CONSISTENCY IN SPORT - WITH PARTICULAR REFERENCE TO CRICKET

Stephen R. Clarke
Mathematics Dept., Swinburne Institute of Technology, Melbourne, Australia

Keywords: Consistency, sport, cricket, golf

Abstract

Sportsmen strive for consistency in the belief this will improve performance. On the contrary, this paper shows that highly variable scores or results will often produce more wins than consistent ones, and that in some sports a more variable performance will produce a higher average score. However it is also shown that inconsistent results will often arise from consistent behaviour. While initially examples from golf and athletics are used, the measurement of consistency in cricket is explored in more depth, with some suggested additional measures of performance evaluated for some players.

INTRODUCTION

Many manufacturing processes strive for reduced variability of measured attributes to keep the percentage of items outside a given tolerance to a minimum. The graph below illustrates the problem.

As the variability increases, the percentage of items outside the given tolerance (defectives) increases. Thus consistency of manufacturing process, or low variability (low standard deviation) is desirable as it reduces the number of defective items.

DISPELLING SOME MYTHS ON CONSISTENCY IN SPORT

Athletes and sports followers also believe an essential characteristic of excellence is consistency, or low variability of performance. For example Pollock (1977) in discussing the mean and standard deviation of scores in relation to consistency of golfers says 'it seems reasonable that a better player would have low values for both'. However, in many sports, it is the exceptional performance, not the average one that is sought after. Thus an increase in variability will often increase the percentage of wins for players competing simultaneously against many opponents. For example, Clarke (1991) shows that a golfer averaging 1 under par will increase the number of
tournament wins from 3.3% to 8.5% if the standard deviation increases from 1.5 shots per round to 2 shots per round (assuming 10 under is required to win).

Furthermore, in many sports, an increase in variability will actually result in an increase in the average measure of performance. For example, in the long jump the athlete's score is the longest out of 3 or even 6 attempts. In such cases, more variable jumping will not only increase the chance of winning the event, but will actually produce a greater mean score. Since the average maximum of 3 standard normals is 0.8463 the expected score for an athlete whose individual attempts have a mean $\mu$ and standard deviation $\sigma$ is $\mu + 0.8463\sigma$. Clarke (1991) applies this to some actual data for long jumpers, which shows that over 18 cm of the final length of a jumper arises through the variation in the jumps.

It is difficult to find analogous situations in production. One possibility may be in reliability, where 1 or more redundant components are installed. The lifetime of the combined unit is the maximum lifetime of the individual component. In this case, paradoxically, larger variability of the individual components would result in longer average lifetimes of the composite unit.

**Consistent Inconsistency**

In judging the consistency or otherwise of a sportsman, the performance of the sportsman is looked at, and consistency judged by the range of performance - the smaller the range the more consistent. Players with a large range of scores are judged to be inconsistent, and in need of some help to improve their consistency. In fact the reverse is often true - a small range in scores can be indicative of inconsistent behaviour, while a large range may be generated from consistent application.

Clarke (1991) shows that in sports like rifle shooting, archery, golf, where a player has a series of attempts which can be classed as a success (bullseye, par, birdie etc) or a failure, the variance of the number of successes is a maximum when the probabilities of a success at each attempt are equal. Thus the maximum variation in total scores occurs when all attempts have the same chance of success - ie when the player plays each attempt the same way. Thus in golf, a player whose concentration wavers, or who tries harder some holes than others, will produce scores that appear more consistent than one who plays each hole the same.

**CRICKET**

Various attempts have been made to define consistency in cricket and fit distributions to batsmen's scores. Elderton (1945) and Wood (1945) argue that a consistent batsman's scores should follow a geometric distribution. Consider J.D Siddons, who batted about No6 for Victoria (Australia) in 1985/6. His scores for the year were

$$33, 17, 76, 5, 74, 7, 7, 107, 1, 45, 17, 2, 36.$$  

While that looks a fairly inconsistent set of scores, it fits a geometric distribution with mean 33 almost exactly. Cricket followers often expect a consistent batsman to have scores with a small standard deviation, like 51, 55, 52, 53, 54. However scores like this mean that a batsman has no chance of going out until he reaches 50, and is almost certain to go out soon after. So in terms of probability of dismissal he is very inconsistent.

In some companies, the number of days between accidents is often charted. Similar arguments could be put forward with regard to this variable. Thus results more consistent than that predicted by the geometric distribution could signify increased vigilance immediately after an accident, and relaxation of standards after a reasonable non-accident period.
Distribution of the number of balls faced.

