A STOCHASTIC ANALYSIS

OF

SCORING SYSTEMS

by

Graham Hilford Pollard

A thesis submitted for the
degree of Doctor of Philosophy
of The Australian National University

December, 1986

The work described in this thesis is my own, with the exception of the coding of the simulations involved in Sections 8.9, 9.4 and 9.5 which were carried out by Dr Ken Noble.

To my parents.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

The work for this thesis was carried out at the Department of Statistics, Institute of Advanced Studies, The Australian National University. I gratefully acknowledge the assistance of this department in providing the facilities for study during the period in which I studied full-time at the department. I also gratefully acknowledge the assistance provided by my employer, The Canberra College of Advanced Education, in allowing me some time on leave to complete part of the research.

I am indebted to Dr Roger Miles for his help and guidance at all stages from the time at which I first expressed an interest in commencing studies.

I would also like to thank Mrs Ann Milligan of Science Text Processors (Canberra) for her help, care and patience during the word-processing phase of the preparation of this thesis.

ABSTRACT

Many scoring systems can be seen as statistical tests of hypotheses. In tennis singles, for example, the scoring system used can be seen as a test involving 2 binomial probabilities p_a and p_b where p_a (p_b) is the probability player A (player B) wins a point initiated by player A (player B). Tennis singles is thus a "bipoints" game. The tennis scoring system is an inefficient test relative to the sequential probability ratio test ($SPRT$) based on pairs of these points. When $p_a + p_b > 1$ (the tennis context), an $SPRT$ based on the "play-the-loser" (PL) rule is super-efficient. Chapter 2 shows that, when $p_a + p_b > 1$, there is in fact a spectrum of super-efficient tests (with even durations) based on "partial-PL" (PPL) rules. The most efficient tests within this spectrum, when $p_a + p_b > 1$, are the $SPRT$ based on the (full) PL rule. Chapter 3 extends this spectrum of tests to produce the total spectrum of tests (including those with odd durations).

Points within the tennis scoring system have different "importances" whereas points within any member of the above (efficient) spectrum of PPL systems are equally "important" when $p_a = p_b$. Intuitively, the differing importances of the points within the tennis scoring system contribute to the inefficiency of that system. Chapter 4 establishes a relationship between the efficiency of a bipoints scoring system and the importances of the points within it; a relationship which is used in Chapter 5 to show that the $SPRT$ based on the PL rule has an optimal efficiency property when $p_a + p_b > 1$.

Chapters 6 and 7 address the question as to whether the super-efficiency of the PL rule carries over to the case of tennis doubles in which there are essentially 4 binomial probabilities p_{a1}, p_{a2}, p_{b1} and p_{b2}. Some asymptotic results are achieved although, generally speaking, they are of little practical relevance.

The particular scoring system used in tennis is analysed in detail in Chapter 8 and the methodology used is seen to be useful for analysing any "nested" scoring system (e.g. tennis is 3-nested: points - games - sets). It was the study of this specific scoring system and its inherent inefficiency which lead to the theory of Chapters 2 to 7. A new tennis scoring system is proposed in Chapter 8.

Chapter 9 contains a brief discussion of some of the characteristics the designer of a scoring system needs to consider and some results are given. The study of the importances of points is extended in Chapter 10 and in Chapter 11 team play with associated countback rules is investigated. The general conclusion is that "upward-nested" countback systems (e.g. points - games - sets, in tennis) are preferable to

"downward-nested" systems (sets - games - points).

In Chapter 12 it is shown that the classical scoring system used in multiple choice examinations can be considerably improved by modifying that scoring system and instructing the examinees to cross any boxes known to be incorrect when the correct box for that question is unknown.

CHAPTER 1
INTRODUCTION

1.1 Background

The properties of different scoring systems have attracted the attention of a number of writers. For example, the scoring systems used in squash have been analysed by ApSimon (1951, 1957), Watson (1970), Hsi and Burych (1971), Kingston (1976), Renick (1976, 1977), Schutz and Kinsey (1977), Phillips (1978), Clarke (1979), Clarke and Norman (1978, 1979) and Pollard (1980, 1983b, 1985a). Some of the above articles also analysed the (related) scoring system used in badminton. Tennis scoring systems and related aspects have been investigated by Kemeny and Snell (1960), Schutz (1970), Gale (1971), Hsi and Burych (1971), Morris (1972, 1973, 1977), George (1973), Carter and Crews (1974), Hannan (1976), Kingston (1976), Anderson (1977), King and Baker (1979), Fischer (1980), Croucher (1981), Pollard (1980, 1983a, 1987) and Miles (1980, 1982, 1983 and 1984). Dibley and Wallis (1981) analysed jai-alai which is a game in which two players (or two teams) compete to win a "match", similar to squash or tennis. However, a tournament structure is imposed on the matches so that more than two players (or two teams) can compete in the tournament even though only two can play at one time in a match. Dibley and Wallis carried out a computer analysis of the starting effect for the case of 8 players (or 8 teams).

The "unfairness" characteristic of some scoring systems (e.g. international squash and badminton) has long been recognised (e.g. ApSimon, 1951). The magnification effect of a scoring system was noted by Kemeny and Snell (1960) and Morris (1977). The concept of the importance of a point was first introduced in the interesting paper by Morris (1977). The concept of the efficiency of a scoring system was formalized in the elegant paper by Miles (1984). The distribution of the number of points played in squash and tennis has been outlined by Phillips (1978) and Pollard (1983a) respectively. These properties or concepts of fairness, magnification, importance, efficiency and distribution of duration (or just variance of duration) have thus been developed in recent times. Modifications to present scoring systems based on such properties or concepts have been proposed by Schutz (1970), Hsi and Burych (1971), Carter and Crews (1974), Miles (1984) and Pollard (1987).

In Chapters 2 to 5 of this thesis a spectrum of efficient scoring systems is devised, a relationship between the efficiency of a scoring system and the importance of the points within it is derived and a scoring system with an optimally efficient property is found. Thus the work by Miles (1984) on efficiency and by Morris (1977) on importances formed the basis or starting point for the early chapters in this thesis. For this reason, these aspects of their work are outlined in the next two sections.

1.2 The Work of Miles (1984)

Miles wrote three earlier papers (1980, 1982, 1983) on scoring systems and on the inefficiency of the tennis scoring method and is currently writing a monograph on scoring systems (Miles, to appear). A brief outline of Miles' 1984 paper analysing the efficiency of scoring systems is now given.

Many sports consist of playing a sequence of points each of which is won either by player A or player B. If there is only one type of point in the match, we have ***unipoints*** in which p (q) is the probability player A wins (loses) each point $(p + q = 1)$. p is a statistical unknown and the aim is efficiently to identify the better player. Player A is the ***better player*** if $p > 0.5$. Considering this sequence of Bernoulli trials, let $X = (X_1, X_2, \ldots)$ be a sequence of encounters between A and B with A $[B]$ representing a point won by player A $[B]$. Let F be a unirule which determines sequentially, for all X, on the basis of $X_1, X_2, \ldots, X_i$ whether A wins or loses or play continues with the sampling of X_{i+1} $(i = 1, 2, \ldots)$. The unirule F is symmetric if its determinations for $\overline{X} = (\overline{X}_1, \overline{X}_2, \ldots)$ where $\overline{X}_i = B$ $[A]$ if $X_i = A$ $[B]$ are the reverse of those of X. This symmetry thus provides the "fairness" structure for the scoring system. Denoting

$$P_F(p) = \text{Prob}\,(A \text{ wins match} \mid F, p)$$

$$\mu_F(p) = E(D \mid F, p)$$

where D is the duration (number of points) of the match, symmetry implies $P_F(0.5) = 0.5$. A symmetric unirule also satisfying the regularity conditions

(i) $P_F(p)$ increases continuously and monotonely from 0.5 to 1 as p increases from 0.5 to 1 and

(ii) $\mu_F(p) < \infty$ $(0 < p < 1)$

is called a ***uniformat***. Thus, a uniformat is just a fair scoring system possessing sensible regularity conditions and involving only one type of point.

Consider, for $0.5 < p' < 1$, the simple hypotheses

$$H_0 : p = p', \;\; q = q'$$

$$H_1 : p = q', \;\; q = p'$$

where $q' = 1 - p'$. Applying Wald and Wolfowitz's (1948) result, Miles concluded that there is a ***unique class of optimal uniformats***, given by the Sequential Probability Ratio Test. It is $\{W_n\}$ $(n = 1, 2, 3, \ldots)$, where W_n is the uniformat in which the winner is the first player to achieve a lead of n points over the other. The key characteristics of W_n are

$$P_{W_n}(p) = p^n/(p^n + q^n)$$

and

$$\mu_{W_n}(p) = n(P_{W_n}(p) - Q_{W_n}(p))/(p - q)$$

where $Q_{W_n}(p) = 1 - P_{W_n}(p)$. Miles then showed that the ***efficiency***, ρ, of a general uniformat with key characteristics P and μ at p is given by

$$\rho = \frac{(P - Q)\ln\frac{P}{Q}}{\mu(p - q)\ln\frac{p}{q}}$$

where $Q = 1 - P$.

Thus Miles (1984) created a formal structure for evaluating the efficiency of any uniformat used to test the above symmetric single statistical hypotheses about the parameter p. He noted that scoring systems such as volleyball, badminton and international squash rackets are not uniformats essentially because of the advantage of serving first; however they may be converted into uniformats by the simple device of playing an initial non-scoring point to determine the server (see also Clarke and Norman (1979)). Miles (1984) also studied the efficiency of a uniformat that forms the basis of many scoring systems: the "first to n points" or "best of $2n - 1$ points" uniformat, B_{2n-1}. He showed that the efficiency of B_{2n-1} at p decreases as n increases to a limiting efficiency of

$$\rho = p\ln(1/4pq)/((p - q)\ln(p/q)) \qquad (\tfrac{1}{2} < p < 1)$$

and thus uniformats B_{2n-1} for large n should be avoided in practice.

Kemeny and Snell (1960, pp.161–167) had earlier noted the "magnification" achieved by scoring systems (thus, for example, "a game of tennis serves to magnify the [point-probability] difference between two players"). Restricting to unipoints, they compared the magnification of the two systems B_{2m-1} and W_n (for (large) values of m and n such that the expected durations were approximately equal) and concluded that B_{2m-1} had about 0.8 times the magnification of W_n when $p =$

$0.5+\epsilon$ ($\epsilon > 0$ small). Thus they concluded that W_n was more efficient than B_{2m-1}. However, they could not directly measure the efficiency, ρ, of B_{2m-1} since they had not set up symmetric hypotheses with α and β equal (α = Prob(Type I error); β = Prob(Type II error)). Schutz, (1970) also attempted to evaluate the efficiency of several unipoints scoring systems and, noting the work of Kemeny and Snell, proposed the use of the W_n system.

We now consider sports such as tennis in which serving is an advantage. There are essentially 2 types of point (a-points (A serves) and b-points (B serves)), and so we have *bipoints* in which p_a (p_b) represents the probability that player A (B) wins an a- (b-) point. Here p_a and p_b are statistical unknowns.

In a study of some tennis data Pollard (1980, pp.71–93) analysed 35 matches (5503 points) from one of Australia's more prestigious tournaments, the 1977 NSW Men's Open Lawn Tennis Championships, and found that the winner of a match won on average 71% of points on service, while the loser won on average 62% of points on service. He also applied numerous statistical tests to test the assumption that the outcome of any point is independent of any other point and concluded that the assumption was reasonable.

The formal structure for evaluating efficiency in this bipoints case, due to Miles (1984), is now given. Player A is the *better player* if $p_a > p_b$. A symmetric *birule*, F, is set up analogously to the symmetric unirule. We also require the initial type of point played to have no bearing on P and μ. That is

$$\begin{aligned} & P^a = P^b = P \\ \text{and} \qquad & \mu^a = \mu^b = \mu \end{aligned}$$

where the superscripts a and b refer to the initial point type. If a symmetric birule F satisfies the above and the condition

$$p_a > p_b \Longrightarrow P > 0.5 \,,$$

we say F is a *biformat*.

By considering (a, b) point-pairs (one a- and one b- point is played (pairwise) with the outcome being win, draw or loss), W_n(point-pairs) (the winner is the first player to achieve a lead of $2n$ points over the other) can be set up as the *standard biformat* and the efficiency of a general biformat F can be measured relative to it. Wald (1947, Chapter 6) recommended the use of point-pairs for comparing the means of two binomial populations. Since we require the initial type of point to have no bearing on P or μ, Miles set up reversal biformats in the following elegant way. Let F^a be a sequential birule starting with an a-point and resulting in a win for either A or B after a point duration of finite mean. Let F^b denote the reverse birule, obtained from F^a by the exchanges $a \leftrightarrow b$, $A \leftrightarrow B$. Then $G = [F^a, F^b]$ denotes a single play which, operating like a point-pair, results in a win by A, a draw or a loss by A. If H is a uniformat then the reversal biformat $H(G)$ is decided by H applied to wins and losses in independent replication of G as "points" (draws are ignored with respect to G but count towards μ). $H(G)$ is a biformat.

For $0 < p_b' < p_a' < 1$, the simple hypotheses

$$H_0 : p_a = p_a',\ p_b = p_b'$$
$$H_1 : p_a = p_b',\ p_b = p_a'$$

are set up and in a similar manner to uniformats, the ***efficiency***, ρ, of a general biformat with key characteristics P and μ at (p_a, p_b), ***relative to*** $\{W_n$***(point-pairs)***$\}$ ***as standard***, is given by

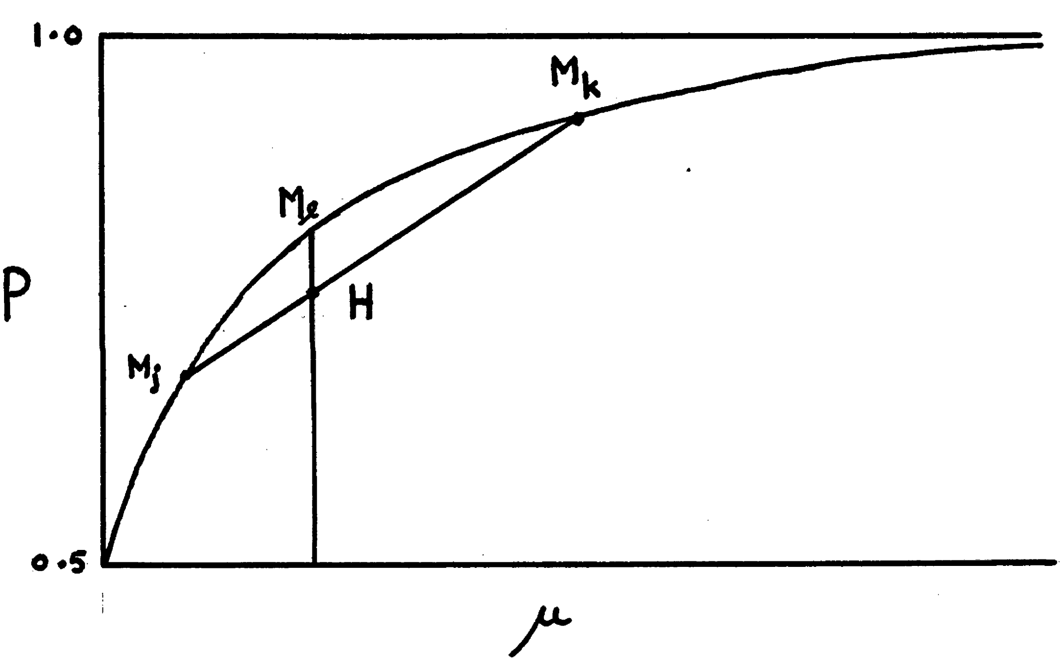
$$\rho = \big(2(P-Q)\ln(P/Q)\big)/\big(\mu(p_a - p_b)\ln(p_a q_b / p_b q_a)\big) \tag{1.2.1}$$

where $q_a = 1 - p_a$ and $q_b = 1 - p_b$. Thus, for given (p_a, p_b), efficiency is a function of P and μ. Using the play-the-loser (PL) discipline in which a win by player A $[B]$ is followed by a b- $[a$-$]$ point, Miles showed that the reversal biformat $\{W_1(W_n(PL^a), W_n(PL^b))\}$ $(n = 1, 2, 3, \ldots)$ was more efficient than $\{W_n(\text{point-pairs})\}$ $(n = 1, 2, 3, \ldots)$ when $p_a + p_b > 1$ (and less efficient when $p_a + p_b < 1$) and $n \geq 2$. He noted that $\{W_1(W_n(PW^a), W_n(PW^b))\}$ $(n = 1, 2, 3, \ldots)$ was more efficient than $\{W_n(\text{point-pairs})\}$ $(n = 1, 2, 3, \ldots)$ when $p_a + p_b < 1$ (and less efficient when $p_a + p_b > 1$) and $n \geq 2$. Here PW^x represents play-the-winner starting with an x-point. (See also Sobel and Weiss (1970) and Zelen (1969)). Miles noted that increases in efficiency are typically minute. He also defined the alternating sequence of a and b points $(a\,b\,a\,b\,a\,b\,\ldots)$, AL, and has shown $\{W_1(W_n(AL^a), W_n(AL^b))\}$ to have the same efficiency as $\{W_n(\text{point-pairs})\}$. To simplify notations, $W_{2n}PL$, $W_{2n}PW$ and $W_{2n}AL$ will be used to represent these three families of reversal biformats respectively.

Miles (1984) noted the inefficiency of the tennis scoring system especially for players with strong services or a strong serving advantage (e.g. men's doubles and matches on "fast" surfaces) and recommended a simple device within the existing framework to remedy this problem:– commence each game not at "love all" but at 0–15, 0–30 or even 0–40 as the "service advantage" (for both players or both pairs of players) increases.

Families of uniformats (or biformats) $\{M_i\}$ $(i = 1, 2, 3, \ldots)$ typically have a (μ, P) relationship at fixed p (or (p_a, p_b)) which is convex upwards as shown in Figure 1.2.1. Consider such a family and suppose a game is played with probability p' under uniformat (or biformat) M_j and, with probability $1 - p'$ under uniformat (or biformat) M_k. Then, it can be seen that, the efficiency of such a hybrid, H, is less than that of the corresponding interpolated point, M_l, on the smooth (μ, P) curve (see Figure 1.2.1).

Figure 1.2.1
A typical (μ, P) relationship for fixed p or (p_a, p_b).



Hsi and Burych (1971) had earlier recognised the bipoints character of tennis and derived the probabilities that a player wins (or loses) a classical set of tennis 6–0, 6–1, 6–2,…,0–6. However, in their conclusions, they made two suggestions (without any mathematical analysis) for shortening the potentially long sets of classical tennis. Firstly, that service points may be alternated rather than service games and secondly, if 5 games all is reached, "a scheme similar to Wald's sequential analysis may be used … the set is determined when one player scores 5 points, say, more than his opponent". However, they did not address the inherent inefficiency of the tennis scoring system.

Carter and Crews (1974) explored the effect of a few different tie-breaker rules in tennis. However, they made the assumption that bipoints tennis could be approximated by unipoints tennis (by using the average $p = (p_a + q_b)/2$); an assumption which was not justified and which led to some incorrect conclusions (e.g. "maximum expected set length of 67.70 points" for classical tennis). They also concluded that "if the game of tennis were thought of as a test of hypothesis, the above information would indicate that it was a powerful test", and that "a tennis match can be considered as a hierarchical series of replicated paired comparisons between the two opponents" and that "a tennis match can be used as an example to illustrate the power of the paired-comparison technique". Thus, without any attempt to state the hypotheses, they recognised the hypothesis testing nature of the tennis scoring system but were incorrect in concluding that scoring in tennis (in its traditional form) was a "powerful test". Pollard (1980, p.10) noted the inefficiency of the tennis scoring method.

1.3 The Work of Morris (1977)

Morris wrote two earlier papers on tennis (1972, 1973), but in his 1977 paper he introduced the elegant concept of the *importance* of a point (for winning a game of tennis) which he defined as "the difference between two conditional probabilities: the probability that the server wins the game given that he wins the point, minus the probability that he wins the game given that he loses that point". He noted that each point is equally important to both players but that some points (games or sets) are more important (to both players) than other points (games or sets). The "time-importance, T_i" of state i within a game was defined as the product of its importance I_i within the game and the expected number of times that state i is entered, n_i, in a single play of the game. Morris noted that if player A increased p to $p+\epsilon$ whenever state i was entered, then the probability of winning the game was increased by $T_i\epsilon$ ($\epsilon > 0$ small). He defined the "magnification factor, M" of a game (at p) as the sum (over all possible non-absorbing states within the game) of the T_i values, and noted the following approximate property

$$P(p+\epsilon) - P(p) \doteq M\epsilon \qquad \text{(small } \epsilon\text{)}$$

where $P(p+\epsilon)$ ($P(p)$) is the probability player A wins the game when the probability of winning every point is $p+\epsilon(p)$. Suppose all possible (non-absorbing) states in a game of tennis are now listed from the most important to the least important (with possible equalities in importances) state. For each non-absorbing state j, the cumulative sums G_j defined by

$$G_j = \sum_{\substack{\text{all } i \text{ s.t.} \\ I_i \geq I_j}} T_i$$

can be formed. If p is increased to $p + \epsilon$ on all states at least as important as state j, the increase in P is approximately $G_j\epsilon$. Now suppose, for example, that all possible (non-absorbing) states are divided into two sets of states; the first set being the set of the most important states, S_1, that occur (about) half the time in the long run whilst the second set is the set of all remaining (least important) ones, S_2, that also occur (about) half the time in the long run. All states in S_1 are thus more important than (or at least as important as) all states in S_2. Now suppose p is increased by ϵ on all states in S_1 and decreased by ϵ on all states in S_2 ($\epsilon > 0$). Then, supposing k is the least important state in S_1, the overall increase in P by the above action is given by $\epsilon(2G_k - M)$ approximately, which is greater than zero unless all states are equally important (in which case it equals zero). Thus Morris (1977) noted that, by increasing p by ϵ on those points in the set S_1 and by decreasing it by ϵ on those points in the set S_2, player A's probability of winning the game was increased even though (it could well be argued that) no extra effort was made. We note that in some realizations of the game the extra effort will need to be made on more than half the points played whereas in other realizations it will need to be made less frequently than half the time. Morris also noted the multiplicative property of importances within a nested scoring system. For example, in tennis, "the importance of a point to winning the match, I_{PM}, can be obtained from the importance of the point to winning the game, I_{PG}, from the importance of the game to winning the set, I_{GS}, and from the importance of the set to winning the match, I_{SM}" in the following way

$$I_{PM} = I_{PG} \times I_{GS} \times I_{SM} \,.$$

1.4 Other Contributions

1.4.1. Squash

Studies of the game of international squash and related scoring systems are firstly considered. For a game of international squash, ApSimon (1957) found explicit expressions for the probability player A wins given that player A is about to serve (or recieve) and needs (say) x points to win while player B needs (say) y points to win. Watson (1970) evaluated the probability player A wins a game of international squash for various values of p. He noted that the scoring used in (say) table tennis could be regarded as a zero-order system, whilst that used in (say) international squash could be regarded as a first-order system since a point can only be won by the server. He noted that we could extend to second-order systems as follows. In such a system the receiver would have to win the rally to gain the serve, win the next rally to retain the serve and win the third rally in order to earn a point. Watson (1970) concluded that "the effect of increasing the order of the system of scoring is to increase the probability of the better player winning the game". Hsi and Burych (1971), using a bipoints structure, attempted to find the distribution of scores in badminton which uses the same service exchange mechanism as in international squash. They introduced the notions of "interruptions" and "exchanges" in their analysis but obtained incorrect results as the number of possible "exchanges" were incorrectly enumerated. Phillips (1978) indicated how this error could be corrected. In his interesting paper Phillips introduced two further notions — "disruptions" and "obstructions". These notions lead to three ways of dividing up the points within a game, which in turn gave three different expressions for the distribution of scores in (first-order) scoring systems such as the game of international squash (approximately B_{2n-1} points where $n = 9$) or the game of badminton. Phillips (1978), noting the work of Kingston (1976), observed the unfairness of the scoring systems used in international squash and badminton, and that the convergence of this unfairness to zero (as the n in B_{2n-1} tends to infinity) was slow. Phillips also investigated the "setting" of a game of international squash when 8–8 is reached for the first time. The receiver has the choice of "no set" ("short" game or "first-to-9") or "set two" ("long" game or "first-to-10"). Supposing player A is the receiver at 8–8, player A should select "no set" if and only if $p_b > p_a/q_a$ (in bipoints). In unipoints play, player A will select "no set" if and only if p (the probability player A wins a rally) is less than (about) 0.38. Clarke and Norman (1978, 1979) obtained the same results for international squash and also did a corresponding analysis for North American squash and Badminton, correcting some earlier work by Renick (1976, 1978). Phillips (1978) also wrote down, for the bipoints case, the distribution of scores in games such as North American squash. Using unipoints, Clarke and Norman (1979) evaluated the expected value and variance of the duration of a game of international squash and noted its larger variance than that of North American Squash (see also Schutz and Kinsey (1977)). Pollard (1980, 1983b, 1985a) studied the distribution of duration and the importances of the points within the international squash scoring system.

1.4.2. Tennis

Articles on tennis that were not discussed in Sections 1.2 or 1.3 are now considered. Gale (1971) investigated the common practice of serving a "strong" serve on the first serve and a "weak" serve on the second serve, if required. Using elementary probability and assuming a continuum of possible serves, then under reasonable assumptions he concluded that it would never make sense to serve a more risky serve as a second serve. Redington (1972) concluded that in the 1971 Wimbledon Final, Newcombe would have done slightly better if he had repeated his first serve instead of using a softer second serve and that the optimum second serve would be somewhere between the two serves used. George (1973) carried out a similar theoretical analysis to that of Gale reaching similar conclusions and then analysed a small data set of men's singles to determine which of the 3 admissible two service strategies, (strong-weak), (strong-strong) and (weak-weak) was best. The data supported the common practice that the (strong-weak) strategy is best and typically the (weak-weak) option was the worst of the three. King and Baker (1979) carried out the same analysis as George for a larger data set of women's singles. For one player the (weak-weak) strategy was preferable, while for another the (strong-strong) strategy was best, but for the remaining 10 players in the analysis it appeared that the (strong-weak) option was best. However, in no case did the data suggest the use of the (weak-strong) strategy, thus agreeing with the earlier-mentioned theory. King and Baker also showed that the statistics of a match could be used to determine how a player could most effectively improve performance. A rather similar approach to the above two-service problem was given by Hannan (1976) who also proposed a rather impractical game theory approach.

Kingston (1976) derived an interesting and possibly surprising property for those competitions which involve an inherent advantage such as service in tennis.

He showed that, for B_{2n-1} systems using bipoints, the initial server has the same probability of winning the match under two different rules (i) A and B serve alternate games and (ii) the winner of one game serves the next. Anderson (1977) gave a simple proof of this result.

Fischer (1980) attempted to find the probability player A wins a set and a match of tennis but obtained inaccurate results as he used a unipoints approximation to a bipoints situation.

Croucher (1981) showed that sporting data was a useful source in providing instruction for students to become familiar with probability concepts. Analysing the first 100 years of Wimbledon tennis finals, he rejected the hypothesis that the players were of equal ability and also concluded that, for men, there was a suggestion of a "back-to-the-wall" effect wherein players who lost the first set and recovered to take their opponents to five sets appeared to "have the edge" in the fifth set.

1.5 The Problems Considered in This Thesis and Some Results

With the exception of W_n(point-pairs), Miles (1984) produced biformats by the use of the reversal biformat structure. Section 2.2 of this thesis shows that it is also possible to produce biformats without using this structure. In particular, this section shows that an equivalent biformat to $W_{2n}PL$ (or $W_{2n}PW$) can be produced using *generalized point-pairs* (*g.p.p.*) rather than points. Section 2.2 also shows that W_n(point-pairs) and $W_{2n}(AL)$ not only have the same efficiency of unity, but in fact have the same distribution of duration.

Miles (1984) showed that the reversal biformat $W_{2n}PL$ has efficiency greater than unity when $n \geq 2$ and $p_a + p_b > 1$. Are there any other super-efficiency biformats? We will see in Sections 2.3 and 2.4 that, within the W_{2n} structure, there are in fact $n-2$ other biformats with $\rho > 1$ when $p_a + p_b > 1$. These $n-2$ other W_{2n} biformats may be called *partial play-the-loser* (PPL) biformats and can be listed in their increasing order of efficiency. When $p_a + p_b > 1$, the most efficient of these $n-1$ super-efficient biformats is in fact $W_{2n}PL$. An intuitive explanation of this super-efficiency is given in Section 2.5. Thus a *"spectrum"* of super-efficient biformats has been produced, each member of which can be obtained by using either individual points or generalized point- pairs.

The duration of the biformats in the above spectrum must be an even number of points. Is it possible to produce biformats with odd durations? Chapter 3 shows that it is. Using the W_{2n-1} structure, $n-1$ super-efficient biformats can be produced without the need for the reversal biformat structure. Thus the above spectrum of super-efficient biformats can be enlarged to include the cases of odd duration. Again, when $p_a + p_b > 1$, the (full) PL cases are seen to be the most efficient ones within this enlarged spectrum of W_n biformats, called the *"total spectrum"*.

Typically there are many different ways of producing the same PPL member of this total spectrum; these different ways are effectively permutations of each other. For example, it is shown in Section 3.3 that a PPL structure can be seen as a "mixture of the PL and AL structures". In fact some PPL members can be obtained by mixing the PL, AL and PW structures, provided the extent of the PL structures exceeds that of the PW structures. (See Sections 3.4 and 3.5.)

Every member of this total spectrum of biformats is shown to have the *constant probability ratio* (*c.p.r.*) property which is defined in Section 2.1 and implies that the random variable, duration, is independent of the indicator variable which equals 1 (0) when player A wins (loses). This property is very useful in generating many of the results of Chapters 2 and 3.

Is it possible to reduce some of the symmetry within the biformat structure? Figures 3.3.1 and 3.4.2 of Sections 3.3 and 3.4 show that it is.

The discussion of the total spectrum of biformats has so far been in terms of additional PPL members. Of course, all the corresponding partial play- the-winner (PPW) members exist and have efficiency greater (less) than unity when $p_a + p_b < 1$ ($p_a + p_b > 1$).

Section 3.6 addresses the question as to whether it is possible to find a super-efficient biformat if it is not known in advance whether $p_a + p_b$ is greater or less than unity. As a result of Theorem 3.1, it appears that it is not possible to find a biformat which is super-efficient in both the regions $p_a + p_b > 1$ and $p_a + p_b < 1$.

