Captains should not employ nightwatchmen

Timothy Trudgian[*]
Mathematical Sciences Institute
The Australian National University, ACT 0200, Australia
timothy.trudgian@anu.edu.au

December 16, 2014

1 Introduction

Imagine a game of cricket in which one team, currently on 1 for 50, loses a wicket with 4 overs left in the day. Normally one would expect the fourth batsman, typically one of the team's strongest, to appear. At times teams may adopt an alternative approach and promote a *nightwatchman.*

Unlike the pinch hitter in baseball — who is promoted to increase the score — the nightwatchman is introduced to minimise the damage. His job, as the name suggests, is to watch over the team as night falls: he is sent in to stop further wickets from being lost.

The nightwatchman is almost always a 'poor, but not too poor' batsman. The rationale for such a selection is as follows. It would be disastrous for the fourth batsman to be dismissed that evening. Since he cannot score any substantial runs it is better to protect him for the next day. The nightwatchman should put up a solid wall of defence in the hope that he survives the evening (hence he must not be too poor); should he be dismissed early on the next day the team has not lost much (hence he must not be too good). The batting captain is faced with the following

Question. *To play, or not to play, a nightwatchman?*

The argument in favour of a nightwatchman is that the benefit of preserving the next-in batsmen outweighs the small loss in losing the nightwatchman the next morning. The argument against centres on the nightwatchman's lack of batting ability: if he is a 'poor' batsman, surely he has a high chance of being dismissed, in which case the batting captain is back to where he started, less one wicket.[1]

There appears to be little mathematical evidence in the literature that could support either argument, though there is some empirical evidence [1]. Practice has varied amongst captains: Steve Waugh never employed a nightwatchman; Michael Clarke used Nathan Lyon as a nightwatchman as recently as February 2014.

This article seeks to offer a mathematical answer in the form of

Theorem 1. *One should never employ a nightwatchman.*

---

[1]Invariably, if the nightwatchman *is* dismissed that evening, the batting captain sends in the next batsmen. There does not appear to be an instance where *two* night-watchmen have been used in the same evening.

This answer is, of course, constrained by some assumptions.

Assumptions. *The following are assumed throughout this article.*

1. *The probability of a batsman being dismissed remains constant throughout the innings.*

2. *Batsmen, when batting in partnerships, face an equal number of deliveries.*

These assumptions are obviously open to debate. One may expect the probability of dismissal to be higher at the start of a batsman's innings. Once he 'plays himself in' and becomes accustomed to the conditions the chance of dismissal should diminish. As for Assumption 2: one batsman could well 'farm' the strike and face far more deliveries than his partner.

The assumptions could be adapted easily enough, though this would make for a more cumbersome exposition. Indeed, one could assemble statistics on the number of times a batsmen has been out after facing 1 ball, 2 balls, etc. This is certainly possible, but requires the accumulation of much more data. Our assumptions enable us to use only two items of information per batsmen, or 22 pieces of data in total, to compile an estimation.

2 Expected number of runs

We first need to consider the number of runs that a batsman, or group of batsmen, is expected to make. Consider a team $\mathcal{T}$ composed of eleven batsmen $\mathcal{B}_1, \mathcal{B}_2, \ldots, \mathcal{B}_{11}$ batting in that order. How many runs do we expect this team to score?[2]

Let A_i denote the batting average of the ith batsman. At first blush one might assume that 'on average each batsman will bat his average', and thus the team $\mathcal{T}$ ought to be expected to score $\sum_{i=1}^{11} A_i$ runs.

This is unsatisfactory as it does not take into account the order in which the batsmen appear. Under this naïve model the expected runs of $\mathcal{T}$ is invariant under any permutation of the $\mathcal{B}_i$'s. That the batting order ought to influence the expected run total is evident upon considering the following example.

2.1 The importance of batting order

Suppose, instead of ten wickets a side, a truncated version of the game had only two wickets a side; that is, a side has only three batsmen $\mathcal{B}_1$, $\mathcal{B}_2$ and $\mathcal{B}_3$. Suppose further that $\mathcal{B}_1$ and $\mathcal{B}_2$ are 'good' batsmen (as indicated by, say, their averages) and that $\mathcal{B}_3$ is a 'bad' batsman.