Wood (1945) fitted the geometric distributions to cricket scores of many county cricket players. While a reasonably close fit was obtained, he found there were more 0-4's and more large scores than predicted. There are several reasons why scores should not be exactly geometric. The score does not increment by 1, but advances by jumps of usually 1, 2, 3, 4, or 6. Furthermore, there is a strong belief that the probability of advancement is smaller early in the innings (until the batsman settles in) and possibly later when fatigue sets in. One would certainly expect it to change from innings to innings, depending on the quality of the pitch, or the opposition. However many of these arguments do not apply to the number of balls a batsman faces. This certainly increments by 1. Because a batsman can alter the degree of risk he takes, it may be that he plays each ball with the same chance of dismissal. Thus early in the innings he is just content to survive, whereas later when he is settled he will take risks in order to score. Similarly on a bad pitch or against good opposition he may play more carefully, and adjust his scoring rate to keep roughly the same risk of dismissal. Thus the problems of combining two geometric distributions with different means would not apply. For these reasons I would expect the distribution of balls faced to follow the geometric more closely than the actual scores.

Unfortunately, the number of balls faced is not regularly published or even kept by scorers. However, a program written by Mark Johnston (Swinburne MSc Student) allows the reconstruction of ball by ball data from a normal score sheet. The batting data from all Australian matches in the 1989-90 Benson & Hedges cup between Australia, Pakistan and Sri Lanka was analysed. All 60 innings of the top Australian batsmen were combined, and the distribution of both the number of balls faced and scores were compared to the geometric distribution using the Kolmogorov test. While scores fitted slightly better than balls faced, both had p values greater than 0.2. Strangely, the results tended to be the reverse of those found by Wood, in that in both cases the number of very low values and very high values was less than expected. This is presumably due to the limited nature of one day cricket, which often results in the premature ending of an innings. It is hoped to repeat the analysis for test cricket.

Coefficient of Variation as a measure of Consistency

Because the mean and standard deviation of the Geometric distribution are (roughly) equal, Wood suggested using the coefficient of variation (CV) as a measure of consistency. The closer to 100, the more consistent a batsman. Pollard (1977) claims a high CV indicates a batsman has problems early, but scores runs more easily later in the innings. In his analysis, Wood found CV's as low as 96, and up to 139, but mainly in the range 100-109. However, the following model shows that perfectly consistent batsmen will have CV's greater than 100. It also shows that perfectly consistent batsman, but with a different scoring profile, will have different coefficients of variation. Thus it is not possible to have a single measure (closeness to 100) which indicates perfect consistency for all batsmen as proposed by Wood.

The total score S a batsman makes in an innings is the sum of the scores he makes off each ball. If $X_i$ is the score off the ith ball then

$$S = X_1 + X_2 + \ldots + X_N,$$

where $N$ is the number of balls in the innings.

For a perfectly consistent batsman the $X_i$'s should be independent and identically distributed.

Feller (p301) gives the mean and variance of the sum of a random number of random variables. Applying these formulae, and using $p$ as $\sigma/\mu$ we get

$$E(S) = E(X) \cdot E(N) \tag{1}$$
$$\text{Var}(S) = E(X)^2 \cdot \text{Var}(N) + E(N) \cdot \text{Var}(X) \tag{2}$$

Thus $p(S)^2 = \text{Var}(S) / E(S)^2$
\[
\rho(S) = \rho(N) \cdot \sqrt[1 + \rho(X)^2 / (E(N) \cdot \rho(N)^2)]{1 + \rho(X)^2 / [E(N) \cdot \rho(N)^2]} \\
\text{i.e.} \quad \text{CV}(S) = \text{CV}(N) \cdot \sqrt[1 + \text{CV}(X)^2 / (E(N) \cdot \text{CV}(N)^2)]{1 + \text{CV}(X)^2 / [E(N) \cdot \text{CV}(N)^2]} 
\] (3)

Thus we see that the coefficient of variation of the batsman’s score is at least that of the coefficient of variation of the number of balls faced.