Many of the results mentioned so far have been achievable by making use of the *c.p.r.* property and without deriving the actual distributions of duration. In Section 3.7 matrix procedures and a probability generating function method are given for deriving the distribution of duration for these PL, AL and PW systems.

Every biformat within the total spectrum of W_n biformats has the following property: each point within that biformat is equally important when $p_a = p_b$. A diagram, the "$P - D - I$ diagram", which was found to be quite useful for the analysis of scoring systems, is introduced in Section 4.2. Morris (1977) noted that points were not equally important in tennis and Miles (1984) noted the inefficiency of the tennis scoring system. Is there a relationship between the efficiency of a biformat or uniformat (at (p_a, p_b) or p) and the importance of the points within it? An exact relationship is given in Section 4.3 as well as an approximate relationship for the case in which $(p_a, p_b) = (p+\delta, p-\delta)$ (or $p = \frac{1}{2} + \delta$ in unipoints) where δ is small. This approximate relationship is developed in Section 4.5 and it is shown that if the points are not equally important when the two players are equal (in unipoints or bipoints), the efficiency at $\frac{1}{2}+\delta$ in unipoints or $(p+\delta, p-\delta)$ in bipoints must be less than unity (δ small). Section 4.4 gives an extension of a well-known theorem in random walk theory.

We noted above that points are equally important (when $p_a = p_b$) for any biformat which is a member of the total spectrum of W_n biformats. Does this property characterize this total spectrum of W_n biformats? Section 5.1 shows that it does and we can conclude that, amongst the set of biformats whose limiting efficiencies at $(p+\delta, p-\delta)$ do not drop below 1 as $\delta \to 0$, the PL (PW) cases of the total spectrum of biformats are in fact *the optimal biformats* when $p_a + p_b > 1$ ($p_a + p_b < 1$). The structure of biformats ensures that the probabilities of the two types of error, α and β, are equal. However, the above optimal property carries over to the cases in which $\alpha \neq \beta$ (the "handicap" case). Section 5.2 uses the above characterization property to show that a certain tie-breaker (TB) system, which is really very similar to the corresponding AL system, can have an efficiency less than unity. A very brief discussion of 2 other bipoints problems: clinical trials (2-drugs) and the 2-armed bandit problem is given in Section 5.3.

Thus, Chapters 2 and 3 introduce generalized point-pairs and produce the total spectrum of W_n biformats (with the c.p.r. property) with or without the use of the reversal structure. Chapter 4 introduces the $P-D-I$ diagram (which was found to be quite useful) and formalizes a relationship between the efficiency of a uniformat or biformat and the importances of the points within it. The relationship is then used in Chapter 5 to produce the family of optimal biformats. These results are perhaps the theoretical highlights of this thesis.

Can the super-efficiency of the PL rule in the tennis context ($p_a + p_b > 1$) be extended to the tennis doubles situation in which there are essentially 4 types of point — "quadri-points"? Extending the notation of Section 1.2, Team A is the better team if $p_{a1} + p_{a2} > p_{b1} + p_{b2}$ and our hypotheses are now

$$H_0 : p_{a1} + p_{a2} > p_{b1} + p_{b2}$$

and

$$H_1 : p_{a1} + p_{a2} < p_{b1} + p_{b2} \, .$$

With 4 types of point, there is now a starting point problem. (We recall that

Miles overcame this starting point problem in the bipoints case by setting up the reversal biformat structure (so that, amongst other things, there would be just one efficiency).) For this quadri-points problem, consider W_n (n finite) for example. P and μ values depend on the initial point selected and on the initial order of selection of points. W_n used in conjunction with PL, PW, or AL structures leads to very complicated exact analyses (for finite n) and, of course, there are multiple efficiencies (depending on the starting point, etc.). These analyses are further complicated since we no longer have the $c.p.r.$ property. However, asymptotically ($n \to \infty$) the starting problem no longer exists and a *"modules"* approach can be used to derive asymptotic efficiencies. These modules act as independent units within the random walk and if a (complex) exchange mechanism (i.e. a rule determining the next type of point to be played) can be decomposed into smaller such independent units, then properties of these units and the theory of the general one-dimensional random walk can be used to obtain the asymptotic efficiency of the random walk with that exchange mechanism. This method of modules is introduced in Chapter 6. In Chapter 7 we consider quadri-points, set up "point-quads" (similar to point-pairs in the bipoints case) as the "standard" module and evaluate the asympototic efficiency of a module based on the PL rule. A single points and generalized point-pairs approach are again shown to be equivalent, although the single points approach is seen to be applicable to the cases of $6, 8, 10, \ldots$ types of point ("multi-points"). As may have been anticipated, this module based on the PL rule (Sections 7.2.1 and 7.2.2) and used with W_n (n large) systems is super-efficient for the quadri-points case if $\overline{p}_a + \overline{p}_b > 1$ and $|p_{a1} - p_{a2}| = |p_{b1} - p_{b2}|$, a rather limiting constraint (here $\overline{p}_a = (p_{a_1} + p_{a_2})/2$ and $\overline{p}_b = (p_{b1} + p_{b2})/2$).

An extension of unipoints is also considered in Section 7.4 and certain quite limited results based on bipoints and quadri-points are established.

Some applications of this module approach to general statistical problems are given in Section 6.4. For example, in statistical inference, the domain of asymptotic super-efficiency of a certain test of a 2-sample Poisson hypothesis testing problem is given. Also, in general, where α and β are type I and type II errors, the module which produces the asymptotically most efficient test for the $\alpha = \beta$ case is also the most efficient for the $\alpha \neq \beta$ case. In Example 6.3.2 the module method is used to provide an intuitive explanation as to why the PL rule is super-efficient when $p_a + p_b > 1$ in bipoints. Section 6.5 shows that the modules approach can be used to find asymptotically optimal strategies and Section 6.6 indicates some possible further studies.

Chapter 8 gives a detailed analysis of classical and tie-breaker tennis. The tennis scoring system can be seen as an inefficient test of the bipoints hypothesis of Section 1.2. Thinking about this specific scoring system and its inherent inefficiency gave rise to the work of Miles (1984) and the theory in Chapters 2 to 7 of this thesis. Using a bipoints approach, Chapter 8 gives expressions for the distribution, the mean and the variance of the number of points in a game, set or match of (best-of-3 sets of) classical or tie-breaker tennis and explains the long matches often previously observed in classical tennis. *The methodology used to derive these expressions could be used to obtain corresponding expressions for other "nested" scoring systems* (squash is 2-nested (game, match) whereas tennis is 3-nested (game, set, match)). In Section 8.9 a "best-of-5 half- sets" scoring system is shown to have a smaller variance of duration than the present "best-of-3 tie-breaker sets" system.

Chapter 9 briefly outlines some aspects in the design of scoring systems. In particular, the roles played by characteristics such as expected duration, μ, variance of duration, σ^2, and efficiency, ρ, are considered. The designer must make a choice between ρ and σ^2. Methods of reducing σ^2 are suggested, and some results relevant to B_{2n-1} and related systems are outlined. My supervisor, Dr Roger Miles, is currently carrying out investigations into this topic of the design of scoring systems (Miles, to appear).

Morris (1977) derived an approximate relationship between the increased probability of winning a point or set of points and the increased probability of winning a match. This relationship is generalized in Section 10.2 and an alternative approach is taken in Section 10.3. It is seen that interactions between states occur; interactions that can be positive or negative. The effect of increasing p in unipoints by δ on some points within the B_{2n-1} and W_n uniformats is investigated in Section 10.5 and a generalization of the results of Kingston (1976) and Anderson (1977) is noted.

In Chapter 11 team play is considered and for simplicity the unipoints case is discussed although the general conclusions undoubtedly also apply to the bipoints case. Player A_i has a probability p_i of winning a point against player B_i. In team play, player A_1 plays player B_1 according to some uniformat until A_1 or B_1 wins; player A_2 plays player B_2 according to the same uniformat until A_2 or B_2 wins. The winning team (A or B) must now be determined by some countback rule. What rule should be used? Also, is there a preferable underlying uniformat or scoring system that should be used for the individual matches? Chapter 11 considers these problems for the cases of 2 or 3 players per team. (Another application related to this problem is the case in which two players play each other 2 or 3 times.) The general conclusion is that "upward- nested" countback systems (e.g. points - games - sets - matches, in tennis) are preferable to "downward-nested" systems (matches - sets - games - points).

In sports scoring systems the concept of efficiency requires the better of two players to have a high probability of obtaining a greater score than that of the weaker player at the completion of the match. The corresponding requirement of examinations (which typically involve more than two persons) is that the rankings of the scores of the examinees should correspond as closely as possible to the rankings of the examinees' respective abilities being measured: the better examinees should have a high probability of obtaining better scores. The multiple choice method of examination and its associated (classical) method of scoring is investigated in Chapter 12. It is shown that the instructions given to the candidates and the current method of scoring commonly used by the examiner can be changed so that the scores obtained by the candidates more closely correspond to their respective abilities being measured. The proposed scoring system involves instructing the examinee to tick the correct box or cross a box or boxes known to be incorrect. Section 12.3 shows that a wide range of scoring systems is available within this structure, each of which, it is argued, removes guessing by examinees, thus eliminating a major source of random variation in examinee's scores.

CHAPTER 2

A SPECTRUM OF W_{2N} BIFORMATS

2.1 Introduction

Miles (1984) produced three reversal biformats using the W_{2n} structure — $W_{2n}PL$, $W_{2n}AL$ and $W_{2n}PW$. Sections 2.3 and 2.4 show that, for a given n (≥ 3), there are in fact $n-2$ reversal biformats 'between' $W_{2n}PL$ and $W_{2n}AL$ which use the W_{2n} structure and can be formed by 'mixing' the PL and AL rules. Correspondingly, for a given n (≥ 3), there are $n-2$ reversal biformats between $W_{2n}AL$ and $W_{2n}PW$. Thus, for a given n, there are in total $2n-1$ reversal biformats using the W_{2n} structure. Allowing n to take the values 1,2,3,..., we thus have a "*spectrum*" of reversal biformats. From Theorems 2.1 and 2.4 it follows that every member of this spectrum has a useful property known as the "*constant probability ratio*" (*c.p.r.*) property which is defined in Section 2.2. Each member of this spectrum has its own *c.p.r.*. Corollary 2.2.1 shows that, for any biformat with this *c.p.r.* property, the distribution of duration, the distribution of duration conditional on player A winning and the distribution of duration conditional on player A losing are identical. Corollary 2.2.2 shows that, for any biformats with the same *c.p.r.*, the distributions of duration are identical. Corollary 2.2.1 is used in Theorems 2.3 and 2.5 to derive a useful relationship between P and μ for each member of this spectrum of reversal biformats. This relationship together with the *c.p.r.* property is used to show that, for a given n, the efficiency of these $2n-1$ reversal biformats is monotonic (Theorem 2.7). It is deduced that, for a given n, $W_{2n}PL$ ($W_{2n}PW$) is the most efficient of these $2n-1$ reversal biformats when $p_a+p_b>1$ ($p_a+p_b<1$). Similarly, $W_{2n}PL$ ($W_{2n}PW$) is the least efficient when $p_a+p_b<1$ ($p_a+p_b>1$).

In this chapter an alternative method of producing, for a given n, biformats with the same distributions of duration (and same *c.p.r.*'s) as the $2n-1$ reversal biformats mentioned above is described. This method makes use of point-pairs rather than points and demonstrates that biformats can be produced without the need for the reversal biformat structure (cf. Miles (1984, p.103)).

2.2 The Constant Probability Ratio Property and Generalized Point-Pairs.

Every member of the spectrum of biformats discussed in Section 2.1 has a convenient property which is now introduced. If $P_k(F)$ ($Q_k(F)$) represents the probability player A wins (loses) in k points under biformat F scoring, the biformat F is said to have the ***constant win-loss probability ratio (c.p.r.)*** property if $P_k(F)/Q_k(F)$ is a constant for all k (in which $Q_k(F) > 0$). Thus, the indicator variable which equals 1 (0) when player A wins (loses), and the variable, duration, are independent when F has the *c.p.r.* property.

For each member of this spectrum of biformats there are two basic structures, one using points and the other using generalized point-pairs, which are now introduced. In bipoints, there are in fact 3 types of point-pairs ($\binom{a}{a}$, $\binom{a}{b}$ and $\binom{b}{b}$) and we take as an example their use with the PL exchange mechanism. A point-pair sequence ***play-the-loser generalized point-pairs*** ($PL\,g.p.p.$) is now defined. Play starts with an $\binom{a}{b}$ point-pair and the PL sequence applies to the points as pairs. Thus, $\binom{A}{A}$ is followed by $\binom{b}{b}$, a draw $\binom{A}{B}$ is followed by $\binom{a}{b}$ and $\binom{B}{B}$ is followed by $\binom{a}{a}$. Note that throughout this chapter, a draw $\binom{A}{B}$ is always followed by the point-pair $\binom{a}{b}$. Using the above exchange mechanism $PL\,g.p.p.$, we can form the family of biformats $\{W_n(PL\,g.p.p.)\}$ $(n = 1, 2, \ldots)$. This family is now shown to have the same *c.p.r.* and the same distribution of duration as the $\{W_{2n}PL\}$ $(n = 1, 2, \ldots)$ family.

Theorem 2.1. *The biformat $W_n(PL\ g.p.p.)$ and the reversal biformat $W_{2n}PL$ have the same constant win-loss probability ratio.*

Proof: Suppose F_1 is used to denote the reversal biformat $W_{2n}PL$ and F_2 is used to denote the biformat $W_n(PL\,g.p.p.)$. Consider firstly a drawn realization of (W_nPL^a, W_nPL^b) in which player A wins the W_nPL^a section in $n + 2m_1$ points $(m_1 = 0, 1, 2, \ldots)$ and loses the W_nPL^b section in $n+2m_2$ points $(m_2 = 0, 1, 2, \ldots)$. Then, for this drawn realization, there is an integer j (≥ 1) satisfying the following table.

Type of point	Outcome A	Outcome B	Total
a	j	$n + m_1 + m_2 - j$	$(1 + m_1) + (n + m_2 - 1)$
b	$n + m_1 + m_2 - j$	j	$(n + m_1 - 1) + (1 + m_2)$
Total	$(n + m_1) + m_2$	$m_1 + (n + m_2)$	$(n + 2m_1) + (n + 2m_2)$

This table uses the fact that the W_nPL^a section must have consisted of $(1 + m_1)$ a-points and $(n + m_1 - 1)$ b-points and that the W_nPL^b section must have had $(1 + m_2)$ b-points and $(n + m_2 - 1)$ a-points. It follows from this table that, for this drawn situation,

$$n_a - n_b = 0$$

where $n_a(n_b)$ is the total number of $a-$ ($b-$) points played. Also, the probability of this draw in the order in which the points occurred is

$$(p_a p_b)^j (q_a q_b)^{n+m_1+m_2-j} .$$

Now consider a drawn realization of $(W_n PL^a, W_n PL^b)$ in which player A loses the $W_n PL^a$ section in $n+2m_3$ points ($m_3 = 0,1,2,\ldots$) and wins the $W_n PL^b$ section in $n+2m_4$ points ($m_4 = 0,1,2,\ldots$). Then, for this drawn realization, there is an integer k (≥ 0) satisfying the following table

Type of point	Outcome A	B	Total
a	k	$n+m_3+m_4-k$	$(1+n+m_3-1)+m_4$
b	$n+m_3+m_4-k$	k	$m_3+(1+n+m_4-1)$
Total	$m_3+(n+m_4)$	$(n+m_3)+m_4$	$(n+2m_3)+(n+2m_4)$

Similarly, for this drawn realization,

$$n_a - n_b = 0$$

and the probability of this draw in the order in which the points occurred is

$$(p_a p_b)^k (q_a q_b)^{n+m_3+m_4-k} .$$

Now finally consider a realization of $(W_n PL^a,\ W_n PL^b)$ in which player A wins both sections and wins in a total of $2n+2m$ points ($m = 0,1,2,\ldots$). Then there is an integer i (≥ 1) satisfying the following table.

Type of point	Outcome A	B	Total
a	i	$1+m-i$	$1+m$
b	$2n+m-i$	$i-1$	$1+(2n+m-2)$
Total	$2n+m$	m	$2n+2m$

The row and column totals in this table can be produced in a similar manner to those of the two earlier tables and it can be seen that, for this realization

$$n_a - n_b = -(2n-2)$$

and the probability of this win to player A in the order in which the points occurred is

$$p_a {q_b}^{2n-1} (p_a p_b)^{i-1} (q_a q_b)^{1+m-i} \, .$$

Thus, noting these three realizations of $(W_n PL^a, W_n PL^b)$, it can be seen that, for F_1, at any point of absorption,

$$n_a - n_b = \begin{cases} -(2n-2) & \text{when player } A \text{ wins} \\ 2n-2 & \text{when player } A \text{ loses} \end{cases}$$

where $n_a(n_b)$ is the total number of a- (b-) points played. Also,

$$P_{2n+2m} = p_a q_b^{2n-1} \{ f_1(p_a p_b, \, q_a q_b) \}$$

where f_1 is a homogeneous polynomial of degree m in $p_a p_b$ and $q_a q_b$. Since any winning path to absorption has a mirror-image losing path to absorption (by making the exchanges, win $\leftrightarrow$ loss and $W_n PL^a \leftrightarrow W_n PL^b$), it follows that

$$Q_{2n+2m} = p_b q_a^{2n-1} \{ f_1(p_a p_b, \, q_a q_b) \}$$

and hence F_1 has the *c.p.r.*

$$(P_{2n+2m})/(Q_{2n+2m}) = (p_a {q_b}^{2n-1})/(p_b {q_a}^{2n-1}) \quad (m = 0, 1, 2, \ldots) \, .$$

It can also be shown that F_2 has the same *c.p.r.*. For example, consider a realization which player A wins with $n+h$ wins, h losses and d draws. For this realization there are non-negative integers i, j, k and l satisfying the following table.

Type of Point-Pair	Outcome AA	AB	BB	Total
aa	i	j	$h-i-j$	h
ab	k	l	$1+d-k-l$	$1+d$
bb	$n+h-i-k$	$d-j-l$	$i+j+k+l-d-1$	$n+h-1$
Total	$n+h$	d	h	$n+2h+d$

The probability of these $n+h$ wins, h losses and d draws in the order in which they occurred is

$$\begin{aligned} &\left\{ (p_a^2)^i (2p_a q_a)^j ({q_a}^2)^{h-i-j} \right\} \times \left\{ (p_a q_b)^k (p_a p_b + q_a q_b)^l (p_b q_a)^{1+d-k-l} \right\} \\ &\quad \times \left\{ ({q_b}^2)^{n+h-i-k} (2p_b q_b)^{d-j-l} ({p_b}^2)^{i+j+k+l-d-1} \right\} \\ &= p_a {q_b}^{2n-1} \left\{ f_2(p_a p_b, \, q_a q_b) \right\} \end{aligned}$$

where $f_2(p_a p_b, q_a q_b)$ is a homogeneous polynomial of degree $2h + d$ in $p_a p_b$ and $q_a q_b$. Thus, by a similar argument to above, F_2 has the same *c.p.r.* as F_1.

Theorem 2.2. *The biformat $W_n(PL\,g.p.p.)$ and the reversal biformat $W_{2n}PL$ have the same distribution of duration.*

Proof: Using the same notation as in the proof of Theorem 2.1, let $P_{(i)}$ be the probability player A wins F_i $(i = 1, 2)$. Then

$$P_{(1)} = p_a {q_b}^{2n-1}\{1 + (c_{11} p_a p_b + c_{12} q_a q_b) + \big(c_{21}(p_a p_b)^2 + c_{22}(p_a p_b)(q_a q_b) \\ + c_{23}(q_a q_b)^2\big) + \ldots\}$$

and

$$P_{(2)} = p_a {q_b}^{2n-1}\{1 + (c'_{11} p_a p_b + c'_{12} q_a q_b) + \big(c'_{21}(p_a p_b)^2 + c'_{22}(p_a p_b)(q_a q_b) \\ + c'_{23}(q_a q_b)^2\big) + \ldots\}$$

where the c_{ij} and c'_{ij} are (integer) constants. Since $P_{(1)} = P_{(2)}$ for all (p_a, p_b) it follows that $c_{ij} = c'_{ij}$ for all i and j and the result follows.

Corollary 2.2.1. *For any biformat having the c.p.r. property, the unconditional distribution of duration, the distribution of duration conditional on player A winning and the distribution of duration conditional on losing are identical.*

Corollary 2.2.2. *For any two biformats with the same c.p.r. , the above three distributions of duration are identical.*

Theorem 2.3. *For the biformats $W_n(PL\,g.p.p.)$ and $W_{2n}PL$,*

$$\frac{P - Q}{\mu} = \frac{p_a - p_b}{2(1 + (n-1)(p_a + p_b))}$$

where μ is the expected duration, P is the probability player A wins and $Q = 1 - P$.

Proof: Suppose D_{2n+2j} $(j = 0, 1, 2, \ldots)$ is the distribution of duration conditional on a win to player A. Also, suppose D'_{2n+2j} $(j = 0, 1, 2, \ldots)$ is the distribution of duration conditional on a loss to player A. These distributions are actually identical (i.e. $D_{2n+2j} = D'_{2n+2j}$ $(j = 0, 1, 2, \ldots)$). Note that if player A wins (loses) in $2n + 2j$ points, $j + 1(2n - 1 + j)$ a-points and $2n - 1 + j(j + 1)$ b-points must have been played. Note also that the expected increase (decrease) in player A's score when an a-point (b-point) is played is $2p_a - 1$ $(2p_b - 1)$. The expected value of player A's final score is $2n(P-Q)$ which is also given by

$$P\left\{\sum_{j=0}^{\infty} D_{2n+2j}\big((j+1)(2p_a-1)-(2n-1+j)(2p_b-1)\big)\right\}$$
$$+Q\left\{\sum_{j=0}^{\infty} D'_{2n+2j}\big((2n-1+j)(2p_a-1)-(j+1)(2p_b-1)\big)\right\}$$
$$=P\left\{\sum_{j=0}^{\infty} D_{2n+2j}2j(p_a-p_b)\right\}+P\left\{(2p_a-1)-(2n-1)(2p_b-1)\right\}$$
$$+Q\left\{\sum_{j=0}^{\infty} D'_{2n+2j}2j(p_a-p_b)\right\}+Q\left\{(2n-1)(2p_a-1)-(2p_b-1)\right\}$$
$$=\mu(p_a-p_b)-(P-Q)(2n-2)(p_a+p_b-1)$$

since $\sum_{j=0}^{\infty} D_{2n+2j}2j=\mu-2n$, where the mean duration conditional on winning, μ_W, equals the mean duration conditional on losing, μ_L, and equals μ. Equating this last expression with $2n(P-Q)$, the result follows.

If we set up the play-the-winner generalized point-pairs sequence ($PW\ g.p.p.$) correspondingly, it follows that $\{W_n(PW\ g.p.p.)\}(n=1,2,\ldots)$ has the same distribution of duration as $\{W_{2n}PW\}$ $(n=1,2,\ldots)$, their *c.p.r.* is given by

$$\frac{P}{Q}=\frac{{p_a}^{2n-1}q_b}{{p_b}^{2n-1}q_a}$$

and

$$\frac{P-Q}{\mu}=\frac{p_a-p_b}{2\big(1+(n-1)(q_a+q_b)\big)}.$$

Miles has shown that $W_{2n}AL$ and W_n (point- pairs) have the same efficiency. However, as with the PL and PW exchange mechanisms, it follows that $\{W_n(\text{point}-\text{pairs})\}$ $(n=1,2,\ldots)$ in fact has the *same distribution* as $\{W_{2n}AL\}$ $(n=1,2,\ldots)$, their *c.p.r.* is given by

$$\frac{P}{Q}=\frac{{p_a}^n{q_b}^n}{{p_b}^n{q_a}^n}$$

and

$$\frac{P-Q}{\mu}=\frac{p_a-p_b}{2n}.$$

(It may be noted in passing that, in unipoints, W_n has the *c.p.r.*

$$\frac{P}{Q}=\frac{p^n}{q^n}$$

and.

$$\frac{P-Q}{\mu}=\frac{p-q}{n}.)$$

2.3 A Biformat between $W_{2N}AL$ and $W_{2N}PL$ (N odd).

Returning the points rather than point-pairs, the outcome of these points can be represented by a sequence of A's and B's, which can be split into two subsequences — one subsequence consisting only of all the A's numbered in sequential order A_1, A_2, A_3, ... and a second subsequence consisting only of all the B's also numbered in their sequential order. Consider the exchange mechanism in which the odd numbered A's $(A_1, A_3, A_5, \ldots)$ are followed by b-points and the odd numbered B's are followed by a-points whilst points alternate $(a\,b\,a\,b\ldots)$ for the even-numbered A's and B's (taken in their realized order in the full sequence rather than in the subsequences). Then $W_nPL^a_{odd}$ represents a scoring system, starting with an a-point, which is won by the first player to achieve a lead of n points over the other using the above exchange mechanism. Figure 2.3.1 gives an example of a realization of $W_3PL^a_{odd}$ won by player A. Such a scoring system can be seen as a "mixture" of AL and PL. Note that $W_nPL^a_{even}$ can be correspondingly defined.

Figure 2.3.1
A realization of $W_3PL^a_{odd}$

Type of Point	a		b		a			b		a	b			AL sub-sequence
		a		b		b	a		b			a	b	PL sub-sequence
Winner			A_1	A_2	A_3		A_4	A_5		A_6		A_7	A_8	A sub-sequence
	B_1	B_2				B_3			B_4		B_5			B sub-sequence
Score	-1	-2	-1	0	1	0	1	2	1	2	1	2	3	

Now returning to point-pairs, the above PL_{odd} exchange mechanism can be applied to point-pairs with the modification that $\binom{A}{A}_2, \binom{A}{A}_4, \binom{A}{A}_6, \ldots$, and $\binom{B}{B}_2$, $\binom{B}{B}_4, \binom{B}{B}_6, \ldots$ are followed by the $\binom{a}{b}$ point-pair, as are all draws $\binom{A}{B}$. Thus, under this PL_{odd} system, the outcomes $\binom{A}{A}_1, \binom{A}{A}_3, \binom{A}{A}_5, \ldots$ of the $\binom{A}{A}$ subsequence are followed by $\binom{b}{b}$ and the outcomes $\binom{B}{B}_1, \binom{B}{B}_3, \binom{B}{B}_5, \ldots$ of the $\binom{B}{B}$ subsequence are followed by $\binom{a}{a}$ whilst all other outcomes are followed by an $\binom{a}{b}$ point-pair. The $\binom{a}{b}$ point-pair is again used to start the process. Thus, the generalized point-pair biformats $\{W_n(PL_{odd}\,g.p.p.)\}$ $(n = 1, 3, 5, \ldots)$ can be set up using the above exchange mechanism. Again, $\{W_n(PL_{even}\,g.p.p.)\}$ $(n = 1, 3, 5, \ldots)$ can be correspondingly defined and, the methods of Section 2.2 can be used to show that, provided n is odd, the odd and even versions have the same distribution of duration,

$$\left.\begin{aligned} c.p.r. = \frac{P}{Q} = \left(p_a^{\frac{n+1}{2}} q_b^{\frac{3n-1}{2}}\right) \Big/ \left(p_b^{\frac{n+1}{2}} q_a^{\frac{3n-1}{2}}\right) \quad \text{and} \\ (P-Q)/\mu = (p_a - p_b) \Big/ \left(2\left(\tfrac{n+1}{2} + \tfrac{n-1}{2}(p_a + p_b)\right)\right) . \end{aligned}\right\} \qquad (2.3.1)$$

Proofs of a more general case are given in Section 2.4. Again, using the methods of Section 2.2, it can be shown that, provided n is odd, the four reversal biformats $\left\{W_1\left(W_n PL^a_{\substack{odd\\or\\even}}\ , W_n PL^b_{\substack{odd\\or\\even}}\ \right)\right\}$ and the two biformats $\left\{W_n\left(PL_{\substack{odd\\or\\even}}\ g.p.p.\right)\right\}$ have identical distributions of duration and the same $c.p.r.$'s. Corresponding results for W_{2n} systems using odd or even versions of the PW mechanism can also be written down.

2.4 A Spectrum of W_{2N} Biformats

The biformats considered in Section 2.3 could be called *partial play-the-loser* (PPL) biformats. However, there are more PPL biformats than just even or odd. For example, consider the W_{10} system. $W_5(PL\,g.p.p.)$ has $c.p.r.$ $\frac{p_a q_b{}^9}{p_b q_a{}^9}$ while W_5 (point-pairs) has $c.p.r.$ $\frac{p_a{}^5 q_b{}^5}{p_b{}^5 q_a{}^5}$. The $\frac{p_a{}^2 q_b{}^8}{p_b{}^2 q_a{}^8}$ $c.p.r.$ can be generated by *not* using PL on (say) every 4th loss to either player; the $\frac{p_a{}^3 q_b{}^7}{p_b{}^3 q_a{}^7}$ $c.p.r.$ is the odds/evens case discussed in Section 2.3 while the $\frac{p_a{}^4 q_b{}^6}{p_b{}^4 q_a{}^6}$ $c.p.r.$ can be generated by using PL only on (say) every 4th loss. Generalizing, $W_n(PL\,g.p.p.)$ has $c.p.r.$ $\frac{p_a q_b{}^{2n-1}}{p_b q_a{}^{2n-1}}$ and W_n(point-pairs) has $c.p.r.$ $\frac{p_a{}^n q_b{}^n}{p_b{}^n q_a{}^n}$. Between these values, there are $(n-2)$ available $c.p.r.$'s each of which can be achieved by using PPL biformats. Consider the $c.p.r.$ $\frac{p_a{}^{n-l} q_b{}^{n+l}}{p_b{}^{n-l} q_a{}^{n+l}}$. Using modulus $n-1$ arithmetic, suppose the $\binom{A}{A}$ wins numbered $i_1, i_2, \ldots, i_l$ (for any non-equal i_j from the set $\{0, 1, 2, \ldots, n-2\}$) in the $\binom{A}{A}$ subsequence are followed by $\binom{b}{b}$ point-pairs and correspondingly suppose the $\binom{B}{B}$ wins numbered (also using modulus $n-1$ arithmetic) $i_1, i_2, \ldots, i_l$ (for the (same as above) non-equal i_j from the set $\{0, 1, 2, \ldots, n-2\}$) in the $\binom{B}{B}$ subsequence are followed by $\binom{a}{a}$ point-pairs. All other outcomes are followed by an $\binom{a}{b}$ point-pair, which also initiates play. Using this exchange mechanism the associated PPL biformat $W_n(PL_l\,g.p.p.)$ $(l = 0, 1, 2, \ldots, n-1)$ has the above $c.p.r.$ and

$$\frac{P-Q}{\mu} = \frac{p_a - p_b}{2(n - l + l(p_a + p_b))} .$$

(Note that $l = 0$ gives us the original point-pairs case, W_n(point-pairs), while $l = n-1$ gives us the full PL case, $W_n(PL\,g.p.p.)$.) Proofs of these results are now given.