The natural batting order would be

$$\{\mathcal{B}_1, \mathcal{B}_2, \mathcal{B}_3\}, \tag{2.1}$$

which, of course, is invariant under the permutation $\mathcal{B}_1 \leftrightarrow \mathcal{B}_2$. Now consider an alternative batting order

$$\{\mathcal{B}_3, \mathcal{B}_1, \mathcal{B}_2\}. \tag{2.2}$$

Under our 'each bats his average' mantra, the batting orders in (2.1) and (2.2) ought to produce the same number of runs. Suppose we consider the *partnerships* in each of the orders in (2.1) and

---

[2]The answer to this question may depend on what we expect 'expect' to mean.

(2.2). Define a *good partnership* as one between two good batsmen and a *mixed partnership* as one between a good batsman and a bad batsmen. One expects more runs to come from a good partnership than from a mixed partnership.

In (2.1), regardless of the order of dismissal, there is one good partnership (viz. between $\mathcal{B}_1$ and $\mathcal{B}_2$) and one mixed partnership (viz. between $\mathcal{B}_3$ and whichever of the openers is not dismissed). Given that $\mathcal{B}_3$ is a bad batsmen it follows that the chances of his being dismissed are higher than those of his good colleagues. Therefore, in (2.2) it is likely that there will be one mixed and one good partnership — the same result as in order (2.1). However, in the event of $\mathcal{B}_1$ being dismissed first in (2.2), we have two mixed partnerships, whence fewer runs scored than in batting order (2.1).

It is apparent that built into the averages of the players is their being able to face sufficiently many balls to score sufficiently many runs. If a good batsman $\mathcal{B}_1$ comes in with only one wicket in hand he is less likely to produce a large score: once his partner is dismissed the innings is over. Taken to its extreme, this is the prevailing reason why one puts the 'best batsmen' at the top of the innings and 'worst batsmen' at the bottom.

Rather than consider the averages of the players, we consider their likelihood of dismissal and the rate at which they accumulate their runs. Consider, therefore, the following quantities. Let B_i, and S_i denote respectively the *average occupancy* and the *strike rate* of $\mathcal{B}_i$. We define the average occupancy to be the average number of balls faced per dismissal; the strike rate is the average number of runs scored per hundred balls faced. The likelihood of dismissal of $\mathcal{B}_i$ is defined to be $\frac{1}{B_i}$, that is, the reciprocal of his average occupancy.

These are the only statistics that enter our model. This is useful since both of these statistics are freely available at [2].[3] We move our attention to the expected number of runs in a single partnership.

2.2 Two batsmen

Consider a partnership between $\mathcal{B}_1$ and $\mathcal{B}_2$. How many runs do we expect from this partnership? For $n \geq 0$ let $X_{1,2}(n)$ denote the probability that the partnership consisting of $\mathcal{B}_1$ and $\mathcal{B}_2$ survives for n balls before it is broken. By Assumption 2, $\mathcal{B}_1$ and $\mathcal{B}_2$ each face $n/2$ balls.[4] Therefore, after these n balls we expect $\mathcal{B}_1$ to be on strike half of the time and $\mathcal{B}_2$ to be on strike half of the time.

By Assumption 1 the probability that $\mathcal{B}_i$ survives $k \geq 0$ balls is $(1 - 1/B_i)^k$, whence

$$X_{1,2}(n) = \frac{1}{2} \left(1 - \frac{1}{B_1}\right)^{n/2} \left(1 - \frac{1}{B_2}\right)^{n/2} \left(\frac{1}{B_1} + \frac{1}{B_2}\right). \tag{2.3}$$

We now turn to the runs expected when the above event occurs: call this $R_{1,2}(n)$. Given that we have assumed that batsmen score runs in proportion to their strike rates we have

$$R_{1,2}(n) = \frac{n}{2} \left(\frac{S_1}{100} + \frac{S_2}{100}\right). \tag{2.4}$$

---

[3]We note the following problem in historical comparison. Until the early 1990s it was not customary to record the number of balls faced by a batsman. Thus, for example, no data appear to be available on the number of balls faced by Sachin Tendulkar.

[4]Allowing 'fractional balls' faced by each batsman is not really an additional assumption since we are already considering averages per hundred balls, average occupancy, etc.