We have already seen that the distribution of the number of balls faced for a consistent batsman is a (truncated) geometric distribution.

\[
\text{Pr} ( N = n ) = (1-p)^{n-1}p \quad n = 1, 2, 3, ... 
\]

with mean \( E(N) = 1/p \), variance \( \text{Var}(N) = (1-p)/p^2 \) and coefficient of variation \( \text{CV}(N) = 100\sqrt{1-p} \), where \( p \) is the probability of dismissal each ball. Using equation (1), we see that a batsman who averages 30 \( (E(S) = 30) \), at an average rate of a run every 2 balls \( (E(X) = .5) \), will face on average 60 balls \( (p = 1/60) \). A weak batsman averaging 5 with a run every 3rd ball has \( p = 1/15 \). Thus in most cases \( p \) is small, and using approximations and ignoring terms in \( p^2 \), we get, substituting in (3)

\[
\rho(S) = (1 - 0.5p)^0.5 [ 1 + p \cdot \rho(X)^2 (1 - p)^{-1} ]^{0.5} \\
= (1 - 0.5p) (1 + 0.5p \cdot \rho(X)^2) \\
= 1 + 0.5p \cdot (\rho(X)^2 - 1) \\
\text{i.e.} \quad \text{CV}(S) = 100 + 50p \cdot (\rho(X)^2 - 1) 
\] (5)

As an example of the accuracy of these approximations, typical figures of \( p = 1/50, \rho(X) = \sqrt{2} \) give \( \text{CV}(S) = 100.995 \) in (4) and 101.000 in (5). Figures of \( p = 1/20, \rho(X) = \sqrt{3} \) give \( \text{CV}(S) = 104.8 \) in (4) and 105 in (5).

Equation (5) can be expressed in a slightly different form. If we use \( R = E(X) \) for the rate at which the batsman scores runs (usually called run rate), \( m = E(S) \) for his average score, then equation (1) gives \( p = 1/E(N) = E(X)/E(S) = R/m \)

Thus \[\text{CV}(S) = 100 + 50(R/m)(\text{CV}(X)^2 - 1) \] (6)

This gives the coefficient of variation of scores in terms of the 3 parameters that describe a batsman’s scoring profile. \( m \) is his average score and describes how many runs he gets, \( R \) is the rate and describes how fast he gets them, and \( \text{CV}(X) \) describes how the runs are distributed between singles, fours etc.

These statistics have been calculated for most batsmen (and the top batsmen combined) in the above series and are shown in Table 1. It is interesting to speculate if perceived differences in player characteristics show up. eg, are players exciting because they have a big \( R \) or a big \( \text{CV}(X) \). Many players considered slow, may just have a small \( \text{CV}(X) \). This can also be looked at for an individual innings. Preliminary results suggest that the standard deviation of \( X \) may be a better indicator of 'excitement' than \( \text{CV}(X) \).
Table 1: Calculated statistics for Australian batsmen:

<table>
<thead>
<tr>
<th></th>
<th>R= E(X)</th>
<th>Std(X)</th>
<th>CV(X)</th>
<th>E(N)</th>
<th>CV(N)</th>
<th>m=E(S)</th>
<th>CV(S)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A Border</td>
<td>0.74</td>
<td>1.04</td>
<td>140.2</td>
<td>40.4</td>
<td>63.8</td>
<td>30.1</td>
<td>59.8</td>
</tr>
<tr>
<td>D Boon</td>
<td>0.66</td>
<td>1.03</td>
<td>155.9</td>
<td>50.0</td>
<td>53.7</td>
<td>33.0</td>
<td>59.7</td>
</tr>
<tr>
<td>D Jones</td>
<td>0.72</td>
<td>1.00</td>
<td>139.7</td>
<td>71.2</td>
<td>61.2</td>
<td>51.2</td>
<td>61.0</td>
</tr>
<tr>
<td>G Marsh</td>
<td>0.57</td>
<td>0.97</td>
<td>170.2</td>
<td>57.0</td>
<td>87.7</td>
<td>32.5</td>
<td>108.2</td>
</tr>
<tr>
<td>I Healy</td>
<td>0.99</td>
<td>0.87</td>
<td>87.7</td>
<td>18.0</td>
<td>48.9</td>
<td>17.8</td>
<td>56.1</td>
</tr>
<tr>
<td>M Taylor</td>
<td>0.60</td>
<td>1.06</td>
<td>174.5</td>
<td>54.0</td>
<td>56.9</td>
<td>32.7</td>
<td>71.3</td>
</tr>
<tr>
<td>M. Waugh</td>
<td>1.00</td>
<td>1.24</td>
<td>124.0</td>
<td>14.0</td>
<td>-</td>
<td>14.0</td>
<td>-</td>
</tr>
<tr>
<td>P Taylor</td>
<td>0.77</td>
<td>0.81</td>
<td>106.2</td>
<td>11.8</td>
<td>97.4</td>
<td>9.0</td>
<td>77.5</td>
</tr>
<tr>
<td>S. O'Donnel</td>
<td>0.81</td>
<td>1.11</td>
<td>137.3</td>
<td>34.8</td>
<td>78.8</td>
<td>28.2</td>
<td>75.9</td>
</tr>
<tr>
<td>S Waugh</td>
<td>0.53</td>
<td>0.95</td>
<td>180.7</td>
<td>24.8</td>
<td>89.8</td>
<td>13.0</td>
<td>86.5</td>
</tr>
<tr>
<td>T Moody</td>
<td>0.74</td>
<td>1.19</td>
<td>159.2</td>
<td>49.0</td>
<td>82.9</td>
<td>36.5</td>
<td>105.3</td>
</tr>
<tr>
<td>Combined</td>
<td>0.69</td>
<td>1.04</td>
<td>150.9</td>
<td>44.6</td>
<td>76.8</td>
<td>30.8</td>
<td>82.2</td>
</tr>
</tbody>
</table>