Theorem 2.4. *The spectrum of biformats* $\{W_n(PL_l\, g.p.p.)\}$ $(n = 1,2,3,\ldots;$ $l = 0,1,\ldots,n-1)$ *have c.p.r.*

$$\frac{{p_a}^{n-l}{q_b}^{n+l}}{{p_b}^{n-l}{q_a}^{n+l}}.$$

Proof: A similar approach is taken to that in the proof of Theorem 2.1. Consider a realization which player A wins on player A's $(k(n-1)+m)^{\text{th}}$ *g.p.p.* win. Suppose there are r i_j's in the set $\{1,2,\ldots,m-1\}$ and that there are d draws in the realization. Then, for this realization, there are non-negative integers e,f,g,h satisfying the following table

Type of point-pair	Outcome AA	AB	BB	Total
aa	e	d_1	f	$(k-1)l+r$
ab	g	d_2	h	$(2k-1)(n-1-l)$ $+2(m-1-r)+1+d$
bb	d_3	d_4	d_5	$kl+r$
Total	$k(n-1)+m$	d	$(k-1)(n-1)+(m-1)$	$2k(n-1)+2m$ $+d-n$

where the d_i's are differences that can be obtained from the row and column totals. The probability of these $k(n-1)+m$ wins, d draws and $(k-1)(n-1)+(m-1)$ losses in the order in which they occurred is

$$\left\{({p_a}^2)^e(2p_aq_a)^{d_1}({q_a}^2)^f\right\} \times \left\{(p_aq_b)^g(p_ap_b+q_aq_b)^{d_2}(p_bq_a)^h\right\} \\ \times \left\{({q_b}^2)^{d_3}(2p_bq_b)^{d_4}(p_b^2)^{d_5}\right\}.$$

which can be shown to be equal to a constant multiplied by

$${p_a}^{n-l}{q_b}^{n+l}(p_ap_b)^{kl+r+e+g-n-f}(q_aq_b)^{kl+r+h+f-l-e}(p_ap_b+q_aq_b)^{d_2}$$

which equals

$${p_a}^{n-l}{q_b}^{n+l}f\big((p_ap_b),(q_aq_b)\big)$$

where $f\big((p_ap_b),(q_aq_b)\big)$ is a homogeneous polynomial of degree $2k(n-1)+2m+d-2n$ in p_ap_b and q_aq_b.

Now consider the mirror image losing path to absorption which can be obtained by the exchanges $aa \leftrightarrow bb$, $AA \leftrightarrow BB$ in the above table. The probability of this path can be shown to equal the same constant as above multiplied by

$$p_b{}^{n-l} q_a{}^{n+l} (p_a p_b)^{kl+r+e+g-n-f} (q_a q_b)^{kl+r+h+f-l-e} (p_a p_b + q_a q_b)^{d_2}$$

which equals

$$p_b{}^{n-l} q_a{}^{n+l} f\big((p_a p_b), (q_a q_b)\big)$$

and hence the result follows as in the proof of Theorem 2.1.

Theorem 2.5. *The spectrum of biformats* $\{W_n(PL_l\, g.p.p.)\}$ $(n = 1, 2, 3, \ldots; l = 0, 1, \ldots, n-1)$ *have the property*

$$\frac{P - Q}{\mu} = \frac{p_a - p_b}{2(n - l + l(p_a + p_b))}\,.$$

Proof: From the proof of Theorem 2.4 we note that, for any realization of the above biformats,

$$n_{aa} - n_{bb} = \begin{cases} -l & \text{when player } A \text{ wins} \\ +l & \text{when player } A \text{ loses} \end{cases}$$

where $n_{aa}(n_{bb})$ is the total number of $aa-$ $(bb-)$ point-pairs played. The proof of this theorem follows in the same way as in Theorem 2.3.

General PPL reversal biformats $W_1(W_n PL_l^a,\, W_n PL_l^b)$ can be set up for individual points (rather than $g.p.p.$'s) using a method corresponding to that described for the $g.p.p.$ biformat $W_n(PL_l\, g.p.p.)$ and also corresponding to the single-points approach described in the first paragraph of Section 2.3. It is now shown that these two biformats have the same distribution of duration.

Theorem 2.6. *The biformats* $W_1(W_n PL_l^a,\, W_n PL_l^b)$ *and* $W_n(PL_l\, g.p.p.)$ *have the same distribution of duration.*

Proof: Firstly consider the $W_n PL_l^a$ section of the reversal biformat and let us suppose that player A wins on player A's $(k(n-1)+m)^{\text{th}}$ point win. As in the proof of Theorem 2.4 let us assume that there are r i_j's in the set $\{1, 2, \ldots, m-1\}$. Then, for this realization, there is a non-negative integer i satisfying the table

Type of point	Outcome A	Outcome B	Total
a	i	$(k-1)l+r+E-i$	$(k-1)l+r+E$
b	$k(n-1)+m-i$	$kl+r+E-k(n-1)-m+i$	$kl+r+E$
Total	$k(n-1)+m$	$(k-1)(n-1)+(m-1)$	$(2k-1)(n-1)+2m-1$

where it is assumed that the number of points in the alternating subsequence, $(2k-1)(n-1-l)+(2m-1-2r)$, is even and is denoted by $2E$. The probability of these $k(n-1)+m$ wins and $(k-1)(n-1)+(m-1)$ losses in the order in which they occurred is equal to

$$\left\{{p_a}^i {q_a}^{(k-1)l+r+E-i}\right\} \times \left\{{q_b}^{k(n-1)+m-i} {p_b}^{kl+r+E-k(n-1)-m+i}\right\} .$$

Now suppose player A wins the other section of the reversal biformat, $W_n PL_l^b$, on player A's $(k'(n-1)+m')^{\text{th}}$ point win, and that there are r' i_j's in the set $\{1,2,\ldots,m'-1\}$. Then a similar table to the above can be written down for this realization and $2E'$ defined by

$$2E' = (2k'-1)(n-1-l)+(2m'-1-2r')$$

must be even whenever $2E$ is even. The probability of these $k'(n-1)+m'$ wins and $(k'-1)(n-1)+(m'-1)$ losses in the order in which they occurred is the same as above with i, k, m, r and E replaced by i', k', m', r' and E' respectively. The product of these two probabilities is

$${p_a}^{n-l}{q_b}^{n+l}(p_a p_b)^{i+i'+l-n}(q_a q_b)^{k'n+kn+m+m'-k-k'-i-i'-n-l} .$$

Thus, using the same mirror image argument as earlier, it follows that, by comparing this win-win (for the two sections of the reversal biformat) case with the loss-loss case, with $2E$ and $2E'$ even, this reversal biformat has $c.p.r.$

$$\frac{{p_a}^{n-l}{q_b}^{n+l}}{{p_b}^{n-l}{q_a}^{n+l}} .$$

For the case in which $2E$ is odd and hence $2E'$ is odd, there is one more a-(b-)point than b- (a-) point in the alternating subsequence of the $W_n PL_l^a (W_n PL_l^b)$ section of the reversal biformat. These two effects compensate each other and the above $c.p.r.$ still results.

In the draw situation (win-loss or loss-win for the two sections of the reversal biformat) the product of the two probabilities is

$$(p_a p_b)^{e_1} (q_a q_b)^{e_2}$$

where e_1 and e_2 are two expressions representing the powers of $p_a p_b$ and $q_a q_b$. Thus the probability of this draw divided by the probability of its mirror image draw is 1 and, as the game is returned to its starting point by a draw, the above analysis of the (win-win) and (loss-loss) cases is sufficient to conclude that $W_n(PL_l\, g.p.p.)$ and $W_1(W_nPL_l^a,\, W_nPL_l^b)$ have the same $c.p.r.$ and same distributions of durations, thus completing the proof.

Using a corresponding PPW exchange mechanism the spectrum of biformats

$$\{W_n(PW_l\, g.p.p.)\} \quad \text{and} \quad \{W_1(W_nPW_l^a,\, W_nPW_l^b)\} \quad (n = 1, 2, 3, \ldots; \; l = 0, 1, 2, \ldots, n-1)$$

can be formed, having the properties

$$c.p.r. = \frac{{p_a}^{n+l}{q_b}^{n-l}}{{p_b}^{n+l}{q_a}^{n-l}}$$

and

$$\frac{P-Q}{\mu} = \frac{p_a - p_b}{2(n - l + l(q_a + q_b))} \, .$$

We now show that if $W_{2n}PL$ is more efficient than $W_{2n}AL$, it is also more efficient than $W_n(PL_l\, g.p.p.)$ for all $l = 1, 2, 3, \ldots, n-2$.

Theorem 2.7. *If $p_a + p_b > 1$, the efficiency of $(W_nPL_l\, g.p.p)$ $(l = 0, 1, 2, \ldots, n-1)$ increases monotonically as l increases.*

Proof: By removing the constants in equation (1.2.1) the relative efficiency of a general biformat with key characteristics P and μ is given by

$$\frac{P-Q}{\mu} \ln \frac{P}{Q} \qquad \text{(Miles (1984, p.97 and p.103))} \, .$$

Using Theorems 2.4 and 2.5, the relative efficiency of $W_n(PL_l\, g.p.p.)$ is thus

$$\big\{(p_a - p_b)/\big(2(n - l + l(p_a + p_b))\big)\big\} \ln\big(({p_a}^{n-l}{q_b}^{n+l})/({p_b}^{n-l}{q_a}^{n+l})\big) \, .$$

Now let $p_a = p + \delta$, $q_a = q - \delta$, $p_b = p - \delta$, $q_b = q + \delta$ where $q = 1 - p$. The relative efficiency of $W_n(PL_l\, g.p.p.)$ is thus

$$\frac{2\delta}{2(n - l + 2lp)} \left[(n-l) \ln \frac{1 + \frac{\delta}{p}}{1 - \frac{\delta}{p}} + (n+l) \ln \frac{1 + \frac{\delta}{q}}{1 - \frac{\delta}{q}} \right] .$$

Expanding the logarithm terms, the relative efficiency is approximately

$$\frac{2\delta}{n-l+2lp}\left[(n-l)\left\{\frac{\delta}{p}+\frac{1}{3}\left(\frac{\delta}{p}\right)^3+\ldots\right\}+(n+l)\left\{\frac{\delta}{q}+\frac{1}{3}\left(\frac{\delta}{q}\right)^3+\ldots\right\}\right]. \quad (2.4.1)$$

In the above expression, the term in δ^2 is $\frac{2\delta^2}{pq}$ for all $l=0,1,2,\ldots,n-1$ whilst the term in δ^4 increases monotonically as l increases provided $2p>1$ (i.e. $p_a+p_b>1$), thus completing the proof.

Corollary 2.7.1. *If $p_a+p_b>1$, the efficiency of $(W_nPW_l\,g.p.p.)$ $(l=0,1,2,\ldots,n-1)$ decreases monotonically as l increases.*

Corollary 2.7.2. *If $p_a+p_b<1$, the efficiency of $W_n(PL_l\,g.p.p.)$ $(l=0,1,2,\ldots,n-1)$ decreases monotonically as l increases, whilst the efficiency of $W_n(PW_l\,g.p.p.)$ $(l=0,1,2,\ldots,n-1)$ increases monotonically as l increases.*

Corollary 2.7.3. *The efficiency of $W_n(PL\,g.p.p.)$ increases (decreases) as n increases if $p_a+p_b>1$ $(p_a+p_b<1)$ whilst the efficiency of $W_n(PW\,g.p.p.)$ increases (decreases) as n increases if $p_a+p_b<1$ $(p_a+p_b>1)$.*

2.5 Summary.

In this chapter two methods, one using single-points with the reversal structure and the other using generalized point-pairs, have been used to produce a spectrum of biformats with $c.p.r.$'s given by

$$\frac{p_a^t q_b^{2n-t}}{p_b^t q_a^{2n-t}} \qquad (n=1,2,3,\ldots;\ \ t=1,2,3,\ldots,2n-1)\,.$$

Every member of this spectrum has the property that the distribution of duration conditional on player A winning, the distribution of duration conditional on player A losing and the unconditional distribution of duration are identical. For a given n, the efficiency of the associated biformats decreases (increases) as t increases if $p_a+p_b>1$ $(p_a+p_b<1)$. Thus, for a given n, the case in which $t=1$ or the PL case ($t=2n-1$ or the PW case) produces the most efficient biformat when $p_a+p_b>1$ $(p_a+p_b<1)$. Also, when $t=1$, the efficiencies of the associated biformats increase (decrease) as n increases if $p_a+p_b>1$ $(p_a+p_b<1)$, whilst, when $t=2n-1$, the efficiencies increase (decrease) as n increases if $p_a+p_b<1$ $(p_a+p_b>1)$. Theorem 2.5 developed a useful relationship between P and μ for this spectrum of biformats.

There is an intuitive explanation for this super-efficiency of these PL and PW mechanisms. It is that, in each case, under the most efficient mechanism a greater number of points with the more variable (Bernoulli) distribution is expected to be selected — an approach used in efficient estimation theory. An alternative explanation is given in Example 6.3.2.

The spectrum of biformats considered in this chapter had durations with an even number of points. Is it possible to extend this spectrum and produce biformats with the $c.p.r.$ property but with *odd* durations? This problem is addressed in Chapter 3.

CHAPTER 3
THE TOTAL SPECTRUM OF W_N BIFORMATS AND SOME FURTHER RESULTS

3.1 Introduction

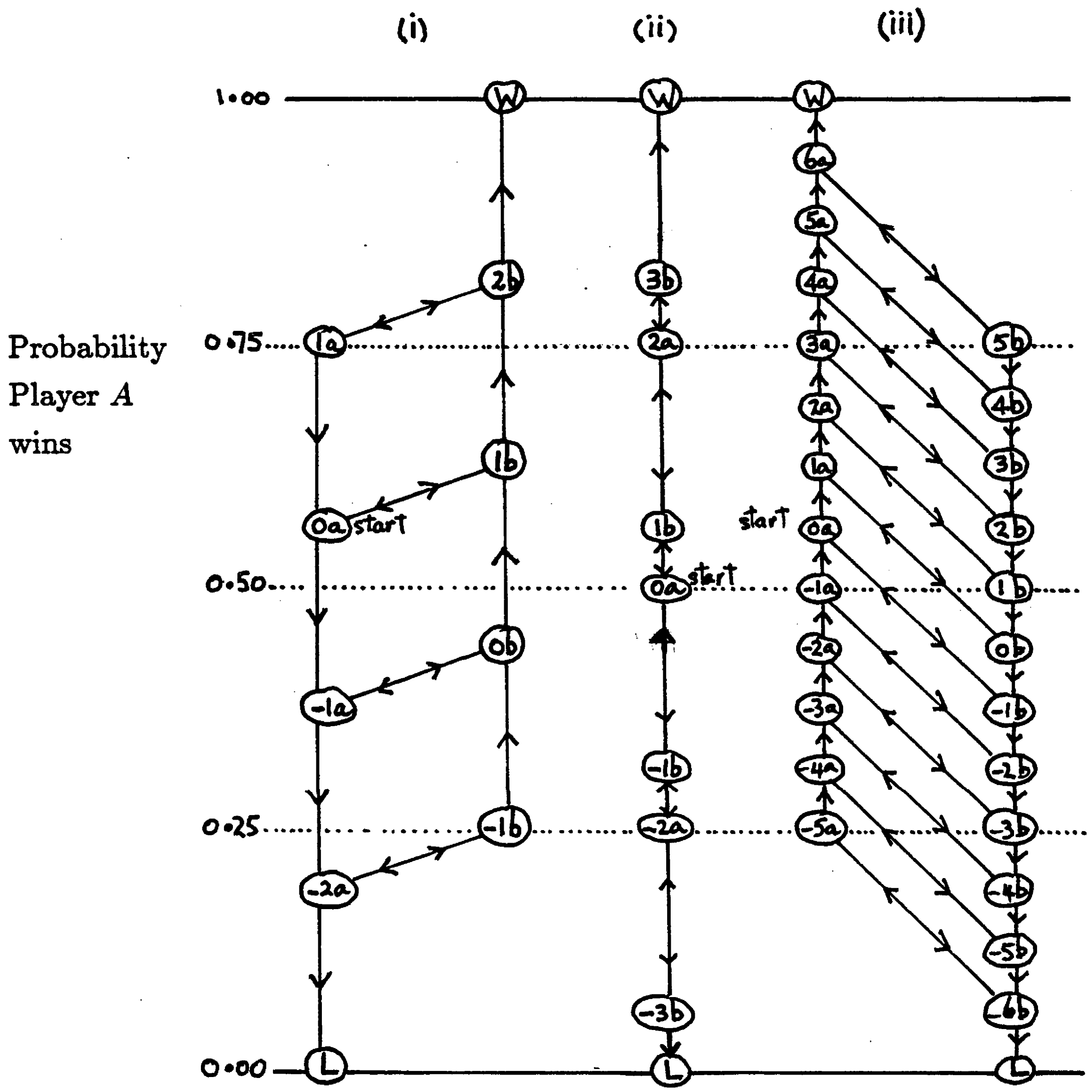
The spectrum of W_{2n} biformats considered in Chapter 2 had the *c.p.r.* property and durations with an *even* number of points. Is it possible to extend this spectrum and produce biformats with the *c.p.r.* property but with *odd* durations? Sections 3.3 and 3.4 show that it is. However, firstly Section 3.2 gives flow diagram representations of the PL-, AL-, and PW-structures. These diagrams are then used in Section 3.3 to show a way of producing biformats within the *c.p.r.* property and odd durations. This method does not involve the use of the "reversal" structure set up by Miles (1984). An alternative method of producing these "odd" cases which does use the reversal structure is then given in Section 3.4. The spectrum of all biformats with the *c.p.r.* property and odd and even durations is called the *total spectrum* of W_n biformats and the methods of Sections 3.3 and 3.4 which are used to produce the odd cases, can in fact be used to generate the "even" cases as well. Thus the total spectrum of W_n biformats can be produced by either of these methods. Section 3.4 also shows how the PL-, AL- and PW-structures can be "mixed" and still produce members of this total spectrum. Some further comments and examples are then given in Section 3.5. Theorem 3.1 of Section 3.6 suggests that no advantage (in terms of efficiency) can be made of the PL and PW rules if it is not known, *a priori*, whether $p_a + p_b$ is greater or less than unity. Finally, Section 3.7 *outlines* some methods that can be used to obtain the actual *distribution* of duration of some of these W_n systems.

3.2 The Basic PL-, AL- and PW-structures.

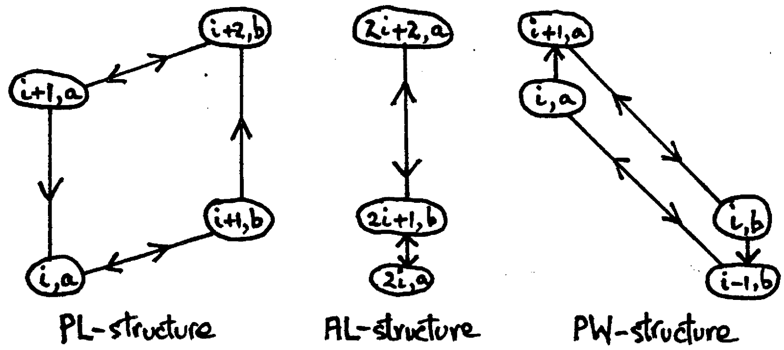
Figure 3.2.1(a) gives flow diagram representations of 3 scoring systems ((i) W_3PL^a (ii) W_4AL^a (ii) W_7PW^a) for the case in which $p_a = p_b = 0.75$. In these diagrams an upwards (downwards) arrow indicates a win (loss) to player A and the probability that player A wins from any state can be read off the vertical or y-axis for that state. Figure 3.2.1(b) highlights the fact that there are essentially 3 basic structures in these diagrams:- the PL-, AL- and PW-structures as shown. These basic structures are used in Sections 3.3 and 3.4 in order to produce W_n biformats with odd durations and the *c.p.r.* property.

Figure 3.2.1

(a) Flow diagrams for (i) W_3PL^a (ii) W_4AL^a (iii) W_7PW^a when $p_a = p_b = 0.75$ (against a probability background).



(b) The basic structures (drawn to scale for the case in which $p_a = p_b = 0.75$).



3.3 Biformats with an odd duration and the *c.p.r.* property that can be obtained without using the reversal structure.

Section 2.4 produced a spectrum of W_{2n} biformats with a range of $c.p.r.$'s given by

$$\frac{p_a^t q_b^{2n-t}}{p_b^t q_a^{2n-t}} \qquad (n=1,2,3,\ldots;\ t=1,2,3,\ldots,2n-1)\,. \tag{3.3.1}$$

However, this section shows that this range of biformats can be extended to produce the W_{2n-1} biformats with $c.p.r.$'s given by

$$\frac{p_a^t q_b^{2n-1-t}}{p_b^t q_a^{2n-1-t}} \qquad (n=2,3,4,\ldots;\ t=1,2,\ldots,2n-2)\,. \tag{3.3.2}$$

To see how this can be achieved, suppose we wish to produce a biformat with $c.p.r.$ $\frac{p_a^t q_b^{2n-1-t}}{p_b^t q_a^{2n-1-t}}$ $(t=2,3,4,\ldots,n-1)$. The W_{2n-1} structure is used and it is assumed without loss of generality that the first point is an a-point. Suppose that the PL-structures are used (in an "inner" zone) when the score is between $-(2n-2t-1)$ and $(2n-2t+1)$ and that the AL-structures are used outside this zone. Figure 3.3.1 gives a flow diagram representing this situation. Such a scoring system has the following properties.

1. During any possible movement (or realization) starting from any state and returning to that state, the following are true:-

(a) the number of a-points won by player A equals the number of b-points won by player B.

(b) the number of a-points lost by player A equals the number of b-points lost by player B.

2.

$$n_a - n_b = \begin{cases} 2t-(2n-1) & \text{if player } A \text{ wins} \\ (2n-1)-2t & \text{if player } A \text{ loses} \end{cases}$$

where $n_a(n_b)$ is the total number of a- (b-) points played.

3. The scoring system is "fair". To see this, suppose $p_a = p_b = p$ and $q = 1-p$. Then, the following recurrence relationships can be written down (using the notation that $P(i,a)$ $(P(i,b))$ represents the probability that player A wins when player A is i points ahead and an a-point (b-point) is about to be played).

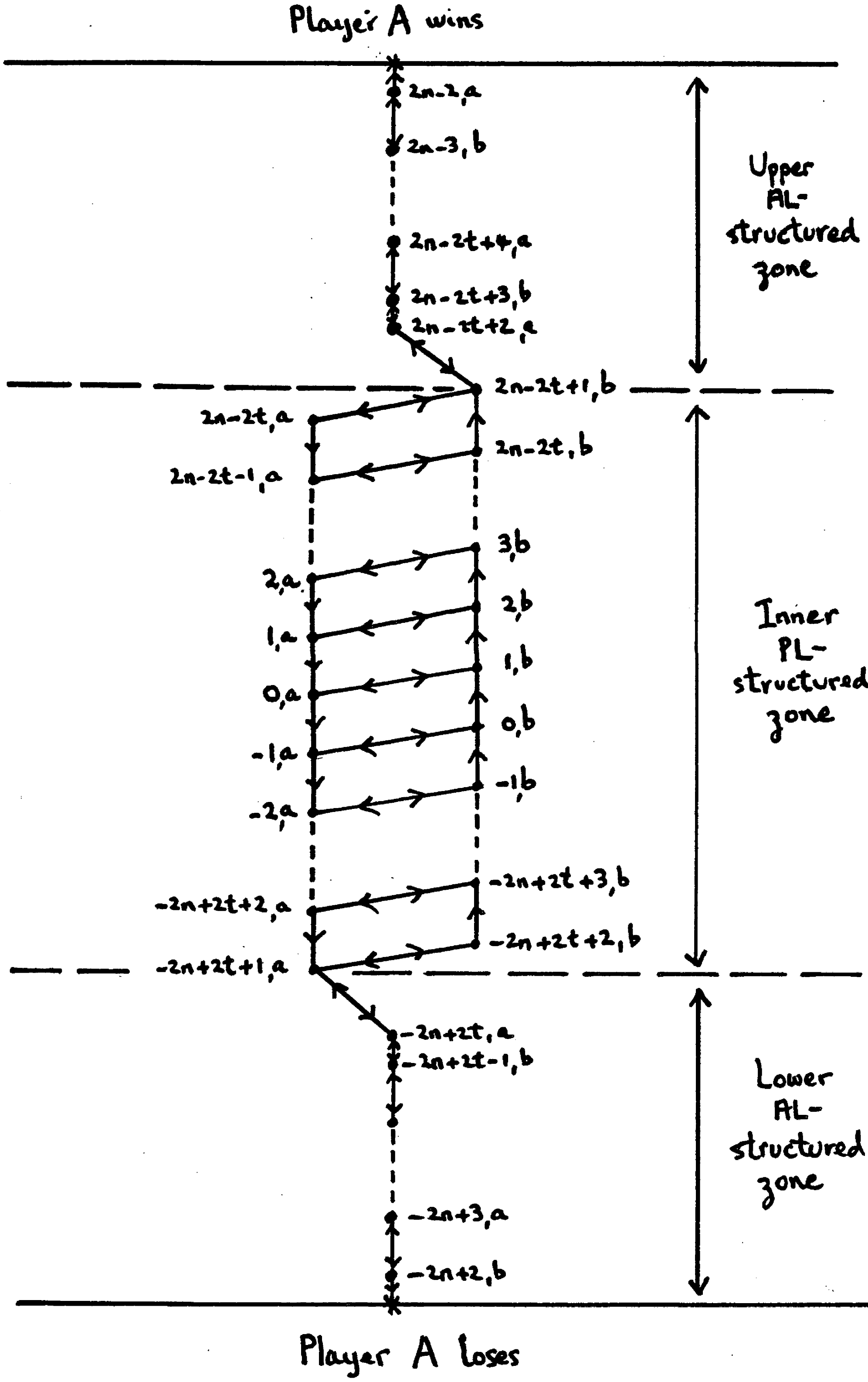
(a) For the PL-structured inner zone,

$$P(i,a) = pP(i+1,b) + qP(i-1,a)$$
$$\text{if } i = -(2n-2t-2), -(2n-2t-3), \ldots, (2n-2t)\,,$$

$$P(i,b) = pP(i-1,a) + qP(i+1,b)$$
$$\text{if } i = -(2n-2t-2), -(2n-2t-3), \ldots, (2n-2t)\,.$$

Figure 3.3.1

A flow diagram representing a W_{2n-1} biformat with $c.p.r.$ $\frac{p_a^t q_b^{2n-1-t}}{p_b^t q_a^{2n-1-t}}$ (drawn to scale for the case in which $p_a = p_b = \frac{2}{3}$; $t = 2, 3, 4, \ldots, n-1$).



(b) For the upper AL-structured zone,

$$P(2n-2t+i,b) = pP(2n-2t+i-1,a) + qP(2n-2t+i+1,a) \quad \text{if } i = 1,3,5,\ldots,2t-3\ ,$$

$$P(2n-2t+i,a) = pP(2n-2t+i+1,b) + qP(2n-2t+i-1,b) \quad \text{if } i = 2,4,6,\ldots,2t-2\ ,$$

$$P(2n-1,b) = 1 \qquad \text{(boundary condition)}\ .$$

(c) For the lower AL-structured zone,

$$P(-(2n-2t-1+i),a) = pP(-(2n-2t-2+i),b) + qP(-(2n-2t+i),b) \quad \text{if } i = 0,2,4,\ldots,2t-2\ ,$$

$$P(-(2n-2t+i),b) = pP(-(2n-2t+i+1),a) + qP(-(2n-2t+i-1),a) \quad \text{if } i = 0,2,4,\ldots,2t-2\ ,$$

$$P(-(2n-1),a) = 0 \qquad \text{(boundary condition)}\ .$$

The solution of these equations is

(a) for the PL-structured inner zone,

$$P(i,a) = \tfrac{1}{2} + \frac{ip}{D} \quad \text{if } i = -(2n-2t-1), -(2n-2t-2), \ldots, (2n-2t)\ ,$$

$$P(i,b) = \tfrac{1}{2} + \frac{q+(i-1)p}{D} \quad \text{if } i = -(2n-2t-2), -(2n-2t-3), \ldots, (2n-2t+1)\ ,$$

where $D = 2t-1+q+(4n-4t-1)p$;

(b) for the upper AL-structured zone,

$$P(2n-2t+i,a) = \tfrac{1}{2} + \frac{q+(2n-2t+1)p+\frac{i-2}{2}}{D} \quad \text{if } i = 2,4,6,\ldots,2t-2\ ,$$

$$P(2n-2t+i,b) = \tfrac{1}{2} + \frac{q+(2n-2t)p+\frac{i-1}{2}}{D} \quad \text{if } i = 1,3,5,\ldots,2t-3\ ;$$

(c) for the lower AL-structured zone,

$$P(-(2n-2t-1+i),a) = \tfrac{1}{2} - \frac{(2n-2t-1)p+\frac{i}{2}}{D} \quad \text{if } i = 0,2,4,\ldots,2t-2\ ,$$

$$P(-(2n-2t+i),b) = \tfrac{1}{2} - \frac{(2n-2t)p+\frac{i}{2}}{D} \quad \text{if } i = 0,2,4,\ldots,2t-2\ .$$

Since we have assumed without loss of generality that the first point is an a-point and since $P(0,a) = \frac{1}{2}$, it follows that the system is "fair".

