the statistics, they like it now for these reasons. Everyone in our program knows where they stand and they know I am committed to helping them reach the level of performance that will correlate with their success.

h. Parent Meetings: While parent meetings are rare in the collegiate ranks, they are a reality in high school and club volleyball. Imagine having a tool at the ready to review an athlete's performance against the known standard when playing time is being disputed. In this way, we as coaches can soothe the relationship and provide tangible, objective evidence for our choices. Statistics even lend themselves to follow-up conversations in that you can tell a parent that their daughter has improved her hitting efficiency since your last call. In my experience this new information will recruit a challenging parent to your side.

i. Team Building: In a given season it is our job to get the absolute most out of our player's collective talents. One way to do this is to set goals for a team based on statistical performance. In our program we will deliver a monthly update so the players can see the team's progress toward our ultimate goal. This is a way for the team to rally together and work towards success. Another way to build a team is to make decisions based on the numbers. Perhaps during tryouts we select a player for the varsity team solely because she has the toughest serve. Another example is to track statistical performance by rotation and then start in the best side-out or point score rotation. There are dozens of ways that a team can be optimized and the fun of understanding the numbers is arranging the pieces of the puzzle to earn the desired outcome.

j. Program Building: It could be argued that building a program, meaning creating a lasting tradition that stretches long after a single team has completed its' season, is the ultimate in coaching. The New York Yankees, the Pittsburgh Steelers, and the Detroit Red Wings are professional examples of this status. By using this data we can track progress over multiple years. We can use it to set expectations and standards for every team. We can use it to develop our JV, freshman, and middle school players and coaches. For instance, we know that our Dif/Set has gone up from -.22 in 2006 to +4.13 in 2007 to +5.55 in 2008. That means we are growing and improving. We can also see that our Kill % has increased while we have been able to decrease our E% in the same span. On the down side we have identified our SA:RE Ratio as an area that needs attention. Therefore, we have worked on developing tougher more consistent serves for our returning players and recruited incoming athletes to address this specific need. In this way our program is developing from year to year based on our understanding of statistical performance.

k. Practice Planning: When we understand the factors that most highly correlate to winning, we can change the nature of our practice. Perhaps more time should be allocated to certain skill areas. For instance serving, serve receive, and attack are very important in terms of correlation to winning and could be given more time each day. Another idea would be to always train around the numbers. Instead of just going through a hitting line, have an individual player attempt to kill 3 of 9 tosses against a blocker and a defender.

l. Drill Design: When we design drills we can accurate set goals based on the numbers necessary to correlate with winning. Here are a few of my favorite statistical based drills:

  i. "Ace and Replace" (SA:RE Ratio Focus): Because SA:RE is of high importance we created this fun, simple drill. Three players begin on Side A of the net. The other competitors in the drill are serving on the opposite side of the net. The passers get to stay on if they can successfully receive the ball from the server. If they are aced (we define this as an "unsettable" ball or a "bad miss", meaning a pass that lands outside of the court) they must exchange places with the server. For every successful pass (not an ace) the passer makes they get a point. For every ace the server get one point and they get to become the passer. Because teams will serve away from the best passer once they figure out the game, it is helpful to rotate passers. Another solution to this problem is to award points to all three passers when a successful pass is made. A final solution would be to force the servers to alternate between all three passers.

ii. "4 v 6 FBK" (Kill% Focus): We know that Kill% has a stronger correlation to winning than limiting Error%, so we created this uneven numbers game to work on our First Ball Kill % (FBK). (It could be played 6 v 6, but this is a drill we used during our spring season when we had smaller numbers.) One team of 4 consisting of an outside or right-side, a middle, a libero, and a setter gets 10 opportunities to receive serve. (Using a free or down ball would make it easier, but would focus on transition percentages.) Their goal is to kill at minimum 4 balls out of this 10. We allow the 4 side to cover a blocked ball and continued to play out that rally. Also, we have the team of 6 transition any ball that is not killed to provide "game-like" scoring repetitions.

iii. "Lambo" (Point Score Differential Focus): If the ratio of +pts/set and -pts./set is examined it is approximately 2 to 1. With that in mind, my friend Chris Lamb, head coach at Wichita State, created this game and I named it after him. For this game the scoreboard starts at 3:3 for each team with the first "3" representing the number of +points and the second "3" representing the number of -points. Each time a team scores a +point the first "3" moves up. Each time they make earn a –point the second "3" is moved down. Play begins with a serve and score is kept for each team based on the result of each play. When a team wins on the +side, they get a BIG POINT and when they lose on the –side a BIG POINT is awarded to the other team. We most often play this game by staying in the same rotation until a team either wins or loses. Once that happens we award the BIG POINT and then the team that lost rotates. The team that is still playing stays in that rotation until they either win or lose. We play for the most points through all 6 rotations. Ties are decided by having each team pick a rotation and playing it out.

iv. "Mary Jo" (+Point Production Focus): This game presents two teams of six an opportunity to only focus on +point production. It is best understood as an alteration to the scoring system and not the flow of play. Meaning, each team plays regular rally score volleyball, but only +points in a row are recorded.

The game begins in the "open" state. Any time during the "open" phase there is a positive point scoring action (kill, block, or ace), the responsible point scoring player, or players, becomes "hot."

The "hot" player(s) must then score a second point in a row in order to earn a BIG POINT (one that is recorded on the scoreboard). Any subsequent point would also result in a BIG POINT for her team. If two or more players were involved in the point scoring action, then they are all "hot." Once any one of these players scores the next point, they become the only "hot" player and they may continue to try to score points for her team.

If the hot hitter's opponent makes an error (attack error, block error, receiving error, ball handling error) the hot hitter remains "hot." If the hot hitter's team makes any type of error (the previously mentioned plus service error) the game returns to the "open" phase. If the hot hitter's opponent scores a point (kill, block, or ace) then that player or players become hot.

v. "5-2-15" (Hitting Efficiency Focus): The numbers in this drill correspond to the number of kills, 5, the number of attack errors, 2, out of 15 attempts that a team must earn to win a point. These numbers are one combination of our winning Kill%, 33%, Error%, 14%, and Hitting Efficiency, .172. Obviously if a team meets the standard they will hit .200 and so we are training at a higher level than we need to execute in a match. Here is how we run the drill. Team A is trying to earn the "5-2-15". Team A receives serve from Team B and attempts to side-out. The result of their attack is stated on the scoreboard. At the termination of the rally they return to the same rotation and try to side-out again until they have a minimum of 15 attack attempts. If at any Team A exceeds 2 hitting errors Team B wins. If Team A earns 5 kills on the first 5 swings we continue playing to see if they can manage their offense.