Note that the only batsmen with CV's of scores over 100 are the openers Marsh and Moody. Clearly it is only the openers who regularly experience conditions similar to first class conditions and have time to develop an innings to its full potential. The CV's of balls faced are all much less than 100, suggesting that players are drastically altering their chance of dismissal through the innings - from cautious early to risky when established. On average, the CV's of the scores are greater than that of the number of balls faced, as suggested by equation 4. However using equation 4 to calculate CV(S) generally underestimates its value. This suggests some extra variation possibly due to the non independence of N and X. Similarly, with the exception of Marsh and Moody, all batsmen have CV's much less than calculated by equation 5 or 6 (perhaps an alternative measure of consistency may be the ratio of the actual CV(S) to that calculated by these equations). This again possibly indicates that the limitations of one day cricket do not allow batsmen to be consistent.

Limits to coefficient of variation.

We will now look at a couple of extreme cases to place limits on values for CV(X) and hence CV(S) for consistent batsmen.

Consider a batsman who averages R runs/ball and hits only 1's and 0's. (a dour batsman)

Then  
\[ X_i = 1 \] with probability R  
\[ X_i = 0 \] with probability 1-R

So  
\[ E(X) = R, \ Var(X) = R - R^2, \text{ and } \rho(X)^2 = (1/R - 1) \]

Substituting in (5) gives \( CV(S) = 100 + 50(1-2R)/m \).

Now for a dour batsman, the greatest value of R would be about 0.5, giving \( CV(S) = 100 \) independent of average. A run rate of \( R = 1/4 \) give \( CV(S) = 100 + 25/m \), or about 101 for a batsman who averages 25.

Now consider a batsman who averages R runs/ball and hits only 4's and 0's. (a Botham)

Then  
\[ X_i = 4 \] with probability R/4  
\[ X_i = 0 \] with probability 1-R/4
So \( E(X) = R, \) \( \text{Var}(X) = 4R - R^2, \) and \( \rho(X)^2 = (4/R - 1) \)

Substituting in (5) gives \( CV(S) = 100 + 50(4-2R)/m. \)

Now for an attacking batsman, the greatest value of \( R \) is about 1, giving \( CV(S) = 100 + 100/m, \) or about 105 for a batsman who averages 20, or 102 for one who averages 50. A run rate of \( R = 0.5 \) would give \( CV(S) = 100 + 150/m, \) or about 108 and 103 for the same two batsmen.

It is clear that consistent batsmen may have coefficient of variations of scores up to about 110. Many of the batsmen evaluated by Wood were in this range, and so may not have been showing inconsistency to the extent indicated by the geometric distribution model. It is also not possible to specify a single value for coefficient of variation of scores that signifies consistency for all batsmen.

CONCLUSION

Some of the beliefs about the value and measurement of consistency in sport need to be carefully examined. Consistent scores will not necessarily maximise number of wins. Consistent behaviour will not necessarily produce consistent scores. In cricket, the notion advanced by Wood that a consistent batsman should have a coefficient of variation of scores of 100 does not apply to the alternative model proposed. This model also demonstrates that a single measure of consistency of scores applicable to all batsmen may not be attainable. For a range of different styles of batsman, all consistent in their own batting performance, the coefficient of variation should lie between 100 and 110. Hence an alternative indication of consistency is a standard deviation at the least equal to the mean, not one equal to zero. Batsmen whose standard deviations are significantly less than the mean are possibly concentrating better early in their innings, or throwing away hands once established. This appears to be borne out by one day results, which show coefficients of variation much lower than 100, particularly for non-opening batsmen.

REFERENCES.