4. The scoring system is a biformat with the $c.p.r.$ property. To see this, suppose $P(Q)$ is the probability player A wins (loses). From properties 1 and 2, it follows that

$$
\begin{aligned}
P = p_a^t q_b^{2n-1-t}\{&1 + (c_{11}p_a p_b + c_{12}q_a q_b) \\
&+ (c_{21}(p_a p_b)^2 + c_{22}(p_a p_b)(q_a q_b) + c_{23}(q_a q_b)^2) + \ldots\}\,, \\
Q = p_b^t q_a^{2n-1-t}\{&1 + (c'_{11}p_a p_b + c'_{12}q_a q_b) \\
&+ (c'_{21}(p_a p_b)^2 + c'_{22}(p_a p_b)(q_a q_b) + c'_{23}(q_a q_b)^2) + \ldots\}\,.
\end{aligned}
$$

Now, considering the case in which $p_a = p_b = p$ and $q = 1 - p$, we have, using property 3,

$$
\begin{aligned}
0.5 &= p^t q^{2n-1-t}\{1 + (c_{11}p^2 + c_{12}q^2) + (c_{21}p^4 + c_{22}p^2q^2 + c_{23}q^4) + \ldots\}\,, \\
0.5 &= p^t q^{2n-1-t}\{1 + (c'_{11}p^2 + c'_{12}q^2) + (c'_{21}p^4 + c'_{22}p^2q^2 + c'_{23}q^4) + \ldots\}\,.
\end{aligned}
$$

Since these equations are true for all p and q such that $0 < p < 1$ and $p + q = 1$, it follows, as in the proof of Theorem 2.2, that $c_{ij} = c'_{ij}$ for all i and j. Hence, the above scoring system is a biformat with $c.p.r.$

$$
\frac{p_a^t q_b^{2n-1-t}}{p_b^t q_a^{2n-1-t}} \qquad (t = 2, 3, 4, \ldots, n-1)\,.
$$

The case in which $t = 1$ can be produced by slightly modifying the general case represented in Figure 3.3.1. The modifications are that the upper AL-structured zone no longer exists (since the point $(2n - 2t + 1), b$ is now an absorbing point) and there is now only one b-point $(-(2n - 2t), b)$ in the lower AL-structured zone. The above four properties are similarly true for this case, $t = 1$.

As in Corollary 2.2.1, it can be seen from property 4 that the distribution of duration conditional on player A winning, the distribution of duration conditional on player A losing, and the unconditional distribution of duration are identical. It follows in the same way as in the proofs of Theorems 2.3 and 2.5 that, for this family of W_{2n-1} biformats

$$
\frac{P - Q}{\mu} = \frac{p_a - p_b}{2t + (2n - 1 - 2t)(p_a + p_b)} \quad (t = 1, 2, 3, \ldots, n-1)
$$

and, similar to equation (2.4.1) in the proof of Theorem 2.7, the relative efficiency of this family of biformats is approximately

$$
\frac{2\delta}{t + (2n - 1 - 2t)p}\left[t\left\{\frac{\delta}{p} + \frac{1}{3}\left(\frac{\delta}{p}\right)^3 + \ldots\right\} + (2n - 1 - t)\left\{\frac{\delta}{q} + \frac{1}{3}\left(\frac{\delta}{q}\right)^3 + \ldots\right\}\right]
$$

($t = 1, 2, 3, \ldots, n-1$) where $2p = p_a + p_b$ and $q = 1 - p$. In the above expression, the term in δ^2 is $\frac{2\delta^2}{pq}$ for all $t = 1, 2, 3, \ldots, n-1$ whilst the term in δ^4 decreases (increases) monotonically as t increases if $p_a + p_b > 1$ ($p_a + p_b < 1$). The above expression has also been used to show that the W_{2n-1} biformat with *c.p.r.*

$$\frac{p_a {q_b}^{2n-2}}{p_b {q_a}^{2n-2}}$$

(i.e. the full PL case) has an efficiency which lies between the efficiencies of the two neighbouring biformats with even durations, $W_{2n-2}PL$ and $W_{2n}PL$.

Returning to the *c.p.r.*'s given in expression (3.3.2), the cases in which $t = n, n+1, n+2, \ldots, 2n-2$ can be produced by using the PW-structure instead of the PL-structure within the "inner" zone.

The methods of this section have been used to produce biformats with odd durations and the *c.p.r.*'s given in expression (3.3.2). However, these methods could equally have been used to produce the biformats with even durations and the *c.p.r.*'s given in expression (3.3.1).

This section has thus given a procedure for producing the total spectrum of W_n biformats with c.p.r.'s given by

$$\frac{p_a^t q_b^{n-t}}{p_b^t q_a^{n-t}} \quad (n = 2, 3, 4, \ldots;\ t = 1, 2, 3, \ldots, n-1)\ . \tag{3.3.3}$$

The results of this section together with those of Section 2.4 have shown that, for a given n, the efficiencies of the biformats in this total spectrum decrease (increase) as t increases if $p_a + p_b > 1$ ($p_a + p_b < 1$). Thus, for a given n, the case in which $t = 1$ ($t = n - 1$) is the most efficient when $p_a + p_b > 1$ ($p_a + p_b < 1$). Also, for $t = 1$, the efficiencies of the biformats in this total spectrum increase (decrease) as n increases if $p_a + p_b > 1$ ($p_a + p_b < 1$) and, for $t = n - 1$, the efficiencies increase (decrease) as n increases if $p_a + p_b < 1$ ($p_a + p_b > 1$).

3.4 Generalized Reversal Biformats with odd or even durations and the *c.p.r.* property.

This section shows that the PL-, AL- and PW-structures can be "mixed" or "linked together" *in any way* and *c.p.r.*'s achieved as long as certain rules are obeyed. These rules are now given and are represented graphically in Figure 3.4.1.

(i) The lower a- (upper b-) point of a PL-structure can also be the upper a- (lower b-) point of the neighbouring PW-structure (see Figure 3.4.1(i)).

(ii) The upper a- (lower b-) point of an AL-structure can also be the lower a- (upper b-) point of a PL-structure (see Figure 3.4.1(ii)).

(iii) The lower a- (upper b-) point of an AL-structure can also be the upper a- (lower b-) point of a PW-structure (see Figure 3.4.1(iii)).

Any diagram can be drawn linking or mixing the PL-, AL- and PW-structures in any of the above 6 ways shown in Figure 3.4.1; upper and lower boundaries (i.e. A winning and A losing boundaries) can be drawn at the upper and lower points in the diagram thus forming a scoring system, S. Any point ((say) an a-point) in the diagram can be selected as the starting point and suppose n_1 (n_2) is the minimum number of points in which player A (player B) can win. S is thus a $W_{n_2}^{n_1}$ system (player A has to reach n_1 points ahead of player B before player B gets n_2 points ahead of player A). The scoring system S_π can be obtained by rotating S through 180° about its "midpoint" (midway between the boundaries). Every a-point in S becomes a b-point in S_π which is a $W_{n_1}^{n_2}$ system. It follows that $W_1(W_{n_2}^{n_1}S^a,\ W_{n_1}^{n_2}S_\pi^b)$ is in fact a reversal biformat. It can be seen that properties corresponding to those of properties 1 and 2 of Section 3.3 hold for this reversal biformat. Also, since these properties are true and since any winning path to absorption has a mirror-image losing path to absorption (by making the exchanges $A \longleftrightarrow B$ and $S \longleftrightarrow S_\pi$), this biformat has the $c.p.r.$ property and is "fair". It can also be seen that, for any given structure S, different starting points will produce, using the above reversal procedure, a generalized reversal biformat with the same $c.p.r.$. Figure 3.4.2 gives an example in which $n_1 = 16$, $n_2 = 9$ and the $c.p.r.$ equals $\frac{p_a^{12}q_b^{13}}{p_b^{12}q_a^{13}}$.

It can be seen that the generalized reversal biformats discussed in this section can be used to produce the total spectrum of W_n biformats with c.p.r.'s given by expression (3.3.3). Thus, we have an alternative procedure to that of Section 3.3 for producing the total spectrum.

3.5 Some Further Comments and Examples

The methods of the two previous sections involve the construction of a scoring system in which the state at any stage and the outcome at that stage together determine the next type of point to be played. This can be compared with the simpler PL and PW rules in which the outcome of any point alone determines the next type of point to be played. Some further examples are now stated.

Suppose the notation $PL_{odd}PW_{even}$ is used to denote the exchange mechanism in which, using the structure of Section 2.3, $\binom{A}{A}_1, \binom{A}{A}_3, \binom{A}{A}_5, \ldots$ of the

Figure 3.4.1

The 6 possible methods of "mixing" or "linking together" the PL-, AL- and PW-structures (drawn to scale for the case in which $p_a = p_b = 2/3$).

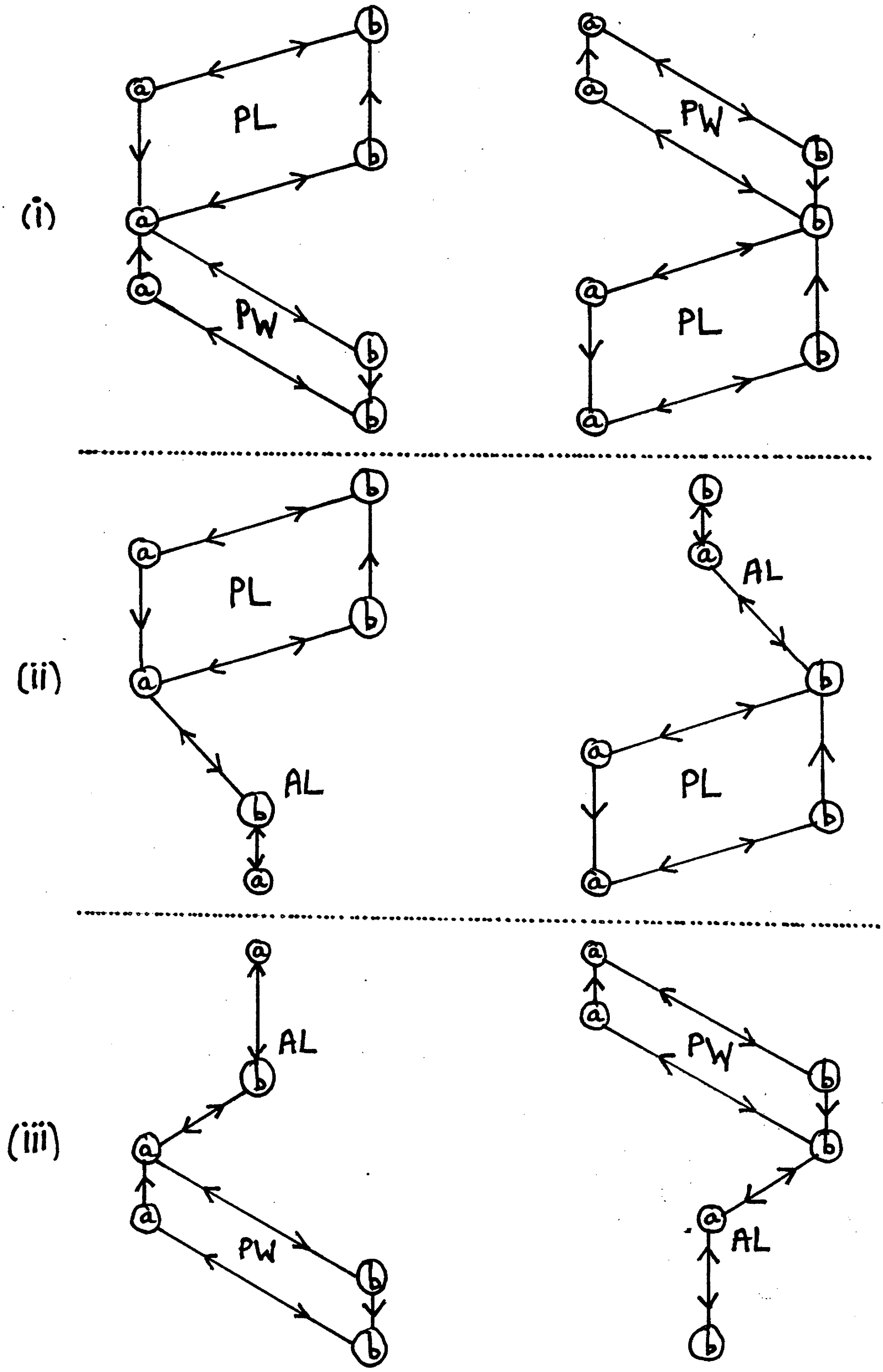
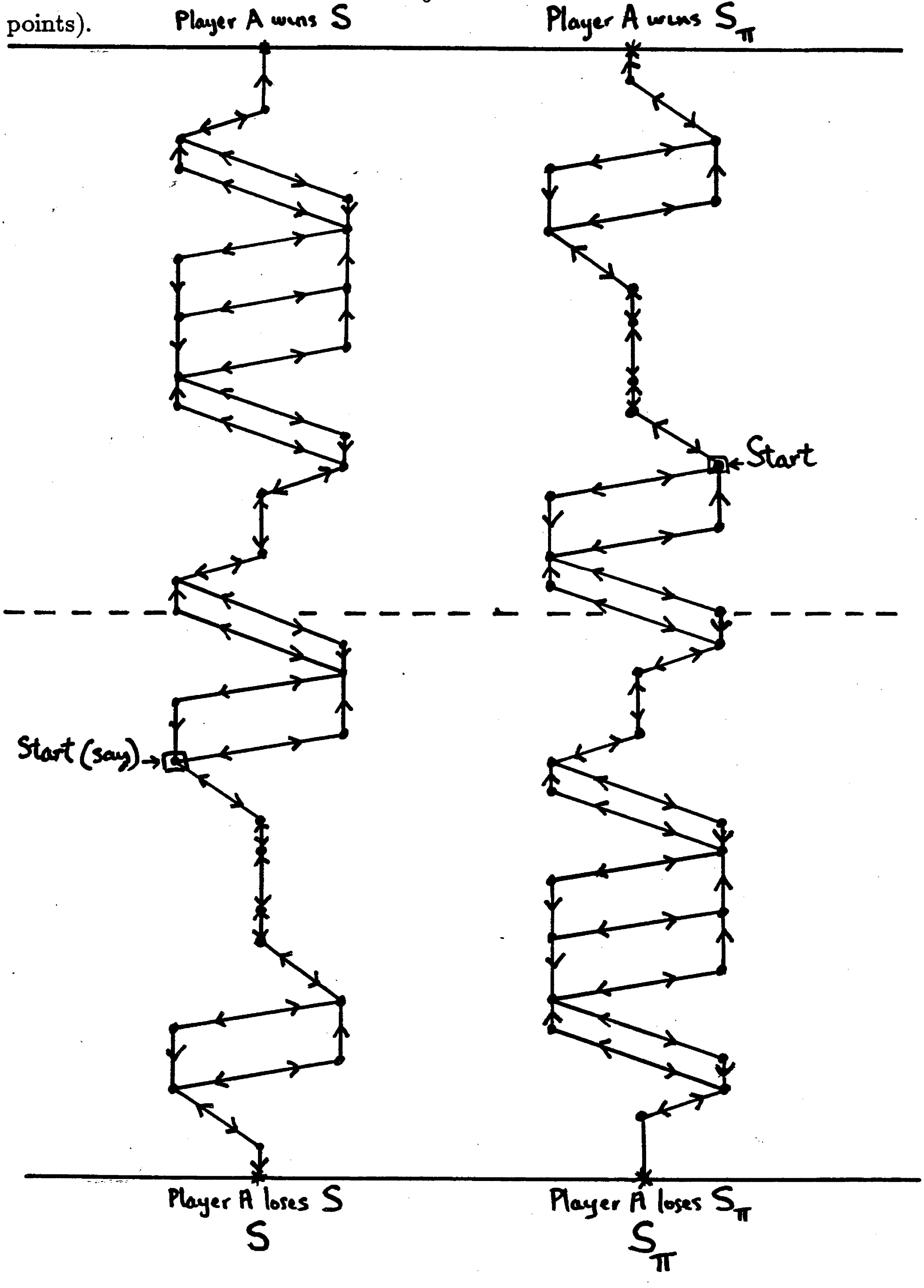


Figure 3.4.2

An example of two scoring systems, S and S_π, that can be used to produce the generalized reversal biformat, $W_1(W_{n_2}^{n_1}S,\, W_{n_1}^{n_2}S_\pi)$, with $c.p.r.$ $\frac{p_a^{12}q_b^{13}}{p_b^{12}q_a^{13}}$ (drawn to scale for the case in which $p_a = p_b = \frac{2}{3}$; $n_1 = 16$ and $n_2 = 9$ for the given starting points).



$\binom{A}{A}$ subsequence are followed by $\binom{b}{b}$ point-pairs whilst $\binom{A}{A}_2, \binom{A}{A}_4, \binom{A}{A}_6, \ldots$ are followed by $\binom{a}{a}$ point-pairs and $\binom{B}{B}_1, \binom{B}{B}_3, \binom{B}{B}_5, \ldots$ of the $\binom{B}{B}$ subsequence are followed by $\binom{a}{a}$ point-pairs whilst $\binom{B}{B}_2, \binom{B}{B}_4, \binom{B}{B}_6, \ldots$ are followed by $\binom{b}{b}$ point-pairs. All $\binom{A}{B}$ outcomes are followed by $\binom{a}{b}$ point-pairs. Using this notation for the point-pairs case and a corresponding one for the single-points case, it can be seen that the two biformats $W_n(PL_{\text{odd}}\, PW_{\substack{\text{even}\\ \text{or}\\ \text{odd}}}\, g.p.p.)$ and the four reversal biformats $W_1(W_n(PL_{\substack{\text{odd}\\ \text{or}\\ \text{even}}}\, PW^a_{\substack{\text{even}\\ \text{or}\\ \text{odd}}}), W_n(PL_{\substack{\text{odd}\\ \text{or}\\ \text{even}}}\, PW^b_{\substack{\text{even}\\ \text{or}\\ \text{odd}}}))$ have, for n odd, the same distribution of duration as W_n (point-pairs).

The example in the previous paragraph can be extended in the same way as Section 2.4 extended Section 2.3. Suppose the notation $PL_{l_1}\, PW_{l_2}$ is used to denote the exchange mechanism in which, using the structure of Section 2.4 and modulus $n-1$ arithmetic, $\binom{A}{A}$ wins numbered $i_1, i_2, i_3, \ldots, i_{l_1}$ (for any non-equal i_j from the set $\{0, 1, 2, \ldots, n-2\}$) in the $\binom{A}{A}$ subsequence are followed by $\binom{b}{b}$ point-pairs whilst $\binom{A}{A}$ wins numbered $k_1, k_2, k_3, \ldots, k_{l_2}$ (for any non-equal k_j from the set $\{0, 1, 2, \ldots, n-2\}$ and such that no k's and i's are equal) are followed by $\binom{a}{a}$ point-pairs. All other $\binom{A}{A}$ wins are followed by $\binom{a}{b}$ point-pairs. The corresponding discipline (with the same i's and k's) is used for the $\binom{B}{B}$ subsequence. Then, using a similar notation for the single-points case (and alternating points when the PL or PW rule is not used, as in Sections 2.3 and 2.4), $W_n(PL_{l_1} PW_{l_2}\, g.p.p.)$ has the same distribution of duration as $W_1(W_n PL_{l_1} PW^a_{l_2}, W_n PL_{l_1} PW^b_{l_2})$, for example. Also, $W_n(PL_{l_1} PW_{l_2}\, g.p.p.)$ has the same distribution of duration as $W_n(PL_{l_1-l_2}\, g.p.p.)$ if $l_1 > l_2$, $W_n(PW_{l_2-l_1}\, g.p.p.)$ if $l_1 < l_2$ whilst, if $l_1 = l_2$, it has the same distribution as W_n(point- pairs).

3.6 A Theorem on PL and PW Biformats

It appears that no advantage (in terms of efficiency) can be made of these PL and PW mechanisms if it is not known, *a priori*, whether $p_a + p_b$ is greater or less than unity, as can be seen from the following Theorem 3.1. Thus this theorem suggests that it is not possible to find a biformat which is super-efficient in both the regions $p_a + p_b > 1$ and $p_a + p_b < 1$.

Theorem 3.1. *The efficiency of $W_1(W_{2n}PW, W_{2n}PL)$ is strictly less than unity unless $p_a + p_b = 1$ or $p_a = p_b$.*

Proof: From Theorems 2.1 and 2.3, it follows that, for $W_{2n}PL$,

$$\begin{aligned} \frac{P_{PL}}{Q_{PL}} &= \frac{p_a q_b^{2n-1}}{p_b q_a^{2n-1}} \\ \text{and} \quad \mu_{PL} &= 2\left(\frac{p_a q_b^{2n-1} - p_b q_a^{2n-1}}{p_a q_b^{2n-1} + p_b q_a^{2n-1}}\right)\left(\frac{1+(n-1)(p_a+p_b)}{p_a - p_b}\right) \end{aligned} \tag{3.6.1}$$

where P_{PL} is the probability player A wins $W_{2n}PL$, μ_{PL} is the expected duration of $W_{2n}PL$ and $Q_{PL} = 1 - P_{PL}$. Using a corresponding notation for $W_{2n}PW$, we have

$$\begin{aligned} \frac{P_{PW}}{Q_{PW}} &= \frac{p_a^{2n-1} q_b}{p_b^{2n-1} q_a} \\ \text{and} \quad \mu_{PW} &= 2\left(\frac{p_a^{2n-1} q_b - p_b^{2n-1} q_a}{p_a^{2n-1} q_b + p_b^{2n-1} q_a}\right)\left(\frac{1+(n-1)(q_a+q_b)}{p_a - p_b}\right). \end{aligned} \tag{3.6.2}$$

Now, for the complete biformat, $W_1(W_{2n}PW, W_{2n}PL)$, it follows that

$$P = \frac{P_{PW}P_{PL}}{P_{PW}P_{PL} + Q_{PW}Q_{PL}} \tag{3.6.3}$$

where P is the probability player A wins the complete biformat. Now recalling the result that, for each section of the complete biformat, the expected duration, the expected duration conditional on player A winning and the expected duration conditional on player A losing are all equal, it follows that

$$\mu = \frac{\mu_{PW} + \mu_{PL}}{1 - (P_{PW}Q_{PL} + Q_{PW}P_{PL})} \tag{3.6.4}$$

where μ is the expected duration of the complete biformat. Using equations (3.6.1) and (3.6.2) in (3.6.3) and (3.6.4), we have

$$\begin{aligned} \frac{P}{Q} &= \left(\frac{p_a q_b}{p_b q_a}\right)^{2n} \\ P - Q &= \frac{(p_a q_b)^{2n} - (p_b q_a)^{2n}}{(p_a q_b)^{2n} + (p_b q_a)^{2n}} \end{aligned}$$

and

$$\mu = \frac{4n\big((p_aq_b)^{2n} - (p_bq_a)^{2n}\big) + 4p_ap_bq_aq_b(n-1)(p_a+p_b-1)\big((p_bq_b)^{2n-2} - (p_aq_a)^{2n-2}\big)}{(p_a-p_b)\big((p_aq_b)^{2n} + (p_bq_a)^{2n}\big)}$$

where $Q = 1 - P$. Thus, the efficiency, ρ, of the complete biformat is given by

$$\begin{aligned}\rho &= \frac{2(P-Q)\ln\left(\frac{P}{Q}\right)}{\mu(p_a-p_b)\ln\left(\frac{p_aq_b}{p_bq_a}\right)}\\ &= \left\{1 + \frac{n-1}{n}(p_ap_bq_aq_b)(p_a+p_b-1)\left(\frac{(p_bq_b)^{2n-2} - (p_aq_a)^{2n-2}}{(p_aq_b)^{2n} - (p_bq_a)^{2n}}\right)\right\}^{-1}.\end{aligned}$$

This expression follows upon substituting the three previous expressions into the formula for ρ. By considering the eight regions within the (p_a, p_b) unit square formed by the lines $p_a = p_b$, $p_a + p_b = 1$, $p_a = 1/2$ and $p_b = 1/2$, it follows that ρ is strictly less than 1 unless $p_a + p_b = 1$ or $p_a = p_b$, thus completing the proof.

3.7 The Distribution of Duration of Some W_N Systems.

This section *outlines* some methods which can be used to obtain the *distribution* of duration of some W_n systems and can be omitted on first reading since the following chapters are not dependent on the results within it.

3.7.1. The W_n uniformat.

Uniformats were introduced in Section 1.2. The distributions of duration of the uniformats $\{W_n\}$ $(n = 1, 2, 3, \ldots)$ are particular cases of those of the $W_{n_2}^{n_1}$ unipoints systems $(n_1 = 1, 2, 3, \ldots; n_2 = 1, 2, 3, \ldots)$ and are available (Feller, 1957, p.318–323). In this section however, an alternative approach is taken, which is suitable, when appropriately expanded, to bipoints cases. Considering the uniformat W_n, define the matrix $\underline{\underline{M}}$, which effectively considers the movement within the uniformat after every two points, by

$$\underset{(n\times n)}{\underline{\underline{M}}} = \begin{bmatrix} 1 & 1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & 2 & 1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & 1 & 2 & 1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & 1 & 2 & 1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ & & & & & & & & & & \\ & & & & & \text{etc.} & & & & & \\ & & & & & & & & & & \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1 & 2 & 1 & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1 & 2 & 1 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1 & 1 \end{bmatrix}$$

where dots represent zeros. $\underline{\underline{M}}$ is a transition matrix in which row i ($i = 1, 2, 3, \ldots, n$) represents movement out of state $n - 2i + 1$ where state j is that state in which player A is j points ahead of player B. Column k refers to the same state as row k. Initializing column vectors $\mathbf{v}_0$ and $\mathbf{v}_1$ (to cover the cases in which n is odd and even respectively) are defined by

$$\underset{(n\times 1)}{\mathbf{v}_0} = \begin{array}{l} \left[\begin{array}{c} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{array}\right] \begin{array}{l} \\ \frac{n-1}{2} \text{ zeros} \\ \\ \\ \\ \\ \\ \frac{n-1}{2} \text{ zeros} \\ \\ \end{array} \end{array} \qquad \text{(n odd)}$$

and

$$\underset{(n\times 1)}{\mathbf{v}_1} = \begin{array}{l} \left[\begin{array}{c} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{array}\right] \begin{array}{l} \\ \\ \frac{n}{2} - 1 \text{ zeros} \\ \\ \\ \\ \\ \frac{n}{2} - 1 \text{ zeros} \\ \\ \\ \end{array} \end{array} \qquad \text{(n even)}\,.$$

For $l = 1, 2, 3, \ldots$ the column vectors $\mathbf{v}_{2l+i}$ are defined by

$$\mathbf{v}_{2l+i} = \underline{\underline{M}}^l \mathbf{v}_i \qquad (i = 0, 1)$$

and, for all n (odd or even), it can be seen that

$$P\left\{\begin{array}{c}\text{Player } A \\ \text{wins the} \\ \text{uniformat} \\ W_n \text{ in } n+2j \\ \text{points}\end{array}\right\} = \mathbf{v}_{n-1+2j}^T \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} p^{n+j}(1-p)^j \quad (j = 0, 1, 2, \ldots) \qquad (3.7.1)$$

where p is the probability player A wins a single point. Thus, explicit expressions giving the distribution of duration conditional on player A winning can be evaluated. Recalling that this uniformat has the $c.p.r.$ property, the distribution conditonal on player A losing and the unconditional distributions are the same.

3.7.2 The W_nAL^a system.

The biformats $W_1(W_nAL^a,\ W_nAL^b)$ and $W_{2n}AL^{a \text{ or } b}$ have the same distribution of duration as W_n (point-pairs), which is available (Cox and Miller, 1964, p.25–33). We now consider a case in which the duration must be odd — say W_nAL^a (n odd). Since n is odd, one more a-point than b-point must have been played at any stage of absorption and hence it follows that, for n odd,

$$P\left\{\begin{matrix}\text{Player } A \text{ wins } W_nAL^a \\ \text{in } n+2j \text{ points}\end{matrix}\right\} = p_a^{\frac{n+1}{2}} q_b^{\frac{n-1}{2}} H_j(p_a p_b,\ q_a q_b)$$

where $H_j(p_a p_b,\ q_a q_b)$ is a homogeneous polynomial of degree j in $p_a p_b$ and $q_a q_b$. *The coefficients of the terms in $H_j(p_a p_b, q_a q_b)$ can be evaluated using the following vector and matrix procedure.* An extension, $\underline{\underline{N}}_1$, of the matrix $\underline{\underline{M}}$ (in which each element in $\underline{\underline{M}}$ gives rise to a sub-matrix in $\underline{\underline{N}}_1$) is defined by

$$\underset{((n(l+1)) \times (n(l+1)))}{\underline{\underline{N}}_1} = \begin{bmatrix} \underline{\underline{1}} & \underline{\underline{1}} & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \underline{\underline{D}} & \underline{\underline{1}}+\underline{\underline{D}} & \underline{\underline{1}} & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \underline{\underline{D}} & \underline{\underline{1}}+\underline{\underline{D}} & \underline{\underline{1}} & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \underline{\underline{D}} & \underline{\underline{1}}+\underline{\underline{D}} & \underline{\underline{1}} & \cdot & \cdot & \cdot & \cdot \\ & & & & \text{etc.} & & & & \\ \cdot & \cdot & \cdot & \cdot & \cdot & \underline{\underline{D}} & \underline{\underline{1}}+\underline{\underline{D}} & \underline{\underline{1}} & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \underline{\underline{D}} & \underline{\underline{1}}+\underline{\underline{D}} & \underline{\underline{1}} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \underline{\underline{D}} & \underline{\underline{D}} \end{bmatrix}$$

where

$$\underset{((l+1) \times (l+1))}{\underline{\underline{1}}} = \begin{bmatrix} 1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & 1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & 1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ & & & & \text{etc.} & & & & \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1 & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1 \end{bmatrix} \quad \text{and}$$

$$\underset{((l+1)\times(l+1))}{\underline{\underline{D}}} = \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & 1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & 1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ & & & & \text{etc.} & & & & \\ \cdot & \cdot & \cdot & \cdot & \cdot & 1 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1 & \cdot \end{bmatrix}.$$

Again dots represent zeros except in the case of the matrix $\underline{\underline{N}}_1$ in which case they represent sub-matrices of zeros. The two submatrices $\underline{\underline{1}}$ and $\underline{\underline{D}}$ are required in order to separate the two possibilities win-loss or loss-win (with probabilities $p_a p_b$ and $q_a q_b$) on two consecutive points. The integer l is selected so that it is large enough to give the number of required elements in the distribution of duration given by equation (3.7.2). Since we are considering the case in which n is odd, the initializing column vector, $\mathbf{w}_0$, is defined by

$$\underset{(n(l+1)\times 1)}{\mathbf{w}_0} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \\ \mathbf{1} \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{bmatrix} \begin{matrix} \\ \frac{n-1}{2} \text{ zero vectors} \\ \\ \\ \\ \\ \frac{n-1}{2} \text{ zero vectors} \\ \\ \end{matrix}$$

where

$$\underset{((l+1)\times 1)}{\mathbf{1}} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad \text{and} \quad \underset{((l+1)\times 1)}{\mathbf{0}} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

For $l = 1, 2, 3, \ldots$, the column vector $\mathbf{w}_{2l}$ is defined by

$$\mathbf{w}_{2l} = \underline{\underline{N}}_1^l \mathbf{w}_0$$

and it can be seen that, for n odd,

$$P\left\{\begin{array}{c}\text{Player } A \text{ wins}\\ W_nAL^a \text{ in}\\ n+2j \text{ points}\end{array}\right\} = \mathbf{w}_{n-1+2j}^T \begin{pmatrix} \mathbf{v}_{l,j} \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} p_a^{\frac{n+1}{2}} q_b^{\frac{n-1}{2}} \qquad j = 0,1,2,\ldots,l \tag{3.7.2}$$

where

$$\mathbf{v}_{l,j} = \begin{bmatrix} (q_a q_b)^j \\ (p_a p_b)(q_a q_b)^{j-1} \\ (p_a p_b)^2 (q_a q_b)^{j-2} \\ \vdots \\ (p_a p_b)^j \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \begin{array}{l} (j+1) \text{ non-zero terms} \\ \\ \\ \\ \\ \\ \\ l-j \text{ zeros} \\ \\ \end{array}.$$

Thus equation (3.7.2) gives explicit expressions for as many terms as are required (by selecting l large enough) in the distribution of duration along player A's winning boundary. The distribution of duration along player A's losing boundary can be obtained by modifying the column vector in equation (3.7.2). An unfortunate characteristic of the above solution (and perhaps an unresolvable problem as each coefficient within the homogeneous polynomials needs to be "kept track of") is that the dimensions of the matrix $\underline{\underline{N}}_1$ need to be increased as more and more terms in the distribution (3.7.2) are required. (In fact, the value of l must be at least the number of terms required in the distribution (3.7.2).)