It follows that the expected number of partnership runs, denoted $P_{1,2}$ is the sum over all possible values of n of the product $X_{1,2}(n)R_{1,2}(n)$. We therefore have

$$P_{1,2} = \frac{1}{400}(S_1 + S_2)\left(\frac{1}{B_1} + \frac{1}{B_2}\right)\sum_{n=0}^{\infty} nY_{i,j}^n, \tag{2.5}$$

where $Y_{i,j} = \sqrt{(1 - 1/B_i)(1 - 1/B_j)}$. Using the series[5] $\sum_{n=0}^{\infty} nx^n = \frac{x}{(1-x)^2}$, valid for $|x| < 1$, we are able to prove

Lemma 1. *Let $P_{1,2}$ denote the number of runs scored by batsmen $\mathcal{B}_1$ and $\mathcal{B}_2$, and let $Y_{i,j} = \sqrt{(1 - 1/B_i)(1 - 1/B_j)}$. We have*

$$P_{1,2} = \frac{1}{400}(S_1 + S_2)\left(\frac{1}{B_1} + \frac{1}{B_2}\right)\frac{Y_{1,2}}{(1 - Y_{1,2})^2}. \tag{2.6}$$

2.3   Three batsmen

Now consider the expected number of runs of three batsmen: $\mathcal{B}_1, \mathcal{B}_2$ and $\mathcal{B}_3$. Let $\mathcal{B}_1$ and $\mathcal{B}_2$ bat for n_1 balls. Suppose $\mathcal{B}_1$ is dismissed and $\mathcal{B}_2$ and $\mathcal{B}_3$ bat for n_2 balls. We find that

$$\sum_{n_1=0}^{\infty}\sum_{n_2=0}^{\infty} Y_{1,2}^{n_1}Y_{2,3}^{n_2}\frac{1}{4B_1}\left(\frac{1}{B_2} + \frac{1}{B_3}\right)\left\{\frac{n_1}{400}(S_1 + S_2) + \frac{n_2}{400}(S_2 + S_3)\right\}. \tag{2.7}$$

The case when $\mathcal{B}_2$ is dismissed and $\mathcal{B}_1$ and $\mathcal{B}_3$ bat for n_2 balls will produce (2.7) with $\mathcal{B}_1 \leftrightarrow \mathcal{B}_2$. To evaluate (2.7) we examine

$$\begin{aligned}
\sum_{n_1=0}^{\infty}\sum_{n_2=0}^{\infty} x_1^{n_1}x_2^{n_2}(a_1x_1 + a_2x_2) &= \sum_{n_1=0}^{\infty} x_1^{n_1}\left(a_1\frac{x_1}{1 - x_2} + a_2\frac{x_2}{(1 - x_2)^2}\right) \\
&= a_1\frac{x_1}{(1 - x_1)^2(1 - x_2)} + a_2\frac{x_2}{(1 - x_1)(1 - x_2)^2}.
\end{aligned} \tag{2.8}$$

This leads us to

Lemma 2. *Let $P_{1,2,3}$ denote the number of runs scored by batsmen $\mathcal{B}_1, \mathcal{B}_2$ and $\mathcal{B}_3$, and let $Y_{i,j} = \sqrt{(1 - 1/B_i)(1 - 1/B_j)}$. We have*

$$\begin{aligned}
P_{1,2,3} = \frac{1}{1600}&\left\{\frac{1}{B_1}\left(\frac{1}{B_2} + \frac{1}{B_3}\right)\left[\frac{Y_{1,2}(S_1 + S_2)}{(1 - Y_{1,2})^2(1 - Y_{2,3})} + \frac{Y_{2,3}(S_2 + S_3)}{(1 - Y_{1,2})(1 - Y_{2,3})^2}\right]\right. \\
&\left.+ \frac{1}{B_1}\left(\frac{1}{B_2} + \frac{1}{B_3}\right)\left[\frac{Y_{1,2}(S_1 + S_2)}{(1 - Y_{1,2})^2(1 - Y_{2,3})} + \frac{Y_{2,3}(S_2 + S_3)}{(1 - Y_{1,2})(1 - Y_{2,3})^2}\right]\right\}.
\end{aligned} \tag{2.9}$$

---

[5]One may object to the oversimplification of allowing batsmen to face infinitely many balls. Since the series is convergent the difference between the infinite sum and, say, the number of runs scored after 900 balls, is negligible.