3.7.3. The W_nPL^a system.

A corresponding approach to that of Section 3.7.2 can be used to find the distribution of duration of W_nPL^a (n odd or even). Using the sub-matrices and sub-vectors defined in Section 3.7.2, the square matrix, $\underline{\underline{N}}_2$, of size $2(n-1)(l+1)$, is defined by

$$\begin{bmatrix}
\underline{\underline{D}} & \underline{\underline{1}} & \underline{\underline{1}} & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\underline{\underline{D}} & \underline{\underline{D}} & \underline{\underline{D}} & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \underline{\underline{1}} & \underline{\underline{D}} & \underline{\underline{1}} & \underline{\underline{1}} & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \underline{\underline{1}} & \underline{\underline{D}} & \underline{\underline{D}} & \underline{\underline{D}} & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \underline{\underline{1}} & \underline{\underline{D}} & \underline{\underline{1}} & \underline{\underline{1}} & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \underline{\underline{1}} & \underline{\underline{D}} & \underline{\underline{D}} & \underline{\underline{D}} & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \underline{\underline{1}} & \underline{\underline{D}} & \underline{\underline{1}} & \underline{\underline{1}} & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \underline{\underline{1}} & \underline{\underline{D}} & \underline{\underline{D}} & \underline{\underline{D}} & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
 & & & & & & & & & \text{etc.} & & & & & & \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \underline{\underline{1}} & \underline{\underline{D}} & \underline{\underline{1}} & \underline{\underline{1}} & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \underline{\underline{1}} & \underline{\underline{D}} & \underline{\underline{D}} & \underline{\underline{D}} & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \underline{\underline{1}} & \underline{\underline{D}} & \underline{\underline{1}} \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \underline{\underline{1}} & \underline{\underline{D}} & \underline{\underline{D}}
\end{bmatrix}$$

$\underline{\underline{N}}_2$ is a transition matrix in which the ith row of sub-matrices ($i = 1, 3, 5, \ldots, 2n-3$) represents movement out of an a-state in which player A is $n-1-i$ points ahead of player B. The ith row of sub-matrices ($i = 2, 4, 6, \ldots, 2n-2$) represents movement out of a b-state in which player A is $n-i$ points ahead of player B. Again the kth column of sub-matrices refers to the same state as the kth row of sub-matrices. The initializing column vectors $\mathbf{x}_0$ and $\mathbf{x}_1$ (to cover the cases in which n is odd and even respectively) are defined by

$$\underset{(2(n-1)(l+1)\times 1)}{\mathbf{x}_0} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \\ \mathbf{1} \\ \mathbf{0} \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{bmatrix} \begin{matrix} \\ (n-2) \text{ zero sub-vectors} \\ \\ \\ \\ \\ \\ (n-1) \text{ zero sub-vectors} \\ \\ \end{matrix} \qquad (n \text{ odd})$$

and

$$\underset{(2(n-1)(l+1)\times 1)}{\mathbf{x}_1} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \\ \mathbf{1} \\ \mathbf{1} \\ \mathbf{0} \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{bmatrix} \begin{array}{l} \\ (n-2)\text{ zero sub-vectors} \\ \\ \\ \\ \\ \\ (n-2)\text{ zero sub-vectors} \\ \\ \end{array} \qquad (n \text{ even}) .$$

For $l = 1, 2, 3, \ldots$, the column vector $\mathbf{x}_{2l+i}$ is defined by

$$\mathbf{x}_{2l+i} = \underline{\underline{N}}_2^l \mathbf{x}_i \qquad (i = 0, 1)$$

and, for all n (odd or even), it can be seen that

$$P\left\{\begin{array}{c}\text{Player } A \\ \text{wins } W_nPL^a \\ \text{in } n+2j \\ \text{points}\end{array}\right\} = \mathbf{x}^T_{n-1+2j} \begin{pmatrix} \mathbf{v}_{l,j} \\ \mathbf{0} \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{pmatrix} p_a {q_b}^{n-1} \quad (j = 0, 1, 2, \ldots, l) . \qquad (3.7.3)$$

Thus, as in Sections 3.7.1 and 3.7.2, the distributions of duration when using the PL rule can be evaluated. Note that the dimensions of the matrix $\underline{\underline{N}}_2$ also need to be increased as more and more terms in the distribution (3.7.3) are required.

3.7.4. Some further comments.

(i) The distributions of duration along the losing boundaries can be evaluated by modifying the column vectors in equations (3.7.1) to (3.7.3).

(ii) The probability of any score between $-n$ and $+n$ (not just $-n$ or $+n$) at any stage can be evaluated similarly. Essentially, this is done by modifying the column vectors in equations (3.7.1) to (3.7.3).

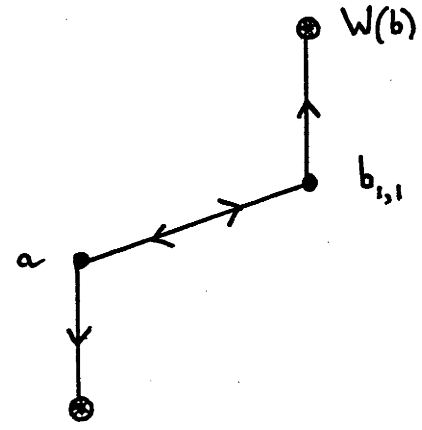
(iii) Similar methods can be used to obtain the distributions of duration of the systems $W_{2n}AL^{a \text{ or } b}$, W_nAL^b, W_nPL^b, $W_n(PL\, g.p.p.)$ and corresponding systems using the PW rule.

(iv) The unsymmetric cases $W^{n_1}_{n_2}$ can also be obtained by modifying the initializing vectors.

3.7.5. An example of the probability generating function for the duration: the W_nPL^b system.

A stagewise approach to finding the probability generating function (p.g.f.) of the duration along player A's winning boundary is now given. Recalling the PL-structure as given in the flow diagram of Figure 3.2.1(b), p.g.f.'s of the duration from "the current b-point until the absorbing (A-winning) b-point" can be generated in stages.

Stage 1: The diagram for this stage is:-



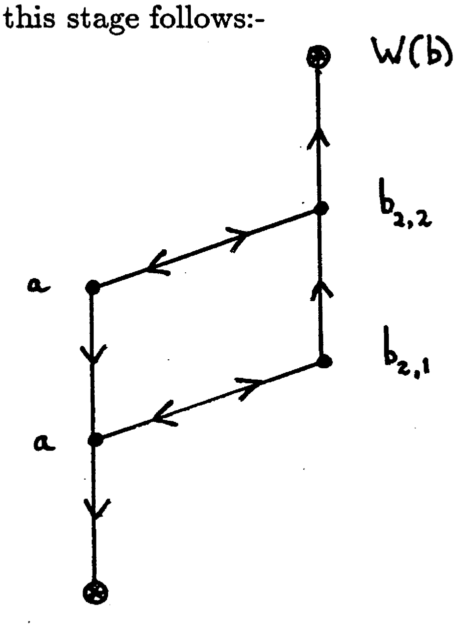
In this diagram $b_{1,1}$ is the only non-absorbing b-state and p.g.f.$\cdot_{1,1}$ is the p.g.f. of the duration starting from state $b_{1,1}$ until absorption in the (A-winning) state, $W(b)$. Thus,

$$\text{p.g.f.}_{\cdot 1,1} = q_b s + p_a p_b s^2 \ \text{p.g.f.}_{\cdot 1,1}$$

where the p.g.f.'s are given as functions of s. Hence,

$$\text{p.g.f.}_{\cdot 1,1} = \frac{q_b s}{D_1} \quad \text{where} \quad D_1 = 1 - p_a p_b s^2 \ .$$

Stage 2: The diagram for this stage follows:-



If p.g.f.$_{\cdot i,j}$ is the p.g.f. (for stage i) of the duration starting from state $b_{i,j}$ until absorption in the (A-winning) state, $W(b)$, we have

$$\text{p.g.f.}_{\cdot 2,1} = q_b s\,\text{p.g.f.}_{\cdot 2,2} + p_a p_b s^2\,\text{p.g.f.}_{\cdot 2,1}$$

and

$$\text{p.g.f.}_{\cdot 2,2} = q_b s + p_a p_b s^2\,\text{p.g.f.}_{\cdot 2,2} + p_a p_b q_a s^3\,\text{p.g.f.}_{\cdot 2,1}\,.$$

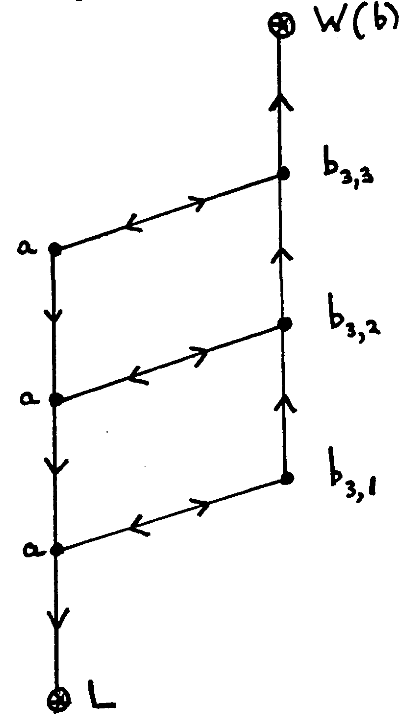
Hence,

$$\text{p.g.f.}_{\cdot 2,1} = \frac{q_b s\,\text{p.g.f.}_{\cdot 2,2}}{D_1}$$

and

$$\text{p.g.f.}_{\cdot 2,2} = \frac{q_b s}{D_2}$$

where $D_2 = 1 - p_a p_b s^2 - \frac{p_a p_b q_a q_b s^4}{1 - p_a p_b s^2}$.
Stage 3: The diagram for this stage is:-



Using a corresponding notation, it can similarly be shown that

$$\text{p.g.f.}_{\cdot 3,1} = \frac{q_b s\,\text{p.g.f.}_{\cdot 3,2}}{D_1}$$
$$\text{p.g.f.}_{\cdot 3,2} = \frac{q_b s\,\text{p.g.f.}_{\cdot 3,3}}{D_2}$$

and

$$\text{p.g.f.}_{3,3} = \frac{q_b s}{D_3}$$

where $D_3 = 1 - p_a p_b s^2 - \frac{p_a p_b q_a q_b s^4}{D_2} - \frac{p_a p_b (q_a q_b)^2 s^6}{D_2 . D_1}$.

Continuing in this fashion and defining for $n = 3, 4, 5, \dots$ and $j = 1, 2, \dots, n$

$$E_{n,j} = D_n . D_{n-1} . D_{n-2} \dots D_{n-j+1}$$

and

$$D_{n+1} = 1 - p_a p_b s^2 - \sum_{j=1}^{n} \frac{(p_a p_b)(q_a q_b)^j s^{2(j+1)}}{E_{n,j}}$$

it follows that the distribution of duration of $W_n PL^b$ has a p.g.f. along player A's winning boundary which is given by

$$\text{p.g.f.}_{2n-2,n-1} = \frac{(q_b s)^n}{D_{n-1} D_n D_{n+1} \dots D_{2n-2}} \, .$$

3.7.6. Explicit results for $W_2 PL^a$.

An enumeration of all possible outcomes has given the following explicit results for $W_2 PL^a$.

$$P \left\{ \begin{array}{l} \text{Player } A \text{ wins } W_2 PL^a \\ \text{in } 4m+2 \text{ points} \end{array} \right\} = (p_a p_b)^m p_a q_b \sum_{i=0}^{m} \binom{2m+1}{2i} (p_a p_b)^{m-i} (q_a q_b)^i$$

$$m = 0, 1, 2, \dots$$

$$P \left\{ \begin{array}{l} \text{Player } A \text{ wins } W_2 PL^a \\ \text{in } 4m \text{ points} \end{array} \right\} = (p_a p_b)^{m-1} p_a q_b \sum_{i=0}^{m} \binom{2m}{2i} (p_a p_b)^{m-i} (q_a q_b)^i$$

$$m = 1, 2, 3, \dots$$

$$P \left\{ \begin{array}{l} \text{Player } A \text{ loses } W_2 PL^a \\ \text{in } 4m \text{ points} \end{array} \right\} = (p_a p_b)^m {q_a}^2 \sum_{i=0}^{m-1} \binom{2m}{2i+1} (p_a p_b)^{m-1-i} (q_a q_b)^i$$

$$m = 1, 2, 3, \dots$$

$$P \left\{ \begin{array}{l} \text{Player } A \text{ loses } W_2 PL^a \\ \text{in } 4m+2 \text{ points} \end{array} \right\} = (p_a p_b)^m {q_a}^2 \sum_{i=0}^{m} \binom{2m+1}{2i+1} (p_a p_b)^{m-i} (q_a q_b)^i$$

$$m = 0, 1, 2, \dots \, .$$

Explicit expressions for $W_2 PL^b$ follow from the above.

3.8 Summary

This chapter has extended the spectrum of biformats devised in Chapter 2 so as to include the biformats with odd durations, thus forming the total spectrum. Two main methods (permutations and other variations within each of these methods is possible), one using the reversal structure and one not, of producing the total spectrum of W_n biformats with *c.p.r.*'s given by

$$\frac{p_a^t q_b^{n-t}}{p_b^t q_a^{n-t}} \qquad (n = 2, 3, 4, \ldots;\ t = 1, 2, 3, \ldots, n-1)$$

have been described. For a given n, the efficiencies of the biformats in this total spectrum decrease (increase) as t increases if $p_a + p_b > 1$ ($p_a + p_b < 1$). Thus, for a given n, the case in which $t = 1$ ($t = n - 1$) is the most efficient if $p_a + p_b > 1$ ($p_a + p_b < 1$). Also, for $t = 1$, the efficiencies of the biformats in this total spectrum increase (decrease) as n increases if $p_a + p_b > 1$ ($p_a + p_b < 1$) and, for $t = n - 1$, the efficiencies increase (decrease) as n increases if $p_a + p_b < 1$ ($p_a + p_b > 1$).

Rules for "mixing" or "linking together" PL-, AL- and PW-structures and a generalized reversal structure have been outlined. It appears that no advantage (in terms of efficiency) can be made of these PL and PW rules if it is not known, *a priori*, whether $p_a + p_b$ is greater or less than unity. Finally, methods that can be used to obtain the *distribution* of duration of some of these W_n systems have been *outlined.*

CHAPTER 4
A RELATIONSHIP BETWEEN THE IMPORTANCES OF THE POINTS WITHIN A BIFORMAT AND THE EFFICIENCY OF THAT BIFORMAT

4.1 Introduction.

Morris (1977) had noted that points in a game of tennis are not equally important and Miles (1984) had noted the inefficiency of the tennis scoring system. Could it be that the differing importances of the points within the tennis scoring system are contributing to its inefficiency? Alternatively, is there a relationship between the efficiency of a biformat and the importances of the points within it? This chapter addresses this question. Firstly however, a diagram called the $P-D-I$ diagram, which was found to be quite useful in the analysis of scoring systems, is introduced in Section 4.2. This section also shows that every biformat within the total spectrum of W_n biformats has the following property: each point within that biformat is equally important when $p_a = p_b$. In Section 4.3 a relationship between the increased probability of winning each point within a biformat and the increased overall probability of winning that biformat is found. This relationship is then used to derive an exact relationship between the efficiency of a biformat, its expected duration and the average importance of the points within it (equation (4.3.2)). An approximation to this relationship (equation (4.3.3)) is then used in conjunction with Theorem 4.4 to show that the limiting value of the efficiency of any biformat F at $(p+\delta, p-\delta)$ as $\delta \to 0$ is given by

$$\lim_{\delta \to 0} \rho(F(p+\delta, p-\delta)) = \frac{\overline{I}^2}{\overline{I^2}}$$

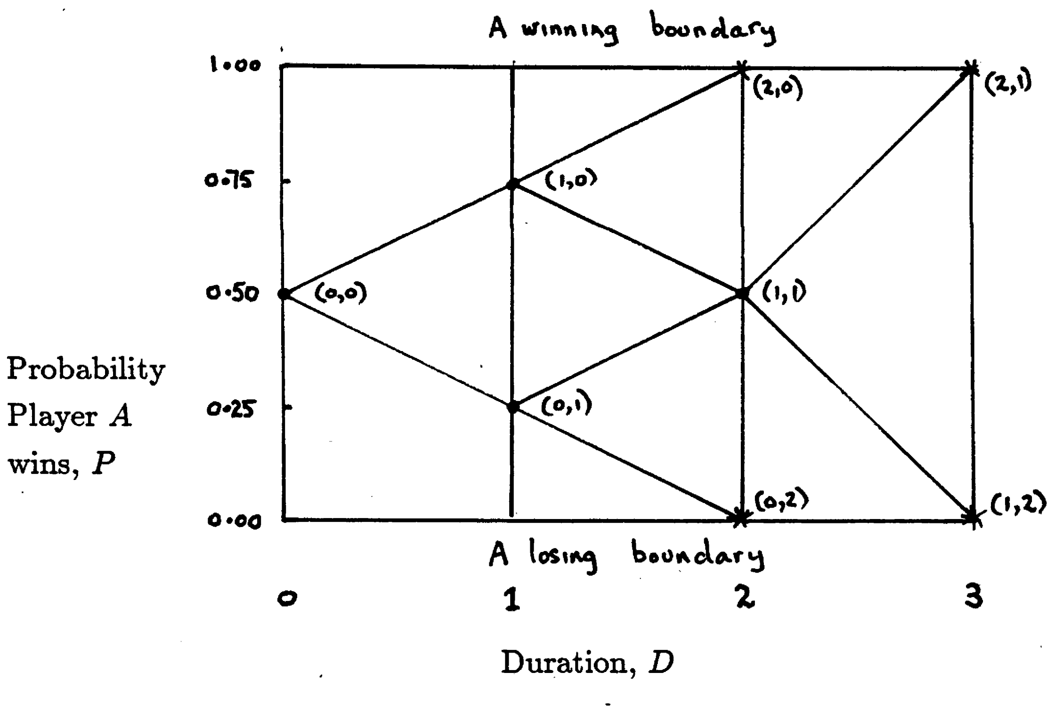
where $\overline{I}^2$ is the square of the average of the importances of the points and $\overline{I^2}$ is the average of the squares of the importances of the points. It follows that it is not possible for a biformat to have a limiting efficiency greater than unity and that the efficiency of a general biformat F at $(p+\delta, p-\delta)$ for sufficiently small δ must be less than unity if points are not equally important at $p_a = p_b = p$. Thus, it can be argued that the differing importances of the points (when $p_a = p_b$) within the tennis scoring system do contribute to its inefficiency.

4.2 The $P-D-I$ Diagram and a property of the total spectrum of W_n biformats.

This is a diagram which gives insight into why some scoring systems are efficient and others inefficient and is a useful tool in understanding this chapter and the next. This diagram can be viewed as a way of decomposing a scoring system into smaller components and monitoring progress towards the acceptance or rejection of H_0.

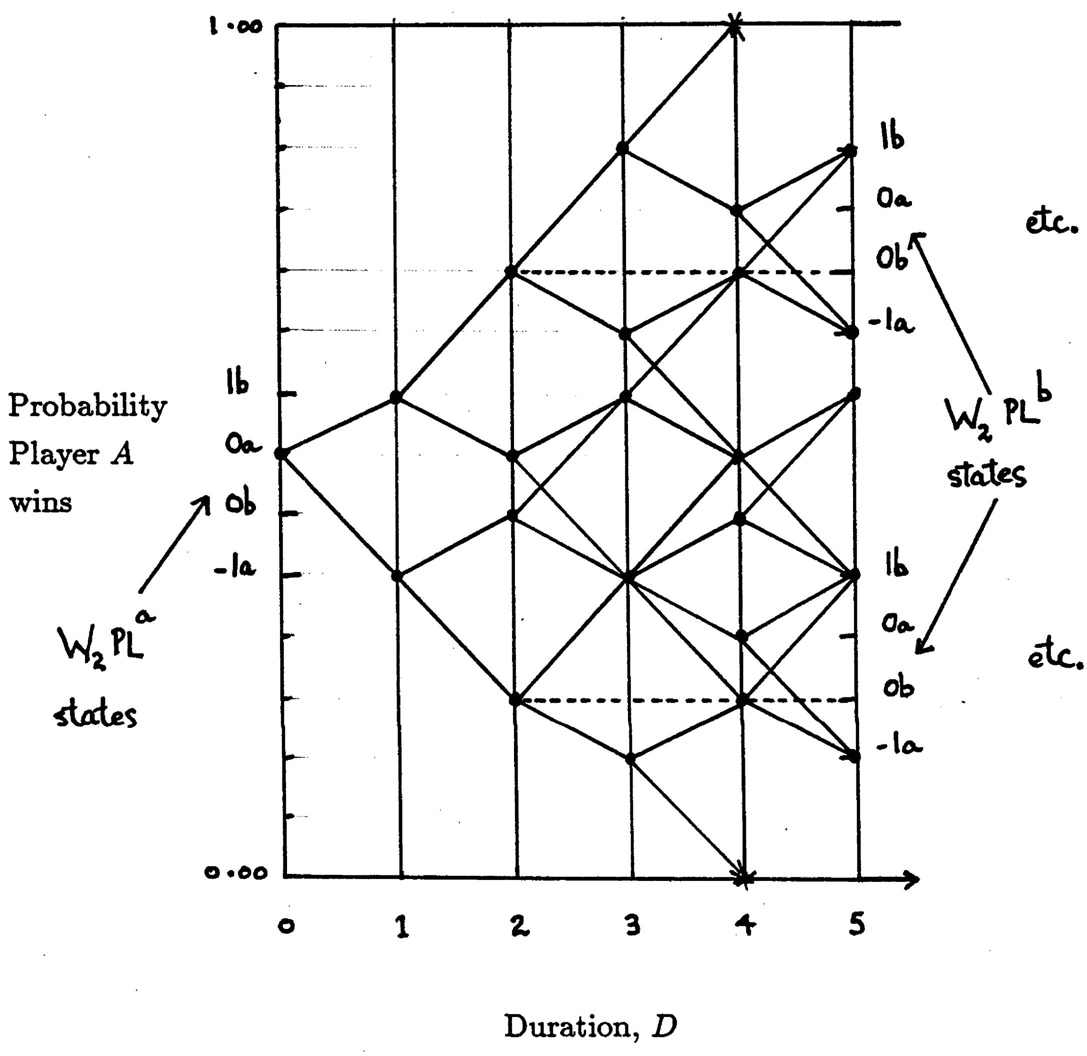
Morris's elegant definition of the importance of a point within a match is now recalled. He defined the *importance*, I_i, of the ith non-absorbing state as the difference between two probabilities:- the probability player A wins the match given a win (to player A) in this state minus the probability player A wins the match given a loss in this state.

Figure 4.2.1.
The $P-D-I$ diagram for the uniformat B_3 when $p = 0.5$.



All realizations of a scoring system can be located on a $P-D-I$ diagram of it. P represents the probability player A wins the match given the current state; D represents the duration so far; I is the importance of the current state and is given by the width of the next branch within the diagram. For example, Figure 4.2.1 gives the $P-D-I$ diagram for 2 equal players for the uniformat B_3. Note the

Figure 4.2.2.
Part of the $P-D-I$ diagram for $W_1(W_2PL^a, W_2PL^b)$ at $p_a = p_b = \frac{2}{3}$, the initial point being an a-point.



isosceles triangles (2 equal players in unipoints!) emerging from each state, the importance of that state being the base of that triangle. Here, for example, state (1,1) is seen to be more important than the other states. It will be seen that these differing importances lead to the inefficiency of B_3 relative to W_n at $p = 0.5 \pm \delta$ (Miles (1984) has shown that $\lim_{\delta \to 0} \rho\{B_3(0.5+\delta)\} = 0.9 < 1$). It can be shown that all states within the uniformat W_n have equal importances $(\frac{1}{n})$ when $p = \frac{1}{2}$. Returning to bipoints, Figure 4.2.2 gives, for example, the $P-D-I$ diagram for the reversal biformat $W_1(W_2PL^a,\ W_2PL^b)$ at $p_a = p_b = \frac{2}{3}$, the initial point being an a-point. All states in Figure 4.2.2 are equally important. This result is now generalized in the following Theorem 4.1 and Corollary 4.1.1.

Theorem 4.1. *The importance of every point in the scoring system W_nPL^x (where $x = a$ or b) is equal to $(1+2(n-1)p)^{-1}$ when $p_a = p_b = p$.*

Proof: Suppose $P(i,a)$ $(P(i,b))$ represents the probability player A wins when the score is i and an a-point (b-point) is about to be played (i.e. the state is (i,a) $((i,b))$).

$$\left.\begin{array}{ll}\text{Then} & P(i,a) = pP(i+1,b) + qP(i-1,a) \\ \text{and} & P(i,b) = qP(i+1,b) + pP(i-1,a)\end{array}\right\} \quad \begin{array}{r} i = n-1, n-2, n-3, \\ \ldots, -(n-1)\,, \end{array}$$

where $q = 1-p$. Boundary conditions are $P(n,b) = 1$ and $P(-n,a) = 0$. (Note that, for simplicity, states $(n-1,a)$ and $(-(n-1),b)$ are defined even though they cannot occur.) The solution of these equations is

$$\left.\begin{array}{l} P(i,a) = \frac{(n+i)p}{1+2(n-1)p} \\ P(i,b) = \frac{(n+i-2)p+1}{1+2(n-1)p} \end{array}\right\} \quad i = n-1, n-2, n-3, \ldots, -(n-1)\,,$$

and the importance of state (i,a) or (i,b) is equal to $P(i+1,b) - P(i-1,a)$ which equals $(1+2(n-1)p)^{-1}$, establishing the result.

Corollary 4.1.1. *The importance, I, of every point in the reversal biformat $W_1(W_nPL^a,\, W_nPL^b)$ is given by*

$$I = \big(2(1+2(n-1)p)\big)^{-1} \qquad \text{when } p_a = p_b = p\,.$$

Thus, for the reversal biformat $W_1(W_nPL^a,\, W_nPL^b)$, P increases (decreases) by qI (pI) whenever player A wins (loses) an a-point. Similarly P increases (decreases) by pI (qI) whenever player A wins (loses) a b-point. Thus, since P increases (decreases) by qI (pI) with probability p (q) when an a-point is played, the variance, V, of the new value of P resulting from this single a-point is given by

$$V = p(qI)^2 + q(-pI)^2 = pqI^2$$

when $p_a = p_b = p$. This is also the variance resulting from a single b-point when $p_a = p_b = p$. Here we have a variance interpretation of the importance of a point. Using a slight extension (in which the random variables are independent (but not identically distributed) with common mean 0 and variance σ^2, $0 < \sigma^2 < \infty$) of the following theorem, it follows that the expected duration of the biformat $W_1(W_nPL^a,\, W_nPL^b)$ is $(4pqI^2)^{-1}$ points when $p_a = p_b = p$.

Theorem 4.2. *(See, for example, Rohatgi (1976, p.598) or Moran (1968, p.481–484).) Let $X_1, X_2, X_3, \ldots,$ be a sequence of i.i.d. r.v.'s with common mean 0 and variance σ^2, $0 < \sigma^2 < \infty$. For any sequential stopping rule in which N is the number of X variables at the stopping point and $E(N) < \infty$, if $E\left\{\left(\sum_{k=1}^{N} |X_k|\right)^2\right\} < \infty$, then*

$$E\left\{\left(\sum_{i=1}^{N} X_i\right)^2\right\} = \sigma^2 EN\ .$$

Figure 4.2.3 gives the $P-D-I$-type diagram for $W_2(PL\,g.p.p.)$ when $p_a = p_b = \frac{2}{3}$. Now each point-pair has 3 possible outcomes (win, draw or loss) and hence the Morris definition of the importance of a state cannot be directly applied. However, a variance interpretation to the importance of a point-pair is possible, as can be seen from the following Theorem 4.3.

Theorem 4.3. *For the biformat $W_1(W_nPL\,g.p.p.)$ when $p_a = p_b = p$, the variance in the new value of P resulting from any single general point-pair (of the type (a,a), (b,b), or (a,b)) is equal to $2pqI^2$ where $I = \big(2(1+2(n-1)p)\big)^{-1}$.*

Proof: Suppose $P(i,aa)$ represents the probability player A wins when player A is i point-pairs ahead and an aa-point-pair is about to be played (i.e. the system is in state (i,aa)). Suppose $P(i,ab)$ and $P(i,bb)$ are correspondingly defined. Then

$$\left.\begin{aligned} P(i,aa) &= p^2P(i+1,bb) + 2pqP(i,ab) + q^2P(i-1,aa) \\ P(i,ab) &= pqP(i+1,bb) + (p^2+q^2)P(i,ab) + pqP(i-1,aa) \\ P(i,bb) &= q^2P(i+1,bb) + 2pqP(i,ab) + p^2P(i-1,aa) \end{aligned}\right\} \begin{aligned} i &= n-1, n-2, \\ &\ldots, -(n-1)\ , \end{aligned}$$

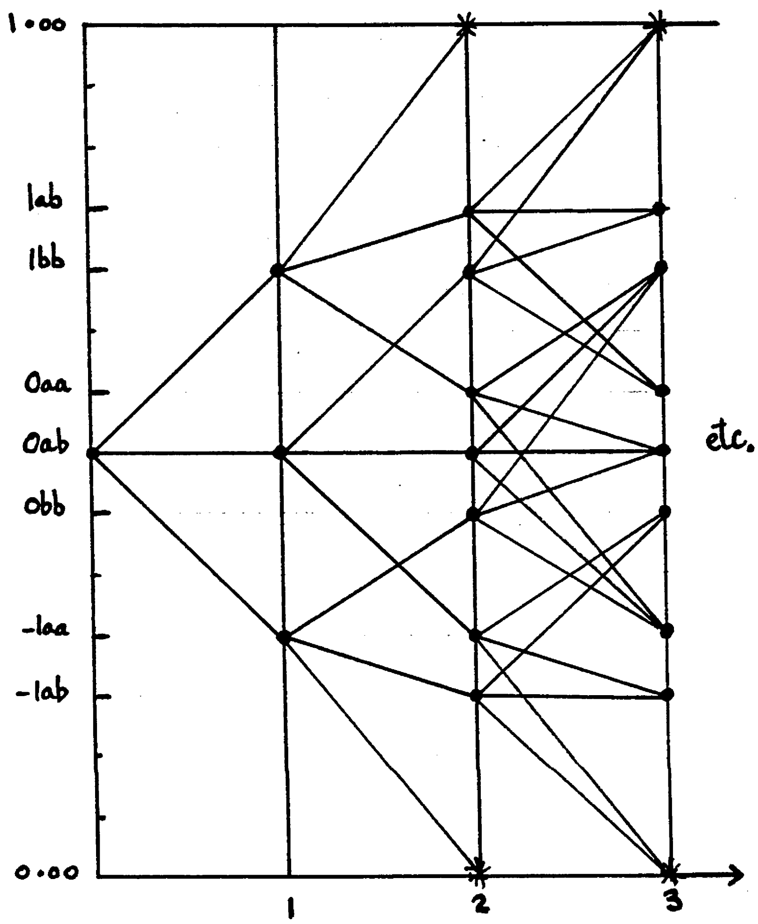
and the boundary conditions are $P(n,bb) = 1$ and $P(-n,aa) = 0$. Note that, again for simplicity, the states $(n-1,aa)$ and $(-(n-1),bb)$ are defined even though they cannot occur. The solution of these equations is

$$\left.\begin{aligned} P(i,aa) &= \tfrac{(n+i)p}{1+2(n-1)p} \\ P(i,ab) &= \tfrac{(n+i-1)p+\frac{1}{2}}{1+2(n-1)p} \\ P(i,bb) &= \tfrac{(n+i-2)p+1}{1+2(n-1)p} \end{aligned}\right\} \quad i = n-1, n-2, \ldots, -(n-1)\ .$$

Now the variance, V, in the new value of P after one point-pair (starting from the state (i,aa)) is given by (noting that the expected value of the new P remains equal to $P(i,aa)$)

$$\begin{aligned} V &= p^2\big(P(i+1,bb) - P(i,aa)\big)^2 + 2pq\big(P(i,ab) - P(i,aa)\big)^2 \\ &\qquad + q^2\big(P(i-1,aa) - P(i,aa)\big)^2 \\ &= \frac{2pq}{4(1+2(n-1)p)^2} \quad \text{after a few steps}\ . \end{aligned}$$

Figure 4.2.3
Part of the $P - D - I$ diagram for $W_2(PL\,g.p.p.)$ at $p_a = p_b = \frac{2}{3}$.



Probability Player A wins

Duration (point-pairs)

It can correspondingly be shown that the variance in the new value of P resulting from the states (i, ab) or (i, bb) is the same as the above and hence the theorem is proved.

Again using Theorem 4.2, the expected duration of $W_n(PL\,g.p.p)$ when $p_a = p_b = p$ is equal to $(8pqI^2)^{-1}$ point-pairs or $(4pqI^2)^{-1}$ points, in agreement with the result for $W_1(W_nPL^a,\ W_nPL^b)$.

Corresponding results are true for the total spectrum of biformats. For example, consider $W_1(W_nPL_l^a,\ W_nPL_l^b)$ when $p_a = p_b = p$. Again points are equally important and have importance, I_l, given by

$$I_l = \frac{1}{2(n - l + 2lp)} ,$$

which can be established in a manner similar to the proof of property 3 in Section 3.3. The expected duration is equal to $(4pqI_l^2)^{-1}$ points. The biformats $W_1(W_nPW_l^a,\ W_nPW_l^b)$ when $p_a = p_b = p$ can be handled similarly as can the additional members of the total spectrum of W_n biformats with odd durations.

4.3 A Relationship Between Efficiency and Importances

Consider, for simplicity, a general unipoints game in which state i (a non-absorbing state) has importance I_i when $p = 0.5$ (all i). Suppose the expected number of times state i occurs when $p = 0.5 + \delta$ is n_i (all i). Then, referring to the $P - D - I$ diagram of this unipoints game (and assuming it is 'fair' (starts at $\frac{1}{2}$ when $p = 0.5$)) it follows that

$$(P - Q)\frac{1}{2} = \sum_{\text{all } i} n_i\left[\frac{I_i}{2}(0.5 + \delta) - \frac{I_i}{2}(0.5 - \delta)\right]$$

where P (Q) is the probability player A wins (loses) the game when $p = 0.5 + \delta$. Both sides of the above equation represent the expected sum of the steps in the $P - D - I$ diagram for start of play until absorption. Thus, if $P = 0.5 + \Delta P$,

$$\Delta P = \sum_{\text{all } i} n_i I_i \delta$$

or

$$\Delta P = M\delta \quad \text{where} \quad M = \sum_{\text{all } i} n_i I_i .$$

Morris (1977) called M the magnification factor and showed $\Delta P \doteq M\delta$. However, by defining I_i to be the importance of state i when $p = 0.5$ and n_i to be the expected number of times state i is entered when $p = 0.5+\delta$, Morris's approximate result becomes exact. Another form of this equation is

$$\frac{\Delta P}{\mu} = \sum_{\text{all } i} \pi_i I_i \delta = \overline{I}\delta \tag{4.3.1}$$

where, when $p = 0.5 + \delta$, π_i is the long-run proportion of time spent in state i and μ is the expected duration of the game. Note that $\overline{I}$ is the average importance of a point, the importances being at $p = 0.5$ and the weights, π_i, being at $p = 0.5 + \delta$.

In the above analysis a "fair" unipoints game has been considered for simplicity. However, for a "fair" bipoints game, Morris's approximate result is correspondingly exact. The reader is referred to Section 10.2 of this thesis for a proof of this result in a more general case. Now returning to bipoints, the efficiency, ρ, of a general biformat $F(P,\mu)$ is given by equation (1.2.1). Consider the case in which player A is the better player. Then there exists a p and a $\delta > 0$ such that $p_a = p + \delta$ and $p_b = p - \delta$. Let $P = \frac{1}{2} + \Delta P$ and $Q = \frac{1}{2} - \Delta P$. Then,

$$\rho = \frac{2(2\Delta P)\ln[(1+2\Delta P)(1-2\Delta P)^{-1}]}{\mu(2\delta)\ln[(1+\frac{\delta}{p})(1-\frac{\delta}{p})^{-1}(1+\frac{\delta}{q})(1-\frac{\delta}{q})^{-1}]} \quad \text{where } q = 1-p$$

$$= 2\overline{I}\frac{\ln\left(\frac{1+2\mu\overline{I}\delta}{1-2\mu\overline{I}\delta}\right)}{\ln\left(\frac{(1+\frac{\delta}{p})(1+\frac{\delta}{q})}{(1-\frac{\delta}{p})(1-\frac{\delta}{q})}\right)} \tag{4.3.2}$$

using equation (4.3.1)

$$= \frac{2\overline{I}[4\mu\overline{I}\delta + \text{powers of } \delta^3 \text{ and higher}]}{[\frac{2\delta}{p} + \frac{2\delta}{q} + \text{powers of } \delta^3 \text{ and higher}]}$$

$$\doteq 4pq\overline{I}^2\mu\,. \tag{4.3.3}$$

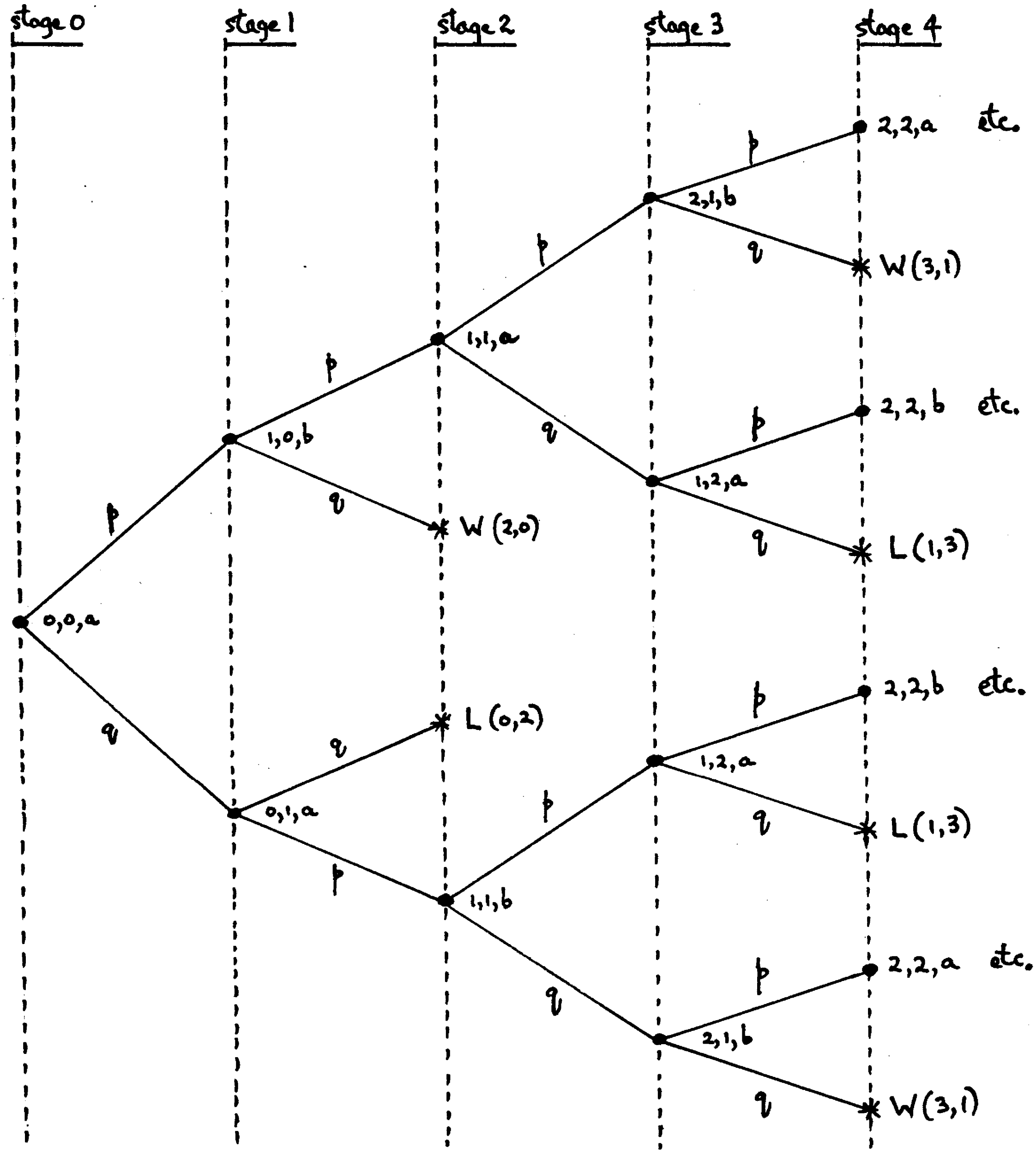
Equation (4.3.2) gives an exact relationship between the efficiency of a general biformat at $(p+\delta,\, p-\delta)$, the expected duration and the average importance of the points within it. Note that δ need not be small. Equation (4.3.3) is the associated approximate relationship for small δ. (In unipoints, the corresponding equation to (4.3.3) is $\rho \doteq \overline{I}^2\mu$). Recall that Section 4.2 showed that, for any member of the total spectrum of W_n biformats considered in Chapters 2 and 3, points were equally important when $p_a = p_b = p$ and expressions for the expected durations were given. These results and equation (4.3.3) implies that the efficiency of any member of this total spectrum of W_n biformats converges to unity as $\delta \to 0$. It will be seen in Chapter 5 that the biformats in this total spectrum are the only biformats with this property.

4.4 Arbitrary Sequential Bipoints Rules and a Related Theorem

Consider a bipoints game in which, at each stage of the game, either an $a-$ or $b-$point is played, quite arbitrarily, or the decision is made to terminate play, quite arbitrarily, with a win or loss, also quite arbitrarily, given to player A. Suppose such a game is called an arbitrary sequential bipoints decision rule. It may not be fair or possess the properties of a biformat, but can be represented by a binary tree.

Figure 4.4.1

A Binary Tree Diagram for W_2PL^a when $p_a = p_b = p$.



Throughout this section it is assumed that the 2 players are equal ($p_a = p_b = p$). The first point (an $a-$ or $b-$point) at stage 0 takes the game to stage 1, etc. At stage i, there are at most 2^i nodes in the tree and each node is either an a or b node (or absorbing node) depending on whether an $a-$ or $b-$point is about to be played (or whether play terminates). Figure 4.4.1 gives a binary tree diagram for W_2PL^a. Absorption nodes are represented by a $*$. Figure 4.4.2 gives a binary tree

Figure 4.4.2

A Binary Tree Diagram for an arbitrary sequential decision rule when $p_a = p_b = p$ and $q = 1 - p$.

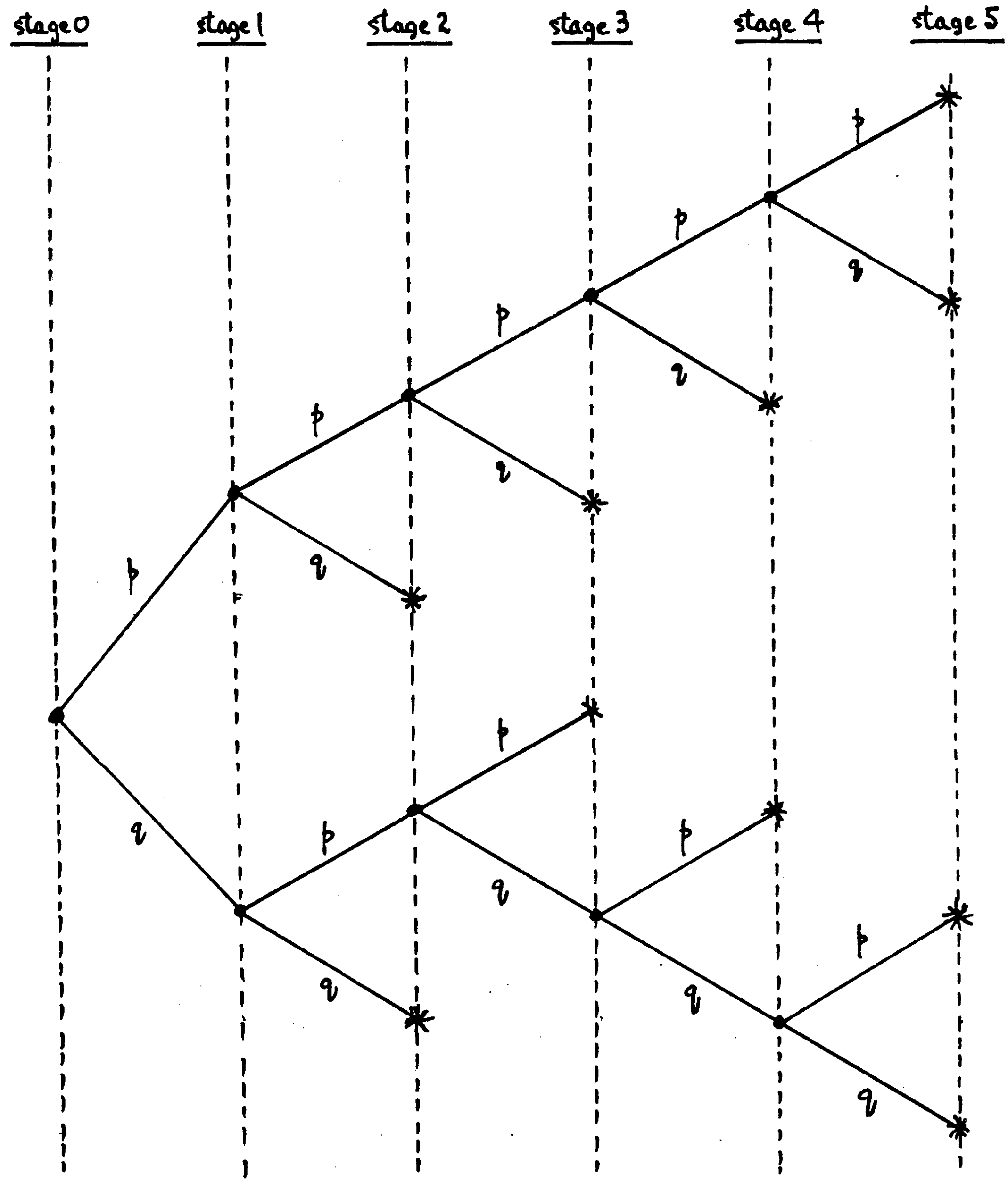


diagram for an arbitrary sequential bipoints decision rule of at most 5 points in which the a- and b- points need not be uniquely identified, neither need wins nor losses at the absorption points.

Now consider the steps (decreases or increases in P) in the $P-D-I$ diagram for an arbitrary sequential bipoints decision rule involving two equal players. These steps are equal to

$$qI_i \text{ with probability } p$$
$$-pI_i \text{ with probability } q$$

if state i is an a-point and has importance I_i,

$$pI_j \text{ with probability } q$$
$$-qI_j \text{ with probability } p$$

if state j is a b-point and has importance I_j. Here $q = 1 - p$. The following theorem, which can be seen as a modification and extension of Theorem 4.2, gives, for any arbitrary sequential bipoints decision rule between two equal players, an expression for the weighted sum of the squares of the importance of the states, the weights being the expected number of times that each state is visited.

Theorem 4.4. *For any arbitrary sequential bipoints decision rule between 2 equal players ($p_a = p_b = p$), there are two expressions for the value of*

$$\sum_{\substack{\text{all paths}\\ \text{to absorption}}} P(\text{path})\{\text{sum of the steps (in the } P-D-I \text{ diagram) of that path}\}^2 .$$

(i) $P(1-P)$ where P is the probability player A wins the arbitrary sequential bipoints decision rule.

(ii) $pq \sum_{\text{all } j} n_j I_j^2$ where the subscript j refers to the jth non-absorbing state, I_j is the importance of that state when $p_a = p_b = p$ and n_j is the expected number of times that state j is visited in one realization when $p_a = p_b = p$.

Proof: (i) This follows immediately as the sum of the steps in the path is either $1-P$ with a (total) probability of P or $-P$ with a (total) probability of $1-P$.

(ii) The proof of (ii) is more complicated and the method of induction is used. Suppose $n_{a1}(i)(n_{b1}(i))$ is the number of non-absorbing a (b) nodes at stage i.

Similarly, let $n_{a0}(i)(n_{b0}(i))$ be the number of absorbing a (b) nodes at stage i. Clearly $n_{a0}(i)+n_{a1}(i)+n_{b0}(i)+n_{b1}(i) \le 2^i$. The strict inequality applies when there exists at least one absorbing node at stage $i-1$ or earlier. The subscript k ($k = 1,2,\ldots,n_{a1}(i)$) is firstly used to denote the kth non-absorbing a-node at stage i and the subscript l ($l = 1,2,\ldots,n_{b1}(i)$) is used to denote the lth non-absorbing b-node at stage i. Suppose $P_{a1k}^{(i)}(P_{b1l}^{(i)})$ is the probability of the path from the start of the binary tree until the kth (lth) non-absorbing a (b) node is reached at stage i. Also, Suppose $s_{a1k}^{(i)}$ ($s_{b1l}^{(i)}$) is the sum of the squares of the i steps within the $P-D-I$ diagram along this path ($k = 1,2,\ldots,n_{a1}(i)$; $l = 1,2,\ldots,n_{b1}(i)$). Similarly, suppose $P_{a0k}^{(i)}$ ($P_{b0l}^{(i)}$) is the probability of the path from the start of the binary tree until the kth (lth) absorbing a (b) node is reached at stage i. Also, suppose $s_{a0k}^{(i)}$ ($s_{b0l}^{(i)}$) is the sum of the squares of the i steps within the $P-D-I$ diagram along this path ($k = 1,2,\ldots,n_{a0}(i)$; $l = 1,2,\ldots,n_{b0}(i)$).

Suppose $I_{a1k}^{(i)}$ ($I_{b1l}^{(i)}$) is the importance of the kth (lth) non-absorbing a (b) node at stage i ($k = 1,2,\ldots,n_{a1}(i)$; $l = 1,2,\ldots n_{b1}(i)$). Suppose S_m is defined by

$$S_m = \sum_{k=1}^{n_{a1}(m)} P_{a1k}^{(m)} s_{a1k}^{(m)} + \sum_{l=1}^{n_{b1}(m)} P_{b1l}^{(m)} s_{b1l}^{(m)} + \sum_{i=0}^{m}\sum_{k=1}^{n_{a0}(i)} P_{a0k}^{(i)} s_{a0k}^{(i)} + \sum_{i=0}^{m}\sum_{l=1}^{n_{b0}(i)} P_{b0l}^{(i)} s_{b0l}^{(i)} .$$

Suppose T_m is defined by

$$T_m = \sum_{i=0}^{m-1}\left[\sum_{k=1}^{n_{a1}(i)} P_{a1k}^{(i)} I_{a1k}^{(i)^2} + \sum_{l=1}^{n_{b1}(i)} P_{b1l}^{(i)} I_{b1l}^{(i)^2}\right] .$$

The method of induction will be used to show $S_m = pqT_m$. Firstly, suppose without loss of generality we assume that the first point is an a-point. Then $T_1 = 1.I^2$ where I is the importance of the first point. Also $S_1 = p(qI)^2 + q(-pI)^2 = pqI^2 = pqT_1$. We will now assume $S_m = pqT_m$ in order to show $S_{m+1} = pqT_{m+1}$. Noting

$$\sum_{i=0}^{m+1}\sum_{k=1}^{n_{a0}(i)} P_{a0k}^{(i)} s_{a0k}^{(i)} = \sum_{k=1}^{n_{a0}(m+1)} P_{a0k}^{(m+1)} s_{a0k}^{(m+1)} + \sum_{i=0}^{m}\sum_{k=1}^{n_{a0}(i)} P_{a0k}^{(i)} s_{a0k}^{(i)}$$

and a similar expression for the b nodes, it follows that

$$S_{m+1} = \sum_{k=1}^{n_{a1}(m+1)} P_{a1k}^{(m+1)} s_{a1k}^{(m+1)} + \sum_{k=1}^{n_{a0}(m+1)} P_{a0k}^{(m+1)} s_{a0k}^{(m+1)}$$
$$+ \sum_{l=1}^{n_{b1}(m+1)} P_{b1l}^{(m+1)} s_{b1l}^{(m+1)} + \sum_{l=1}^{n_{b0}(m+1)} P_{b0l}^{(m+1)} s_{b0l}^{(m+1)}$$
$$+ \sum_{i=0}^{m} \sum_{k=1}^{n_{a0}(i)} P_{a0k}^{(i)} s_{a0k}^{(i)} + \sum_{i=0}^{m} \sum_{l=1}^{n_{b0}(i)} P_{b0l}^{(i)} s_{b0l}^{(i)}$$

which, since nodes which are non-absorbing or absorbing at (exactly) stage $m+1$ must have resulted from a binary branch from a non-absorbing node at stage m,

$$= \sum_{k=1}^{n_{a1}(m)} P_{a1k}^{(m)} \left\{ p\left(s_{a1k}^{(m)} + (qI_{a1k}^{(m)})^2 \right) + q\left(s_{a1k}^{(m)} + (-pI_{a1k}^{(m)})^2 \right) \right\}$$
$$+ \sum_{l=1}^{n_{b1}(m)} P_{b1l}^{(m)} \left\{ p\left(s_{b1l}^{(m)} + (-qI_{b1l}^{(m)})^2 \right) + q\left(s_{b1l}^{(m)} + (pI_{b1l}^{(m)})^2 \right) \right\}$$
$$+ \sum_{i=0}^{m} \sum_{k=1}^{n_{a0}(i)} P_{a0k}^{(i)} s_{a0k}^{(i)} + \sum_{i=0}^{m} \sum_{l=1}^{n_{b0}(i)} P_{b0l}^{(i)} s_{b0l}^{(i)} .$$

Noting

$$\left\{ p\left(s_{a1k}^{(m)} + (qI_{a1k}^{(m)})^2 \right) + q\left(s_{a1k}^{(m)} + (-pI_{a1k}^{(m)})^2 \right) \right\} = \left\{ s_{a1k}^{(m)} + pqI_{a1k}^{(m)^2} \right\} ,$$

and a similar expression for the b nodes, it follows that

$$S_{m+1} = \sum_{k=1}^{n_{a1}(m)} P_{a1k}^{(m)} \left\{ s_{a1k}^{(m)} + pqI_{a1k}^{(m)^2} \right\}$$
$$+ \sum_{l=1}^{n_{b1}(m)} P_{b1l}^{(m)} \left\{ s_{b1l}^{(m)} + pqI_{b1l}^{(m)^2} \right\}$$
$$+ \sum_{i=0}^{m} \sum_{k=1}^{n_{a0}(i)} P_{a0k}^{(i)} s_{a0k}^{(i)} + \sum_{i=0}^{m} \sum_{l=1}^{n_{b0}(i)} P_{b0l}^{(i)} s_{b0l}^{(i)}$$
$$= S_m + pq\left(\sum_{k=1}^{n_{a1}(m)} P_{a1k}^{(m)} I_{a1k}^{(m)^2} + \sum_{l=1}^{n_{b1}(m)} P_{b1l}^{(m)} I_{b1l}^{(m)^2} \right)$$
$$= pq\left(T_m + \sum_{k=1}^{n_{a1}(m)} P_{a1k}^{(m)} I_{a1k}^{(m)^2} + \sum_{l=1}^{n_{b1}(m)} P_{b1l}^{(m)} I_{b1l}^{(m)^2} \right)$$

assuming $S_m = pqT_m$

$= pqT_{m+1}$ and hence the result follows .

Note that S_n and T_n are monotonic increasing.

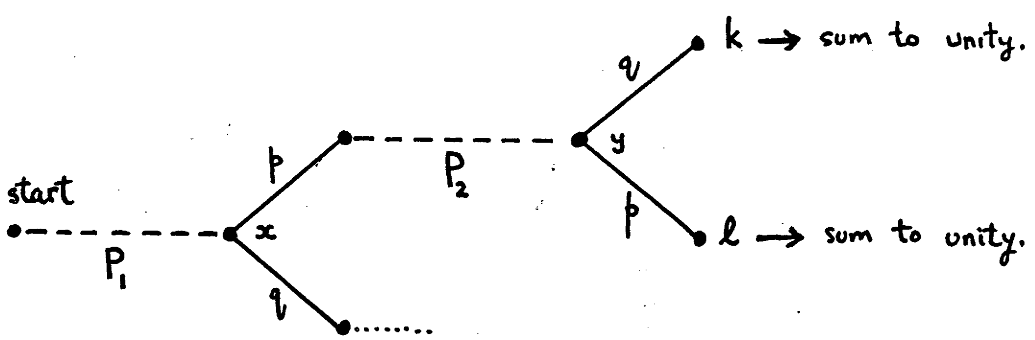
Now, let us assume that the arbitrary sequential decision rule has a maximum possible duration, N. Then,

$$S_N = \sum_{i=0}^{N} \left[\sum_{k=1}^{n_{a0}(i)} P_{a0k}^{(i)} s_{a0k}^{(i)} + \sum_{l=1}^{n_{b0}(i)} P_{b0l}^{(i)} s_{b0l}^{(i)} \right]$$

since, at stage N, $n_{a1}(N)$ and $n_{b1}(N)$ are zero. Now, if $t_{a0k}^{(i)}$ ($t_{b0l}^{(i)}$) is the sum of the i steps within the $P-D-I$ diagram along the kth (lth) path to absorption at stage i ($k = 1, 2, \ldots, n_{a0}(i)$; $l = 1, 2, \ldots, n_{b0}(i)$), it can be shown that

$$S_N = \sum_{i=0}^{N} \left[\sum_{k=1}^{n_{a0}(i)} P_{a0k}^{(i)} t_{a0k}^{(i)^2} + \sum_{l=1}^{n_{b0}(i)} P_{b0l}^{(i)} t_{b0l}^{(i)^2} \right]$$

since cross product terms either do not exist or sum to zero. The cross product term does not exist if 2 nodes in the binary tree do not share a path (to absorption). Now consider the case of 2 nodes ((say) x and y) which do share a path. Suppose node x (y) has importance I_x (I_y). Let P_1 be the probability of the path from the start of the binary tree to node x, which is assumed to be at an earlier stage than node y. Without loss of generality, assume node x (y) corresponds to (say) an a-(b-) point. Let P_2 be the probability of the path from the node reached after a win to player A at node x to node y. Diagrammatically,



Suppose node k (l) is reached after a win (loss) to player A at node y. Note that the sum of probabilities of all the paths to absorption from node k (l) is 1. Thus,

$$\sum_{\substack{\text{all paths}\\ \text{to absorption}}} (xy \text{ cross product terms})$$
$$= P_1 p(qI_x) P_2 \{q(pI_y) + p(-qI_y)\}$$
$$= 0\,.$$

(Similarly, the sum of any cross product terms conditional on a loss in state x is zero.) Now, since $S_N = pqT_N$, we have

$$\sum_{i=0}^{N}\left[\sum_{k=1}^{n_{a0}(i)} P_{a0k}^{(i)} t_{a0k}^{(i)^2} + \sum_{l=1}^{n_{b0}(i)} P_{b0l}^{(i)} t_{b0l}^{(i)^2}\right]$$
$$= pq\sum_{i=0}^{N-1}\left[\sum_{k=1}^{n_{a1}(i)} P_{a1k}^{(i)} I_{a1k}^{(i)^2} + \sum_{l=1}^{n_{b1}(i)} P_{b1l}^{(i)} I_{b1l}^{(i)^2}\right].$$

Noting that the probability of a path from the start of the binary tree to a node is also the expected number of times that that node is reached, the result follows.

Finally, if there does not exist a maximum possible duration, then, by taking the limit ($n \to \infty$) of both sides of the equation $S_n = pqT_n$, the result similarly follows. This completes the proof.