2.4 Generalisation to n batsmen for $2 \leq n \leq 11$

To continue the investigation into further partnerships we first need to prove

Lemma 3. *For any $j \geq 2$ we have*

$$\sum_{n_1=0}^{\infty} \sum_{n_2=0}^{\infty} \cdots \sum_{n_j=0}^{\infty} x_1^{n_1} x_2^{n_2} \cdots x_j^{n_j} (a_1 x_1 + a_2 x_2 + \cdots a_j x_j) =$$

$$\frac{a_1 x_1}{(1-x_1)^2(1-x_2)\cdots(1-x_j)} + \frac{a_2 x_2}{(1-x_1)(1-x_2)^2\cdots(1-x_j)} + \cdots \qquad (2.10)$$

$$+ \frac{a_j x_j}{(1-x_1)(1-x_2)\cdots(1-x_j)^2}.$$

Proof. The case $j = 2$ was proved in (2.8); (2.10) follows by induction. $\qquad \square$

Let $\mathcal{W}$ denote the order in which all ten wickets fall; that is, $\mathcal{W} = (i_1, i_2, \ldots, i_{10})$ where $1 \leq i_1 < i_2 < \ldots < i_{10} \leq 11$. For a given $\mathcal{W}$ we first deduce the partnerships involved in the innings. For example if $\mathcal{W} = (1, 3, 4, 5, 2, 6, 8, 9, 10, 7)$ we know that the partnerships must have been

$$(\mathcal{B}_1 \,\&\, \mathcal{B}_2), \quad (\mathcal{B}_2 \,\&\, \mathcal{B}_3), \quad (\mathcal{B}_2 \,\&\, \mathcal{B}_4), \quad (\mathcal{B}_2 \,\&\, \mathcal{B}_5), \quad (\mathcal{B}_2 \,\&\, \mathcal{B}_6),$$
$$(\mathcal{B}_6 \,\&\, \mathcal{B}_7), \quad (\mathcal{B}_7 \,\&\, \mathcal{B}_8), \quad (\mathcal{B}_7 \,\&\, \mathcal{B}_9), \quad (\mathcal{B}_7 \,\&\, \mathcal{B}_{10}), \quad (\mathcal{B}_7 \,\&\, \mathcal{B}_{11}).$$

We can apply Lemma 3 with $x_i = Y_{j,k}$ and $a_i = (S_j + S_k)/200$, where $\mathcal{B}_j$ and $\mathcal{B}_k$ are the batsmen in the ith partnership. We then sum over all possible $\mathcal{W}$ to calculate the expected number of runs.

This may be of some independent interest in predicting the combined score of several partnerships. It would take considerable effort to run a statistical analysis on this model. One would have to recalculate batsmen's strike rates and average occupancies before each innings.

2.5 Current examples

Take the following Australian team as at 27 March 2014. The expected number of runs is 325.15.

Table 1: Australian statistics

$\mathcal{B}_i$	B_i	S_i
DA Warner	63.45	73.35
CJL Rogers	80.40	47.44
SE Marsh	73.93	44.45
MJ Clarke	92.24	55.83
SPD Smith	77.44	51.69
SR Watson	68.23	53.13
BJ Haddin	60.39	58.39
MG Johnson	37.94	58.29
RJ Harris	31.60	61.13
PM Siddle	30.49	46.72
NM Lyon	41.05	39.34

Were one to bring in, Siddle, say, to bat ahead of Watson as a nightwatchman, the expected number of runs becomes 321.66.

3 Nightwatchman and Conclusion

We are now in a position to comment on the efficacy of a nightwatchman. The effect of promoting a nightwatchman, say $\mathcal{B}_{10}$, over the next-in batsman, say $\mathcal{B}_4$ is to push every batsman $\mathcal{B}_i$ with $4 \leq i \leq 9$ down one position in the batting order. This means that good batsmen appear fewer times in the sum over all orders $\mathcal{W}$. Therefore, presuming that the nightwatchman is a worse batsman than the one he is replacing, it follows that it is poor practice to promote a nightwatchman. This establishes Theorem 1.

The decision to employ the nightwatchman is a gamble; it is always worthwhile to know the odds before gambling. Of course there are exceptional cases of success: Jason Gillespie scored a double century as a nightwatchman. But, all things considered, employing a nightwatchman is a bad bet.

References

[1] Davis, C., *The myth of the nightwatchman*, `http://www.sportstats.com.au/nightwatchman.html` **accessed** 24 March 2014.

[2] *ESPNcricinfo*, `www.espncricinfo.com` **accesssed** 24 March 2014.