4.5 A Further Relationship Between Efficiency and Importances

For any biformat F and two equal players ($p_a = p_b = p$), it follows from Theorem 4.4, since $P = 0.5$, that

$$pq\sum_{\text{all } y} n_j I_j^2 = 0.25$$

$$\text{or} \quad 4pq\mu\sum_{\text{all } j} \pi_j I_j^2 = 1$$

$$\text{or} \quad \mu = \frac{1}{4pq\overline{I^2}} \qquad (4.5.1)$$

where π_j is the long run proportion of time in non- absorbing states that the biformat is in non-absorbing state j when $p_a = p_b = p$, $\overline{I^2} = \sum_{\text{all } j} \pi_j I_j^2$ and μ is the expected duration of the biformat when $p_a = p_b = p$.

Now consider the limiting value of the efficiency of the general biformat F at $(p+\delta, p-\delta)$ as $\delta \to 0$. Assuming the continuity of μ and π_j in δ and using equation (4.5.1) in equation (4.3.3), this limiting efficiency is given by

$$\lim_{\delta\to 0} \rho(F(p+\delta,\, p-\delta)) = \frac{\overline{I}^2}{\overline{I^2}}\,. \qquad (4.5.2)$$

Thus, it is not possible for a biformat to have limiting efficiency greater than 1. By the continuity of efficiency in δ, it follows from (4.5.2) that the efficiency of a general biformat F at $(p+\delta,\, p-\delta)$ for small δ is less than unity if points are not equally important when $p_a = p_b = p$. Thus, in our search for the most efficient biformat, we will restrict attention to those biformats with equally important points for two equal players.

4.6 An application of the relationship in Section 4.5

Consider the unipoints or bipoints situation in which the W_{n_1} framework is used up until some duration d where, if the game is unfinished, the W_{n_2} framework is then used. Then, *as a first approximation* (noting the results of Section 4.2), it can be seen that points played after duration d have greater (smaller) importance than those played before that stage if n_2 is less (greater) than n_1. Thus, it can be seen using equation (4.5.2) that when a mixing of W_n systems in the above sense produces a uniformat or biformat, then the efficiency of this uniformat at $(\frac{1}{2}+\delta)$ or biformat at $(p+\delta,\ p-\delta)$ is less than 1 (δ small).

4.7 Summary.

Section 4.2 introduced the $P-D-I$ diagram which has been found (not only in this chapter but also in Chapters 5, 9 and 10) to be quite useful in the analysis of scoring systems. This diagram showed that although the definition of importance could not directly be applied to point-pairs, a variance interpretation of the importance of a point-pair was possible. This section also showed that every biformat within the total spectrum of W_n biformats has the following property: each point within that biformat is equally important when $p_a = p_b$. In Section 4.3 a relationship between the increased probability of winning every point and the increased probability of winning a biformat was found (equation (4.3.1)). This relationship was then used to derive an exact relationship between the efficiency of a biformat, its expected duration and the average importance of the points within it (equation (4.3.2)). An approximation to this relationship at $(p_a, p_b) = (p+\delta, p-\delta)$ was seen to be

$$\rho \doteq 4pq\overline{I}^2\mu \qquad \text{for } \delta \text{ small}\,. \tag{4.3.3}$$

Arbitrary sequential bipoints decision rules were introduced in Section 4.4 and a related theorem, Theorem 4.4, was proved. As a consequence of this theorem, for any biformat F and two equal players ($p_a = p_b = p$), it follows that

$$\mu = \frac{1}{4pq\overline{\overline{I^2}}}\,. \tag{4.5.1}$$

Thus the limiting value of the efficiency of any biformat F at $(p+\delta, p-\delta)$ as $\delta \to 0$ is given by

$$\lim_{\delta\to 0} \rho(F(p+\delta, p-\delta)) = \frac{\overline{I}^2}{\overline{\overline{I^2}}}\,. \tag{4.5.2}$$

Hence it can be seen that it is not possible for a biformat to have limiting efficiency greater than unity. Also, it follows that the efficiency of a general biformat F at $(p+\delta, p-\delta)$ for sufficiencly small δ must be less than unity if points are not equally important when $p_a = p_b = p$.

It will be shown in Chapter 5 that the property of equally important points when $p_a = p_b$ is a property that *characterizes* the total spectrum of W_n biformats.

CHAPTER 5
A FAMILY OF OPTIMALLY EFFICIENT BIFORMATS: THE OPTIMAL 2-SAMPLE BINOMIAL TEST

5.1 The Family of Optimal Biformats

We are interested in finding the optimally efficient biformat. However, suppose we restrict ourselves to the set of biformats which have "very good" efficiency near the equality diagonal $p_a = p_b$. Formally, suppose we restrict to the set of biformats, S, which have the property

$$\lim_{\delta \to 0} \rho(p+\delta, p-\delta) = 1 \;, \quad \text{for all } p \in (0,1) \;, \tag{5.1.1}$$

then Theorem 5.1 will show that any member of this set of biformats must have the same distribution of duration (and hence the same P, μ and ρ values) as a member of the total spectrum of biformats considered in Chapter 3. Thus, equation (5.1.1) is a property which characterizes the total spectrum of biformats. Recall that it follows from equation 4.5.2 that it is not possible to find a biformat for which the above limit is greater than 1.

Theorem 5.1. *If F is a biformat which is a member of the above set of biformats S, then F must have the same distribution of duration (and hence the same P, μ and ρ values) as a member of the total spectrum of biformats.*

Proof: Consider a finite interval along the equality diagonal, (p_1, p_1) to (p_2, p_2) where $0 < p_1 < p_2 < 1$. From equation (4.5.2) it follows that F must have equally important points at any point in this interval. Suppose (p_a, p_b) is a point within this finite interval such that

$$p_a/q_a = p_b/q_b = i_1 \;, \quad \text{an irrational number where } p_a = p_b \;.$$

Also, suppose for this biformat, n is the number of points in which player A can win by winning all $\frac{n+r}{2}$ a-points and all $\frac{n-r}{2}$ b- points (for some integer r (positive or negative) determined by the biformat). As all points are equally important and the $P-D-I$ diagram starts at $P = 0.5$, we have

$$\frac{n+r}{2}c + \frac{n-r}{2}i_1 c = 0.5$$

where c is player A's increase in P by winning a single a-point. c is thus the irrational number $\big((n+r)+(n-r)i_1\big)^{-1}$ determined by F and p_a $(= p_b)$. The importance of all points at (p_a, p_b) for this biformat F is $(1+i_1)c$. For this biformat and for this (p_a, p_b) point, $\frac{n+r}{2}$ and $\frac{n-r}{2}$ are 2 integers satisfying the relationship

$$\frac{n+r}{2}i_2 + \frac{n-r}{2}i_3 = 0.5 \tag{5.1.2}$$

where $i_2 = c$ and $i_3 = i_1 c$ are two irrational numbers determined by F and p_a $(= p_b)$. Another integer solution $\frac{n'+r'}{2}$ and $\frac{n'-r'}{2}$ for this (p_a, p_b) point and F such that

$$\frac{n'+r'}{2}i_2 + \frac{n'-r'}{2}i_3 = 0.5$$

cannot exist (or i_2 and i_3 would be rational). Thus, given all points are equally important, the way in which movement in the $P - D - I$ diagram finally *exactly* reaches the W (winning) boundary is as follows. Player A must win l (say, for some l) a- points and lose l b-points and must win k (say, for some k) b-points and lose k a-points (the P-effect of all these points cancel out in the $P - D - I$ diagram!) as well as the necessary winning of $\frac{n+r}{2}$ a-points and $\frac{n-r}{2}$ b-points. Such a realization takes $n + 2l + 2k$ points. Thus, for every possible winning (for player A) state, the number of a-points played minus the number of b-points played is r. Now, from equation (5.1.2), it follows that,

$$\frac{n+r}{2}(-i_2) + \frac{n-r}{2}(-i_3) = -0.5 \tag{5.1.3}$$

where i_2 is player A's decrease in P by losing a single b-point and i_3 is player A's decrease in P by losing a single a-point. Thus, since F is a biformat and the $P - D - I$ diagram starts at 0.5, equation (5.1.3) shows, for the same reasons as above, that the number of a-points played minus the number of b- points played is $-r$ for every possible losing (for player A) state. Thus, for this biformat (and for all (p_a, p_b), not just (p_a, p_b) on the equality diagonal), it follows that

$$\begin{cases} P\ (A \text{ wins in } n \text{ points}) = {p_a}^{(n+r)/2}{q_b}^{(n-r)/2} \\ P\ (B \text{ wins in } n \text{ points}) = {p_b}^{(n+r)/2}{q_a}^{(n-r)/2} \end{cases}$$

$$\text{and} \quad \begin{cases} P\ (A \text{ wins in } n+2m \text{ points}) = {p_a}^{(n+r)/2}{q_b}^{(n-r)/2}(P_m(p_a p_b, q_a q_b)) \\ P\ (B \text{ wins in } n+2m \text{ points}) = {p_b}^{(n+r)/2}{q_a}^{(n-r)/2}(P_m(p_a p_b, q_a q_b)) \end{cases}$$

for $m = 1, 2, \ldots$, where $P_m(p_a p_b, q_a q_b)$ is a homogeneous polynomial of degree m in $p_a p_b$ and $q_a q_b$. Note that these homogeneous polynomials of degree m are equal since F is a biformat and $P(A$ wins in $n + 2m$ points) equals $P(B$ wins in

$n + 2m$ points) for all (p_a, p_b) within the equality interval considered. Thus, this biformat has *c.p.r.*

$$\frac{p_a^{\frac{n+r}{2}} q_b^{\frac{n-r}{2}}}{p_b^{\frac{n+r}{2}} q_a^{\frac{n-r}{2}}}$$

and this completes the proof.

Now, as we already know from Chapter 3 which are the most efficient members of the total spectrum of W_n biformats for each of the cases $p_a+p_b \gtrless 1$, it follows that we know the most efficient members of S for each of these cases. Our conclusion is that *amongst the set of biformats whose limiting efficiencies do not drop below 1 along the equality diagonal, the W_n biformats with c.p.r.'s*

$$\frac{p_a q_b^{n-1}}{p_b q_a^{n-1}} \left(\frac{p_a^{n-1} q_b}{p_b^{n-1} q_a} \right)$$

are optimally efficient when $p_a + p_b > 1$ ($p_a + p_b < 1$) for any given n (≥ 2). However, the following theoretical possibility is noted. There may be (although it seems to me to be most unlikely) a biformat which is more efficient at some point (p_a, p_b) (where $p_a \neq p_b$) than any member of the total spectrum of W_n biformats but it can be seen that such a biformat, if it exists, must have an efficiency less than unity "near" the equality diagonal, $p_a = p_b$, and hence cannot be globally optimal.

The structure of the above optimal biformats for the 2-sample binomial test when $\alpha = \beta$ (i.e. the symmetrical case) is such that it can be seen from diagrams such as Figure 3.2.1 that the biformats or tests pass through non-symmetrical situations and yet the 1-step PL (PW) rule remains optimal when $p_a + p_b > 1$ ($p_a + p_b < 1$). These non-symmetrical situations, as a starting point, correspond to the $\alpha \neq \beta$ case in hypothesis testing or the "handicap" situation in sport. Thus, Bellman's (1957) "principle of optimality" (dynamic programming) leads us to conclude that *the PL (PW) rule used in conjunction with $W_{n_2}^{n_1}$ systems has the above optimally efficient property for the 2-sample binomial problem when $\alpha \neq \beta$ provided $p_a + p_b > 1$ ($p_a + p_b < 1$).* An analysis of the asymptotic case (n_1, n_2 large) is given later in Example 6.4.2.

5.2 An Application of Section 5.1 to the Tie-breaker System

Miles (1983), noting the tennis tie-breaker game, has defined the tie-breaker exchange mechanism, TB, (*abbaabb*... or *baabbaa*...) and formed the reversal biformat $W_1(W_nTB^a, W_nTB^b)$. It can be seen that when n is even, the duration of each section of this reversal biformat must be even and we effectively have the $W_{2n}AL$ reversal biformat and hence $\rho = 1$. However, for n odd ($n = 2m - 1$), the probability ratio is no longer constant. For example,

$$\frac{P(\text{player } A \text{ wins in } 4m-2 \text{ points})}{P(\text{player } A \text{ loses in } 4m-2 \text{ points})} = \frac{{p_a}^{2m-1}{q_b}^{2m-1}}{{p_b}^{2m-1}{q_a}^{2m-1}}$$

whereas, it can be shown that

$$\frac{P(\text{player } A \text{ wins in } 4m \text{ points})}{P(\text{player } A \text{ loses in } 4m \text{ points})}$$

$$= \frac{{p_a}^{2m-2}{q_b}^{2m-2}\left\{\left((m-1)p_a p_b + m q_a q_b\right){q_b}^2 + \left(m p_a p_b + (m-1) q_a q_b\right){p_a}^2\right\}}{{p_b}^{2m-2}{q_a}^{2m-2}\left\{\left((m-1)p_a p_b + m q_a q_b\right){q_a}^2 + \left(m p_a p_b + (m-1) q_a q_b\right){p_b}^2\right\}}.$$

Since, for n odd, the probability ratio is not constant, it follows from Section 5.1 that points cannot be equally important when $p_a = p_b$ and hence, using equation (4.5.2), the efficiency of $W_1(W_nTB^a, W_nTB^b)$ at $(p+\delta, p-\delta)$ (δ small) must be less than 1 when n is odd and greater or equal to 3 (cf. Miles (1983) pp.315). This has been computationally verified for the case in which $n = 5$. (In fact, the computations also showed that the restriction, δ small, was not necessary.) In a similar way, suppose we set up the alternating pairs exchange mechanism, AP, ($aa\,bb\,aa\,bb\ldots$ or $bb\,aa\,bb\,aa\ \ldots$) and formed the biformat $W_1(W_nAP^{aa}, W_nAP^{bb})$. In a similar way to above, for n odd the efficiency is 1 whereas for n even and greater or equal to 2 the efficiency at $(p+\delta, p-\delta)$ (δ small) is less than 1. However, it is clear that both these reversal biformats have an asymptotic efficiency of 1 as $n \to \infty$.

5.3 Two Other Bipoints Problems: the Clinical Trials (2-Drugs) Problem and the 2-Armed Bandit Problem

5.3.1. The clinical trials (2-drugs) problem.

A paper by Simon (1977) has a set of references on this problem. The clinical trials problem has two time periods.

(i) a trial period (or hypothesis testing phase) during which testing is carried out to determine (which is more likely to be) the better of the two drugs.

(ii) the remaining period (usually with an infinite horizon) during which the drug evaluated in (i) as the better drug is (the only one) administered as it has been deemed to be the better one.

The time period mentioned in (ii) above has no analogy in the sporting context. Also, the hypothesis testing phase mentioned in (i) above has an added complication in that, during this period, it is required (for ethical reasons) that the poorer drug should ideally be administered no more often than is necessary. This consideration is not involved in the bipoints sporting context. For example, in the case of tennis (when $p_a + p_b > 1$), the test (or biformat) with the optimal efficiency property of Section 5.1 (i.e. the PL case) is such that the expected number of serves of the weaker player is greater than that of the better player. Thus, in the clinical trials problem when $p_a + p_b > 1$, there is a conflict between the desirability of using the most efficient test on the one hand, and the requirement that the poorer treatment be administered no more often than is necessary, on the other hand. There is no such conflict however when $p_a + p_b < 1$, since the test (or biformat) with the optimal efficiency property of Section 5.1 (i.e. the PW case) has a greater expected number of trials with the better drug. Note however that although there is no conflict in this case, it does not necessarily follow that the PW rule is optimal for the clinical trials problem since the formulation of this problem is different to that of the bipoints sporting problem, which is entirely a hypothesis testing problem.

It can be seen that, *in the laboratory situation*, the hypothesis testing phase mentioned in (i) above may not have the added ethical complication that the poorer drug be administered no more often than necessary. In such a situation, the laboratory testing problem is mathematically identical to the bipoints problem and hence Section 5.1 gives a W_n test for this laboratory situation with an optimally efficient property. (It will be seen later (Chapters 8 and 9) that all of the W_n tests of Chapters 2 to 5 have large variances of duration which is a disadvantage in the sporting context. These large variances however may not be a major disadvantage in this laboratory context.)

5.3.2. The 2-armed bandit problem.

A paper by Gittins (1979) has a set of references on this problem. In this problem the gambler has 2 poker machines: one which pays more than the other but which is which is unknown. The gambler's objective is to maximize either the finite horizon expected return or the discounted infinite horizon expected return. This problem thus has a conflict between the hypothesis testing component (which is the better machine to play?) and the desirability of minimizing the number of plays on the poorer machine. It can be seen that, in the sporting context, there is no specific finite horizon for many scoring systems, infinite horizons are not relevant and neither is discounting.

Notes in proof.

A. The duration p.g.f.'s of the members of the total spectrum of biformats have, very recently (December 1986), been derived by R. E. Miles. For the member with c.p.r.

$$\left(\frac{p_a}{p_b}\right)^i \times \left(\frac{q_b}{q_a}\right)^j \qquad (*)$$

it is

$$g_{i,j}(s) = \frac{(p_a^i q_b^j + p_b^i q_a^j)\, s^{i+j}}{\{1-h_+(s)\}^i h_-(s)^j + \{1-h_-(s)\}^i h_+(s)^j} \qquad (**)$$

$(i,j = 1,2, \ldots)$, where

$$h_{\pm}(s) = \tfrac{1}{2}[1 + (q_a q_b - p_a p_b)s^2 \pm \{1 - 2\,(q_a q_b + p_a p_b)s^2 + (q_a q_b - p_a p_b)^2 s^4\}^{\frac{1}{2}}]$$

Despite the simplicity of the result, the derivation was rather extended.

Thus, my theory (Corollary 2.2.2) has shown that, for <u>any</u> biformat with c.p.r. (*), its duration p.g.f. must be (**).

B. The proof of Theorem 4.4 utilized a "forward" induction procedure. Subsequently R. E. Miles has found that the corresponding "backward" induction procedure affords a much simpler proof.

CHAPTER 6
A METHOD FOR DETERMINING THE ASYMPTOTIC EFFICIENCY OF W_N SYSTEMS AND SOME EXAMPLES OF ITS USES

6.1 Introduction

We have seen in Chapter 1 that the efficiency of any uniformat or biformat can be calculated once P and μ have been evaluated. For example, we saw that for $W_{2n}PL$, P and μ and hence ρ could be found, and then by letting $n \to \infty$ in our expression for ρ, we could obtain the asymptotic efficiency of $W_{2n}PL$. *This chapter shows that some complex W_n systems can be decomposed into smaller independent components called "modules", which can in turn be analysed to produce values from which the approximate asymptotic values of P, μ and ρ can be derived.* These independent modules effectively become the steps of a general one-dimensional random walk in discrete time and this approach can be a useful method for obtaining the asymptotic efficiencies of W_n systems when exact P and μ values are intrinsically more difficult to derive for a given n. Examples of such cases will be seen not only in this chapter but also in Chapter 7 when an extension of bipoints to multi-points is considered. In Example 6.3.2 a modules approach is used to give an explanation as to why the PL rule is more (less) efficient than the PW rule when $p_a + p_b > 1$ $(p_a + p_b < 1)$ and, in Section 6.5, the modules approach is also used to determine the best of a set of strategies available to a player.

6.2 The "Module" Method

We make use of the theory of the general one-dimensional random walk in discrete time with two absorbing barriers and the reader is referred to (say) Cox and Miller (1965, p.46–58) for the analysis used in this section. The steps in the random walk, Z_i, are mutually independent random variables on the integers $\ldots, -2, -1, 0, 1, 2, \ldots$. Unlike the simple random walk in which the possible values for Z_i are $-1, 0, +1$, the general random walk can jump over the two absorbing barriers $-b$ and a, and so we define the absorbing states to be $(-\infty, -b]$ and $[a, \infty)$. The moment generating function (m.g.f.) of Z_i is defined by

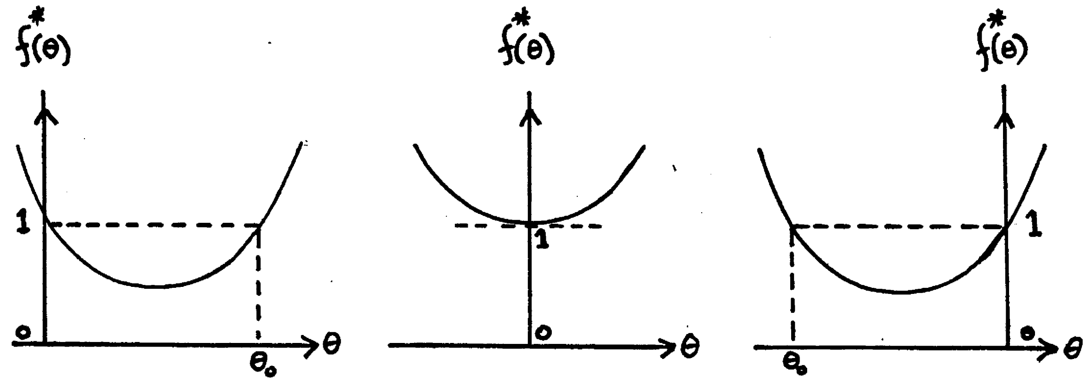
$$f^*(\theta) = \sum_{j=-\infty}^{\infty} e^{-\theta j} P(Z_i = j)$$

and assuming Z_i takes both positive and negative values with positive probability, $f^*(\theta)$ is a convex (downwards) function of θ. Consider the roots of the equation,

$f^*(\theta) = 1$. $\theta = 0$ is clearly one root. If $E(Z_i) \neq 0$ there is a unique second real root $\theta_0 \neq 0$ which has the same sign as $E(Z_i)$. If $E(Z_i) = 0$ then $\theta = 0$ is a double root. Figure 6.2.1 give diagrams for the m.g.f. of Z_i when (a) $E(Z_i) > 0$, (b) $E(Z_i) = 0$ and (c) $E(Z_i) < 0$.

Figure 6.2.1

The moment generating function $f^*(\theta)$ of a random variable Z_i for (a) $E(Z_i) > 0$, (b) $E(Z_i) = 0$ and (c) $E(Z_i) < 0$.



(a) $E(Z_i) > 0$ (b) $E(Z_i) = 0$ (c) $E(Z_i) < 0$

Now if $P(Q)$ represents the probability of absorption in states $[a, \infty)$ $((-\infty, -b])$ and $E(N)$ is the expected number of steps to absorption, then, neglecting the excess over the barriers,

$$\frac{P}{Q} \doteq \begin{cases} \frac{1-e^{\theta_0 b}}{e^{-\theta_0 a}-1} & \text{when } E(Z_i) \neq 0 \\ \frac{b}{a} & \text{when } E(Z_i) = 0 \end{cases} \tag{6.2.1}$$

and

$$E(N) \doteq \begin{cases} \frac{(a+b)-ae^{\theta_0 b}-be^{-\theta_0 a}}{E(Z_i)(e^{-\theta_0 a}-e^{\theta_0 b})} & \text{when } E(Z_i) \neq 0 \\ \frac{ab}{\text{var}(Z_i)} & \text{when } E(Z_i) = 0 \end{cases} .$$

6.3 Examples of the Application of the "Module" Method to Scoring Systems

Example 6.3.1. Consider the $W_{2n}PL$ reversal biformat with $p_a \neq p_b$. Suppose the module is defined as the component from one a-point until the next a-point. Then such modules are clearly mutually independent. In this and the following examples, the random variable S is the increase (or shift) in player A's score during one module and the random variable D is the duration of the module. For this

"a-point until next a-point PL module", S and D have the distributions

$$\begin{cases} P(S=-1)=q_a \\ P(S=i)=p_a p_b {q_b}^i \qquad (i=0,1,2,\dots) \end{cases}$$

and

$$\begin{cases} P(D=1)=q_a \\ P(D=j)=p_a p_b {q_b}^{j-2} \qquad (j=2,3,4,\dots) \end{cases}$$

where $q_a = 1-p_a$ and $q_b = 1-p_b$. Now, noting that S plays the role of Z_i in Section 6.2, the m.g.f. of S is given by

$$f^*(\theta) = q_a e^{\theta} + p_a p_b/(1-q_b e^{-\theta})$$

and the equation $f^*(\theta) = 1$ gives a quadratic equation in e^{θ} with roots 1 and q_b/q_a. It can be shown that

$$\frac{E(D)}{E(S)} = \frac{p_a + p_b}{p_a - p_b} \tag{6.3.1}$$

and, using the results of Section 6.2, it follows that, for $W_{2n}PL$,

$$\frac{P}{Q} \doteq \left(\frac{q_b}{q_a}\right)^{2n}$$

and

$$\mu \doteq \frac{2n(p_a+p_b)}{|p_a-p_b|}$$

where $P(Q)$ is the probability player A wins (loses) and μ is the expected duration of the $W_{2n}PL$ reversal biformat (n large).

Now consider the reversal biformat $W_{2m}AL$. In this case a module is just a pair of points, $\{a,b\}$, and now S has the distribution

$$\begin{cases} P(S=-2)=p_b q_a \\ P(S=0)=p_a p_b + q_a q_b \\ P(S=+2)=p_a q_b \end{cases}$$

and D equals 2. The m.g.f. of S, $f^*(\theta)$, can be written down and the equation $f^*(\theta)=1$ gives a quadratic equation in $e^{2\theta}$ with roots 1 and $p_a q_b/p_b q_a$. Noting

$$\frac{E(D)}{E(S)} = \frac{1}{p_a - p_b},$$

it follows that, for $W_{2m}AL$ (m large),

$$\frac{P'}{Q'} \doteq \left(\frac{p_a q_b}{p_b q_a}\right)^{m}$$

and

$$\mu' \doteq \frac{2m}{|p_a-p_b|}$$

where the dashes in the above expressions refer to the $W_{2m}AL$ biformat. Now, taking the approach of Miles (1984), we require $\frac{P}{Q}$ to equal $\frac{P'}{Q'}$ which gives the equation

$$n \doteq \frac{m \ln \frac{p_a q_b}{p_b q_a}}{2 \ln \frac{q_b}{q_a}} .$$

It follows that the asymptotic efficiency of $W_{2n}PL$ as $n \to \infty$ is given by

$$\rho \doteq \frac{2 \ln \frac{q_b}{q_a}}{(p_a + p_b) \ln \frac{p_a q_b}{p_b q_a}}$$

in agreement with Miles (1984) who derived an explicit expression for the efficiency of $W_{2n}PL$ (n finite) and then let $n \to \infty$ in that expression.

Example 6.3.2. This example gives an intuitive explanation as to why the PL rule is more (less) efficient than the PW rule when $p_a + p_b > 1$ ($p_a + p_b < 1$).

Consider a PL module defined in the following slightly different way to that in Example 6.3.1. Starting with an a-point and using the PL rule, a b-point will at some stage occur; b-points possibly continue until at some stage an a-point will occur for the first time since the first b-point occurred, thus completing the cycle $a \ldots b \ldots a$ and forming the module. This module can be seen to consist of two components, each possessing a geometric distribution structure. Defining S and D in the same way as in Example 6.3.1, it follows that

$$\begin{aligned} E(S) &= \frac{p_a - q_a}{p_a} - \frac{p_b - q_b}{p_b} \\ \mathrm{Var}(S) &= \frac{q_a}{{p_a}^2} + \frac{q_b}{{p_b}^2} \\ E(D) &= \frac{1}{p_a} + \frac{1}{p_b} \\ \text{and} \qquad \frac{E(S)}{E(D)} &= \frac{\delta}{p} \end{aligned}$$

where $p = (p_a + p_b)/2$ and $\delta = (p_a - p_b)/2$. The corresponding PW module has the following properties.

$$\begin{aligned} E(S) &= \frac{p_a - q_a}{q_a} - \frac{p_b - q_b}{q_b} \\ \mathrm{Var}(S) &= \frac{p_a}{{q_a}^2} + \frac{p_b}{{q_b}^2} \\ E(D) &= \frac{1}{q_a} + \frac{1}{q_b} \\ \text{and} \qquad \frac{E(S)}{E(D)} &= \frac{\delta}{q} \end{aligned}$$

where $q = 1 - p$.

Consider the W_m system where $m = nq$ for the PL case and $m = np$ for the PW case and n is large. It follows from the respective expressions above for $\frac{E(S)}{E(D)}$ (or from Theorem 2.3 and its ensuing paragraph) that these two systems have the same asymptotic (n large) expected durations. Now the value of Var(S) for the PL module can be "scaled" so that it can be *directly* compared with the value of Var(S) for the PW module. The purpose of this scaling is to allow for (i) the different values of $E(D)$ for the two modules, and (ii) the different values of m in W_m for the two systems. The appropriately scaled variance for the PL module is

$$\left(\frac{q_a}{{p_a}^2} + \frac{q_b}{{p_b}^2}\right) \times \left(\frac{\frac{1}{q_a} + \frac{1}{q_b}}{\frac{1}{p_a} + \frac{1}{p_b}}\right) \times \left(\frac{p}{q}\right)^2$$

which equals
$$\left(\frac{q_a}{{p_a}^2} + \frac{q_b}{{p_b}^2}\right)\left[\frac{p_a p_b}{q_a q_b} . \frac{p_a + p_b}{q_a + q_b}\right] . \qquad (6.3.2)$$

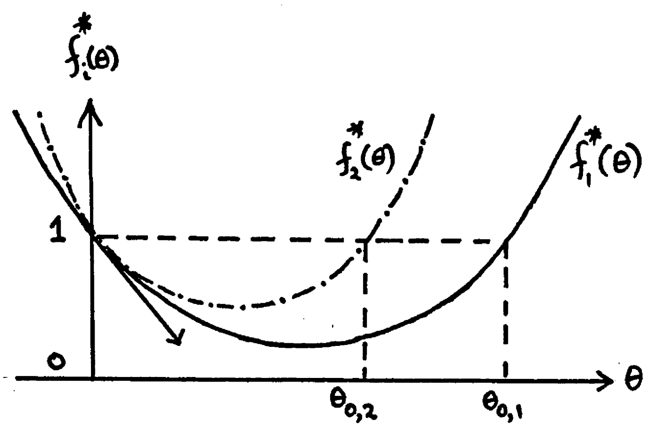
If $p_a \neq p_b$ and $p_a + p_b > 1$, (6.3.2) is less than $\frac{p_a}{{q_a}^2} + \frac{p_b}{{q_b}^2}$, the value of Var($S$) for the PW module. Thus, for the same asymptotic expected durations, the smaller scaled (or "relative") variance for the PL module (when $p_a \neq p_b$ and $p_a + p_b > 1$) results in the better player having a higher probability of winning and hence the PL system has greater asymptotic efficiency. (It may be noted that, if the variance of S could be reduced to zero, the better player would win with probability 1!) It can also be seen that expression (6.3.2) is equal to $\frac{p_a}{{q_a}^2} + \frac{p_b}{{q_b}^2}$ when $p_a = p_b$ or $p_a + p_b = 1$. (This corresponds to the fact that the efficiencies (or limiting efficiencies) of W_m systems when used with PL or PW rules are equal (and equal to 1) along either diagonal within the (p_a, p_b) unit square.) Finally, if $p_a + p_b < 1$ and $p_a \neq p_b$, expression (6.3.2) is greater than $\frac{p_a}{{q_a}^2} + \frac{p_b}{{q_b}^2}$ and hence the PW structure is the more efficient in this case.

The above explanation can be generalized. Suppose we assume without loss of generality that player A is the better of two players. Consider two W_n scoring systems SS_1 and SS_2 which have modules with the same (appropriately scaled) expected rate of progress towards the A-winning (or "correct-decision") boundary (i.e. $E(S_1) = E(S_2)$). Thus, SS_1 and SS_2 have the same expected durations. Suppose these modules have m.g.f.'s $f_1^*(\theta)$ and $f_2^*(\theta)$ respectively. Since, for $i = 1$ or 2, $E(S_i) = \left\{-\frac{df_i^*(\theta)}{d\theta}\right\}_{\theta=0}$, we have that these two m.f.g.'s have the same slope at $\theta = 0$ as shown in Figure 6.3.1. Now suppose $f_2^*(\theta)$ has a greater rate of change of slope at $\theta = 0$ than does $f_1^*(\theta)$. That is, suppose

$$\left.\frac{d^2 f_2^*(\theta)}{d\theta^2}\right|_{\theta=0} > \left.\frac{d^2 f_1^*(\theta)}{d\theta^2}\right|_{\theta=0} > 0 . \qquad (6.3.3)$$

Then, as shown in Figure 6.3.1, it is reasonable to believe that the root $\theta_{0,2}$ of $f_2^*(\theta) = 1$ would typically be less than the root $\theta_{0,1}$ of $f_1^*(\theta) = 1$. Using the first of equations (6.2.1), $\theta_{0,1} > \theta_{0,2}$ implies that SS_1 is more efficient than SS_2 (since they have the same expected durations). However, from (6.3.3) it follows that variance$(S_2) >$ variance(S_1) since $E(S_2) = E(S_1)$. *Thus, given two W_n systems with the same expected rate of progress towards the "correct-decision" boundary (by using appropriate scaling) it is reasonable to believe that the system whose associated (appropriately scaled) module has the smaller variance is typically (at least asymptotically) the more efficient.* This is related to the result for the case in which steps have a normal distribution with mean μ and variance σ^2 (see (say) Cox and Miller (1965), p.50). In this case, θ_0 equals $\frac{2\mu}{\sigma^2}$ indicating that, for a given μ, θ_0 and hence the probability that the better player wins (or the correct decision is made) increases as σ^2 decreases.

Figure 6.3.1
The m.g.f.'s of two random variables with the same (positive) mean
but differing variances and values of θ_0.



Example 6.3.3. Here we consider the module approach to the asymptotic efficiency of the $W_1(W_nPL^a_{even}, W_nPL^b_{even})$ reversal biformat. This rather complex PPL_{even} exchange mechanism can be converted to a 1-step dependent process involving 14 states by noting the following. If at each point being played, a record is kept of the following:-

(i) the previously played AL point (a or b) when the present point being played is a PL point,

(ii) the current type of point being played (aAL, aPL, bAL or bPL),

(iii) whether player A's (and B's) last wins were even or odd ((A_{odd}, B_{odd}), (A_{odd}, B_{even}), (A_{even}, B_{odd}) or (A_{even}, B_{even})), where the interpretation of even and odd is defined in Section 2.3, then, of the apparent 32 possible states, an enumeration shows that only 14 can in fact occur. These states are defined in Table 6.3.1. Note, for example, that state 1 is followed by state 11 or 4 (and state 5, for example, is followed by state 14 or 1) depending on whether player A wins or loses respectively.

Table 6.3.1
The fourteen types of points in the module for the
$W_1(W_nPL^a_{even}, W_nPL^b_{even})$ biformat.

State	(Previous AL state	;	Current point	;	Current Status of A and B wins)
1	(b	;	aAL	;	$A_{odd}B_{odd}$)
2	(b	;	aAL	;	$A_{odd}B_{even}$)
3	(b	;	aAL	;	$A_{even}B_{odd}$)
4	(a	;	aPL	;	$A_{odd}B_{even}$)
5	(b	;	aPL	;	$A_{odd}B_{even}$)
6	(a	;	aPL	;	$A_{even}B_{even}$)
7	(b	;	aPL	;	$A_{even}B_{even}$)
8	(a	;	bAL	;	$A_{odd}B_{odd}$)
9	(a	;	bAL	;	$A_{odd}B_{even}$)
10	(a	;	bAL	;	$A_{even}B_{odd}$)
11	(a	;	bPL	;	$A_{even}B_{odd}$)
12	(b	;	bPL	;	$A_{even}B_{odd}$)
13	(a	;	bPL	;	$A_{even}B_{even}$)
14	(b	;	bPL	;	$A_{even}B_{even}$)

A branching process was devised on a computer, starting with state 1 and continuing along all possible branches until state 1 was revisited for the first time (at the end of every branch), thus creating a (rather complex) module for this PPL_{even} exchange mechanism. The distributions of S and D (as defined in Example 6.3.1) were thus generated for numerous values of p_a and p_b and it was verified numerically that

$$\frac{E(D)}{E(S)} = \frac{1}{2}\left(\frac{p_a + p_b}{p_a - p_b} + \frac{1}{p_a - p_b}\right)$$

and that the roots of the equation $f^*(\theta) = 1$ were $e^\theta = 1$ and $\sqrt[4]{\frac{p_a q_b{}^3}{p_b q_a{}^3}}$. These values may have been anticipated from the results in Example 6.3.1 since we know from Chapter 2 that the PPL_{even} rule can be seen as a mixture of "half" PL and "half" AL. In a similar manner to Example 6.3.1, the asymptotic efficiency of this $W_1(W_n PL^a_{even}, W_n PL^b_{even})$ biformat as $n \to \infty$ is given by

$$\rho = \frac{\ln \frac{p_a}{p_b} + 3 \ln \frac{q_b}{q_a}}{(1 + p_a + p_b)(\ln \frac{p_a}{p_b} + \ln \frac{q_b}{q_a})}$$

which has been shown to agree with the standard approach using equations (2.3.1).

6.4 Examples of the Application of the "Module" Method to Statistical Inference

Example 6.4.1. Here we consider a 2-sample Poisson test. Suppose we have two mutually independent sets of Poisson variables

$$X_{1i} \sim \text{Poisson}(\lambda_1) \quad (i = 1, 2, 3, \ldots)$$
$$X_{2j} \sim \text{Poisson}(\lambda_2) \quad (j = 1, 2, 3, \ldots)$$

and our interest is in testing

$$H_0 : \lambda_1 > \lambda_2$$
$$\text{against} \qquad H_1 : \lambda_1 < \lambda_2 \,.$$

We can set up a "standard module" consisting of one observation from each distribution and, using the same notation as in Example 6.3.1, form $S_i = X_{1i} - X_{2i}$ $(i = 1, 2, 3, \ldots)$. This module has a duration of 2, $E(S_i) = \lambda_1 - \lambda_2$ and the m.g.f. of S_i is given by

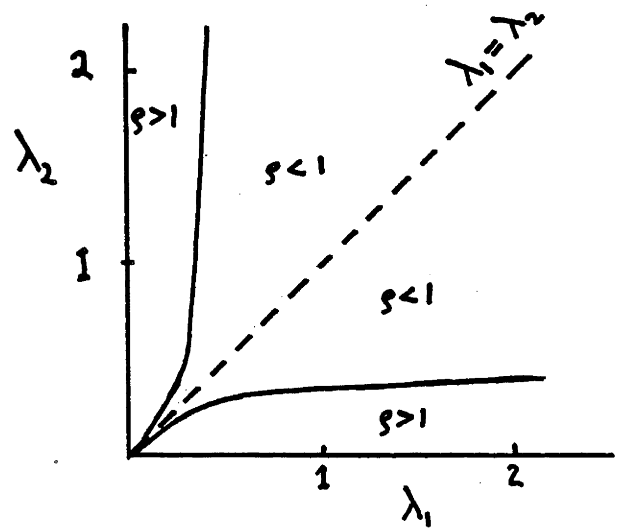
$$f^*(\theta) = e^{\lambda_1(e^{-\theta}-1)+\lambda_2(e^{\theta}-1)} \,.$$

The equation $f^*(\theta) = 1$ gives a quadratic in e^θ with solutions 1 and $\frac{\lambda_1}{\lambda_2}$.

(i) Analogously to the 2-sample binomial case, we define the "PW module" in the following way. The PW module commences with an X_1 observation and independent X_1 observations continue as long as $X_1 \geq 1$; as soon as $X_1 = 0$ this section of the module stops and the second section of the module commences with a $-X_2$ observation; independent $-X_2$ observations continue as long as $-X_2 \leq -1$ and as soon as $-X_2 = 0$, the module stops. A branching process was set up on a computer in order to numerically evaluate $E(D)$, $E(S)$ and θ_0 for specified values of λ_1 and λ_2. Hence, the asymptotic efficiency, under the W_n structure, of the PW module relative to the standard module was evaluated for numerous values of λ_1 and λ_2. Figure 6.4.1 shows the domain in which the PW structure is asymptotically super-efficient. As in the case of the 2-sample binomial test, only very small increases in asymptotic efficiency above unity occur (see Miles (1984)). As can be seen from Figure 6.4.1, the domain of asymptotic super-efficiency is of little practical relevance although it could be argued that the case in which λ_1 and/or λ_2 are small is of some interest.

Figure 6.4.1

The asymptotic efficiency of the PW structure for the 2-sample Poisson test



(ii) Again, analogously to the 2-sample binomial case, we define the "PL module" in the following way. The PL module commences with an X_1 observation and independent X_1 observations continue until an $X_1 \geq 1$ for the first time when this section of the module stops and the second section of the module commences with a $-X_2$ observation; independent $-X_2$ observations continue until $-X_2 \leq -1$, when the module stops. Unlike the PW case, the analysis of this PL module can be carried out explicitly and it can be shown that, for this module,

$$\frac{E(D)}{E(S)} = \frac{2 - e^{-\lambda_1} - e^{-\lambda_2}}{\lambda_1(1 - e^{-\lambda_2}) - \lambda_2(1 - e^{-\lambda_1})}$$

and $$f^*(\theta) = \frac{e^{-(\lambda_1+\lambda_2)}(e^{\lambda_1 e^{-\theta}} - 1)(e^{\lambda_2 e^{\theta}} - 1)}{(1 - e^{-\lambda_1})(1 - e^{-\lambda_2})}.$$

The equation $f^*(\theta) = 1$ has solutions $e^\theta = 1$ and $\frac{\lambda_1}{\lambda_2}$. It follows that the asymptotic efficiency, under the W_n structure, of this PL module relative to the standard module, as $n \to \infty$, is given by

$$\rho = \frac{2\{\lambda_1(1 - e^{-\lambda_2}) - \lambda_2(1 - e^{-\lambda_1})\}}{(\lambda_1 - \lambda_2)\{2 - e^{-\lambda_1} - e^{-\lambda_2}\}}$$

which is less than 1 for all λ_1 and λ_2. Note however that as both λ_1 and λ_2 increase, $\rho \to 1$. This is expected as the PL module approaches the standard module as λ_1 and λ_2 become larger.

Example 6.4.2. We now consider an example in which the probabilities of type I and type II errors are not équal and take the 2-sample binomial hypothesis testing problem

$$H_0 : p_a > p_b$$
$$H_1 : p_a < p_b$$

as the example. The probability of the type I error, α, is defined as $P(H_0$ is rejected$/H_0$ is true) and the probability of the type II error, β, is defined as $P(H_0$ is accepted$/H_0$ is false). So far we have restricted ourselves to cases in which $\alpha = \beta$. However, the case in which $\alpha \neq \beta$ is relevant to handicap scoring systems and statistical inference. Suppose we are interested in the relative efficiency of two scoring systems, SS with absorbing barriers $-b$ and a and SS' with absorbing barriers $-b'$ and a' (a, b, a', b' large). (SS could, for example, be based on the AL module and SS' based on the PL module.) In general however, we suppose that the module approach to the SS system can be taken and θ_0 is the non-zero solution of the corresponding equation $f^*(\theta_0) = 1$. We also suppose that we restrict to modules with the following reasonable property:- if H_0 (H_1) is true, then $\theta_0 > (<) 0$. Then, for the SS system, if H_0 is true, $e^{\theta_0} > 1$ and $\alpha \doteq e^{-b\theta_0}$ whereas if H_1 is true, $e^{\theta_0} < 1$ and $\beta \doteq e^{a\theta_0}$. Now, suppose the module approach can also be taken to the SS' system and let θ_0' be the non-zero solution of the appropriate equation $f^*(\theta_0') = 1$ for this sytem. Then, again restricting to modules with the above property, we have that if H_0 is true, $e^{\theta_0'} > 1$ and $\alpha \doteq e^{-b'\theta_0'}$ whereas, if H_1 is true, $e^{\theta_0'} < 1$ and $\beta \doteq e^{a'\theta_0'}$. It follows that if we require the probabilities of the type I errors to be equal ($\alpha = \alpha'$) and the probabilities of the type II errors to be equal ($\beta = \beta'$), then

$$\frac{\theta_0'}{\theta_0} = \frac{a}{a'} = \frac{b}{b'} \,.$$

Suppose μ (μ') is the expected duration of the SS (SS') system. Then, if H_0 is true,

$$\mu \doteq \frac{aE(D)}{E(S)}$$

$$\text{and} \qquad \mu' \doteq \frac{a'E(D')}{E(S')}$$

where, as in Example 6.3.1, S (S') is the shift and D (D') is the duration of one SS (SS') module. Similarly if H_1 is true,

$$\mu \doteq \frac{-bE(D)}{E(S)}$$

$$\text{and} \qquad \mu' \doteq \frac{-b'E(D')}{E(S')}\cdot$$

Thus, if $\alpha = \alpha'$, $\beta = \beta'$ and H_0 or H_1 is true,

$$\frac{\mu}{\mu'} = \frac{E(D)/(\theta_0 E(S))}{E(D')/(\theta_0' E(S'))}\,. \tag{6.4.1}$$

However μ/μ' is the ratio of the expected durations of the two systems, and hence, as in the definition of efficiency in Chapter 1, since $\alpha = \alpha'$and $\beta = \beta'$, represents the asymptotic efficiency of SS' relative to SS. Note that this relative asymptotic efficiency does not depend on the respective values of α ($= \alpha'$) and β ($= \beta'$) but is a function of characteristics of the associated modules.

Our conclusion is that, *given two modules, the one which produces the asymptotically more efficient W_n system (for the case in which $\alpha = \beta$) is the same one as that which produces the asymptotically more efficient $W_{n_2}^{n_1}$ system (for the case in which $\alpha \neq \beta$).* This result applies to *any* W system (e.g. Poisson, multipoints (see Chapter 7),...) which can be decomposed into modules, where the symbol W is used to cover all W_n and $W_{n_2}^{n_1}$ systems. It can also be seen from equation (6.4.1) that *the asymptotically most efficient of a set of W systems is the one whose associated module has the smallest value of $E(D)/(\theta_0 E(S))$*. For example, it can be verified that in Example 6.3.2 the PL module has a smaller (larger) value of this quantity than does the PW module when $p_a + p_b > 1$ ($p_a + p_b < 1$), by noting that $\theta_0 = \ln(p_a/p_b)$ and $\ln(q_b/q_a)$ for the PW and PL modules respectively.

6.5 Selecting an Optimal Strategy

Suppose the better of two players has a set of strategies available, each of which can be represented by a module. Then, when playing under the W_n (n fixed and large) system, it can be seen from equation (6.2.1) that *the better player* *should select the strategy whose module has the greatest value of θ_0, the non-zero solution of the associated equation $f^*(\theta_0) = 1$.*

Example 6.5.1. Suppose, for example, a tennis player (player A) has 3 types of service; a service to the backhand (B), a service to the forehand (F) and a surprise service (S). Suppose, for simplicity, that the probability that this player wins any such point given the type of the previous point is given by the following matrix of conditional probabilities.

Conditional Probability		Current service B	F	S
Previous service	B	0.695	0.705	0.8
	F	0.705	0.695	0.8
	S	0.6	0.6	0.6

For W_n (n large), what is the best sequence of services available to this player? Some of the available sequences and the values of θ_0 for their associated modules are:-

(i)	B only	$\theta_0 = 0.8236$
(ii)	F only	$\theta_0 = 0.8236$
(iii)	S only	$\theta_0 = 0.4055$
(iv)	$B, F, \ldots$	$\theta_0 = 0.8712$
(v)	$B, S, \ldots$	$\theta_0 = 0.8959$
(vi)	$F, S, \ldots$	$\theta_0 = 0.8959$
(vii)	$B, F, S, \ldots$	$\theta_0 = 0.8874$
(viii)	$F, B, S, \ldots$	$\theta_0 = 0.8874$
(ix)	$B, S, F, S, \ldots$	$\theta_0 = 0.8959$
(x)	$F, B, S, B, \ldots$	$\theta_0 = 0.8832$
(xi)	$S, F, B, F, \ldots$	$\theta_0 = 0.8832$

where the notation $B, F, S, \ldots$ (for example) represents modules of these 3 services in that order (B, F, S; B, F, S; B, F, S; $\ldots$). Amongst this set of 11 sequences, the asymptotically optimal sequences for this player can be seen to be (v), (vi) or (ix). There are of course infinitely many other sequences which can be seen as "combinations" (in full or in part) of any of the above eleven sequences and would be asymptotically sub-optimal (apart from combinations of any of the above optimal ones (v), (vi) and (ix)).

It can be seen that the optimal sequences (v), (vi) and (ix) correspond to the cases in which the "p-value" alternates or cycles through a set of values which are quite different (.8, .6, .8, .6, .8, .6, ...). However, it can also be seen that the sub-optimal sequences (vii) and (viii), for example, have probabilities (.705, 0.8, 0.6; ...) with an average p-value greater than that of the optimal sequences. This may appear to be a surprising result, especially for this asymptotic W_n case. However, a closer examination shows that 6 points (or 3 modules) of the optimal sequences (v) or (vi) have a shift in player A's score with an expected value of 2.4 and a variance of 4.8 whilst 6 points (or 2 modules) of the sub-optimal sequences (vii) or (viii) have values of 2.42 and 4.8638 respectively. It is this larger variance of sequences (vii) and (viii) which leads to their sub-optimality. This variance plays a corresponding role to that in Example 6.3.2. Note that, although the assumption of underlying normality is unreasonable, the formula $\theta_0 = \frac{2\mu}{\sigma^2}$ (Cox and Miller (1965, p.50), Example 6.3.2) gives values for θ_0 of 1 and 0.9951 for the optimal ((v) and (vi)) and sub-optimal ((vii) and (viii)) sequences respectively. Thus, it can be seen that the smaller variance of the optimal sequences more than compensates for the lower expected shift. *Thus, for W_n systems (n fixed and large), it appears that the better player increases the probability of winning by alternating or cycling through a set of p-values which are increasingly different whilst maintaining the same average p-value.* This result leads to two questions. Firstly, can this result also apply to other scoring systems and, secondly, can it apply in the non-asymptotic case? The answers are yes. Consider, for example, B_3 in which a player has the choice of keeping p fixed at 2/3 or allowing 2 of the p-values to equal 1 and the other zero. The second choice with the same average p-value is clearly preferable to this player. This result for B_3 (and the one above for W_n (n large)) is related to Poisson's binomial (see, for example, Moran (1968, p.15–16)) in which the variance of the total number of successes in n trials (n fixed) when the p_i's are not equal is less than that for the corresponding Bernoulli case.

6.6 Further Studies

We have seen in Chapters 2 to 5 that the PL rule leads to the most efficient scoring system when p_a and p_b are *constants* ($p_a \neq p_b$), $p_a + p_b > 1$ and W_n is used. However, the case in which these probabilities are random variables could be examined.

Suppose P_a is a random variable representing player A's probability of winning an a-point and that P_a has mean p_a and variance v_a. Suppose P_b, p_b and v_b are correspondingly defined. It follows that, for the W_n biformats using the PL, PW and AL rules (and many other sensible scoring systems), the probability player A wins is 0.5 when $p_a = p_b$ irrespective of the relative sizes of v_a and v_b. This can be seen from expressions for $E(S)$ for such systems (see Examples 6.3.1 and 6.3.2 for example) and from equation (6.2.1). Thus, such systems remain "fair" even in the event that P_a and P_b are random variables (with possibly different variances).

However, does the PL rule lead to a more efficient scoring system than the PW rule when $p_a + p_b > 1$ and $v_a = v_b$? It quite possibly does. One possible investigation would be to suppose P_a can take on (only) two possible values p_{a_i} with probability p_i $(i = 1, 2)$ and to suppose P_b can correspondingly take on (only) two possible values p_{b_i} with probability p_i $(i = 1, 2)$. The cases in which $|p_{a_1} - p_{a_2}|$ equals $|p_{b_1} - p_{b_2}|$ (so that $v_a = v_b$) could firstly be considered. It is conjectured that, in such cases, the PL rule would lead to an asymptotically more efficient W_n system than does the PW rule provided $p_1(p_{a1} + p_{b1}) + p_2(p_{a2} + p_{b2}) > 1$ and $p_1 p_{a_1} + p_2 p_{a_2} \neq p_1 p_{b_1} + p_2 p_{b_2}$. A tedious branching process/module approach to this problem could be carried out.

It would appear that the conclusions for the cases in which $v_a \neq v_b$ are more complex and could very well depend on the relative sizes of v_a and v_b as well as the values of p_a and p_b. An initial investigation using the above branching process/module approach without the constraint that $|p_{a1} - p_{a2}|$ equals $|p_{b1} - p_{b2}|$ could be carried out to suggest possible conclusions.

6.7 Summary

This chapter has shown that some complex W_n systems could be decomposed into independent components called modules whose characteristics could be used to determine the approximate asymptotic (n large) values of P, μ and ρ. These modules effectively became the steps of a general one-dimensional random walk in discrete time which was discussed in Section 6.2.

Section 6.3 gave three examples of the use of this approach to scoring systems. In particular, the PL and PW modules were defined and analysed in Example 6.3.2 in order to give an intuitive explanation as to why the PL rule is more (less) efficient than the PW rule when $p_a + p_b > 1$ ($p_a + p_b < 1$). The explanation was a particular case of a more general observation: Given two W_n systems with the same expected rate of progress towards the "correct-decision" boundary, it seemed reasonable to believe that the system whose associated (appropriately scaled) module had the smaller variance was (at least asymptotically) the more efficient.

Section 6.4 gave two examples of the application of this module method to statistical hypothesis testing. The first example considered a 2-sample Poisson test. A "standard module" was set up for this situation and, analogously to the 2-sample binomial test, PW and PL modules were also defined. The PL module was shown to always produce an asymptotically less efficient W_n system than the standard module, whilst in some restricted cases the PW module produced an asymptotically more efficient system. These cases appeared to be of little practical relevance. The second example in Section 6.4 considered a case in which the probabilities of type I and type II errors, α and β, were not equal. It was shown that, given two modules, the one which produced the asymptotically more efficient W_n system (for the case in which $\alpha = \beta$) was the same one as that which produced the asymptotically more efficient $W_{n_2}^{n_1}$ system (for the case in which $\alpha \neq \beta$). Also, the asymptotically most efficient of a set of W systems (W_n or $W_{n_2}^{n_1}$) was the one whose associated module had the smallest value of $E(D)/(\theta_0 E(S))$ where these quantities were defined in Sections 6.2 and 6.3.

Section 6.5 considered the problem of a player selecting the optimal strategy from a set of strategies that could be represented by modules. The conclusion was that, when playing under the W_n (n fixed and large) system, the better (poorer) player should select the strategy whose module has the greatest (smallest) value of θ_0, the non-zero solution of the associated m.g.f. equation $f^*(\theta) = 1$. An example was taken in which the variance of the shift in a player's score during one module played a corresponding role to that in Example 6.3.2. Thus, in line with the observation made in paragraph 2 of this section, the following observation was made: For W_n systems (n fixed and large), it appeared that the better player increased the probability of winning by alternating or cycling through a set of p-values which were increasingly different whilst maintaining the same average p-value. This result was also seen to apply to some other scoring systems.

In the next chapter, tennis doubles is seen as a game in which there are essentially four types of point and the module approach set up in this chapter will be seen as a useful tool in the asymptotic analysis of such "multi-points" scoring systems.

CHAPTER 7
AN APPROACH TO THE ASYMPTOTIC EFFICIENCY OF MULTI-POINTS SYSTEMS

7.1 Introduction

In tennis doubles there are essentially 4 types of points a_1, a_2, b_1 and b_2 (corresponding to each of the players' services) and team A is the better team if $p_{a1}+p_{a2} > p_{b1}+p_{b2}$. Extending to N pairs of points, team A is defined to be the better team if $\sum_{i=1}^{N} p_{a_i} > \sum_{i=1}^{N} p_{b_i}$. The rationale behind this definition is that, if a very large (and equal) number of each type of point is played, then, by the law of large numbers, the probability that team A wins more points than team B will tend to 1 if $\sum_{i=1}^{N} p_{a_i} > \sum_{i=1}^{N} p_{b_i}$.

In Section 7.2 the case in which $N = 2$ ("quadri-points") is considered, "point-quads" are set up as the "standard" module and the asymptotic efficiency of a module based on the PL mechanism is evaluated. A single points and a g.p.p. approach are shown to be equivalent although the single points approach is seen to be more useful when extending to the case in which $N \geq 3$ which is considered in Section 7.3. In Section 7.2.3, an alternative module suitable for quadri-points and also based on the PL rule, is investigated and is shown to have an asymptotic efficiency which surprisingly tends to infinity at certain boundary points. In Section 7.4 an extension of unipoints is considered and some limited results based on bipoints and multi-points are stated.

7.2 Quadri-Points

7.2.1. A module using generalized point-pairs.

Following Miles' approach to bipoints, it is natural in this quadri-points case to set up $\{a_1, b_1, a_2, b_2\}$ as the "standard module", called $M1$. Note that any ordering of the four points within the module is allowed. It can be seen that, for this module, the duration is 4 points and, using the notation of Example 6.3.1, $E(S) = 2(p_{a1} + p_{a2} - p_{b1} - p_{b2})$. This module is "fair" since $E(S) = 0$ when $p_{a1}+p_{a2} = p_{b1}+p_{b2}$. Using the module method of Section 6.2, there appears to be no simple formula for the non-zero solution of the associated equation $f^*(\theta_0) = 1$ although, for a given set of values for the parameters p_{a_i} and p_{b_i}, θ_0 can be evaluated numerically.

We now consider a second module, $M2$, using the PL mechanism between teams (PL_{teams}). The first section of the module starts with an (a_1, b_1) point-pair and the PL mechanism operates in the following way. A point-pair win to team A, (AA), is followed by a (b_1, b_2) point-pair whilst a point-pair win to team B, (BB), is followed by an (a_1, a_2) point-pair. This process stops as soon as a draw, (AB), occurs and the second and final section of the module begins. This section starts with an (a_2, b_2) point-pair and proceeds in the same manner as above until a draw is reached, thus completing the module. It can be seen that this module has the property that the number of a_1-points played equals the number of a_2-points played. A corresponding property also holds for the b-points. This $M2$ module was analysed numerically using a branching process on a computer and, again using the same notation as in Example 6.3.1, it was verified that

$$\frac{E(D)}{E(S)} = \frac{\overline{p}_a + \overline{p}_b}{\overline{p}_a - \overline{p}_b} \tag{7.2.1}$$

where $\overline{p}_a = \frac{p_{a1}+p_{a2}}{2}$ and $\overline{p}_b = \frac{p_{b1}+p_{b2}}{2}$. Equation (7.2.1) may have been anticipated from equation (6.3.1). It can also be seen from equation (7.2.1) that this module is "fair". As with the $M1$ module, no simple formula for θ_0 appears to exist although its value was calculated numerically for several parameter sets. It was verified that the value of θ_0 and that of $E(D)/E(S)$ (and hence ρ (see equation (6.4.1))) were unaffected by reversing the roles of b_1 and b_2; namely the case in which the first section of the module started with an (a_1, b_2) point-pair and the second section with an (a_2, b_1) point-pair.

Thus, using the same approach to efficiency as that used in Chapter 6, the asymptotic efficiency of the W_n system (n large) using the $M2$ module relative to that of the $M1$ module (with an efficiency defined to be 1) was evaluated numerically for numerous parameter sets. Assuming without loss of generality that $p_{a1} \geq p_{a2}$ and $p_{b1} \geq p_{b2}$, then, using the parameterization

$$\begin{aligned} p_{a1} &= p_a + \delta_a \\ p_{a2} &= p_a - \delta_a \\ p_{b1} &= p_b + \delta_b \\ p_{b2} &= p_b - \delta_b \end{aligned}$$

the conclusions for the asymptotic efficiency using the $M2$ module and the W_n system (relative to that using the $M1$ module) were:-

(i) When $\delta_a = \delta_b = 0$, the asymptotic efficiency is of course equal to that of the biformat $W_{2n}PL$ as $n \to \infty$.

(ii) For the case in which team A is the better team, the asymptotic efficiency increases as δ_a increases (for a given δ_b). Also, for this case and for a given δ_a, the asymptotic efficiency decreases as δ_b increases. For the case in which team B is the better team, the asymptotic efficiency increases as δ_b increases (for a given δ_a) and, for a given δ_b, decreases as δ_a increases. These conclusions are related to the conclusion in Section 6.5.

(iii) When $\delta_a = \delta_b = \delta > 0$, the asymptotic efficiency is slightly greater than in the bipoints case ($\delta_a = \delta_b = 0$) if $p_a + p_b > 1$, and is slightly less than in the bipoints case if $p_a + p_b < 1$.

(iv) Figures 7.2.1a, 7.2.1b and 7.2.1c show, for three values of (p_{b1}, p_{b2}), the domains in which the asymptotic efficiency using the $M2$ module is greater or less than unity.

Figure 7.2.1a

A diagram showing the regions in which the asymptotic efficiency of W_n using the $M2$ module is greater or less than unity when $(p_{b1}, p_{b2}) = (0.7, 0.5)$.

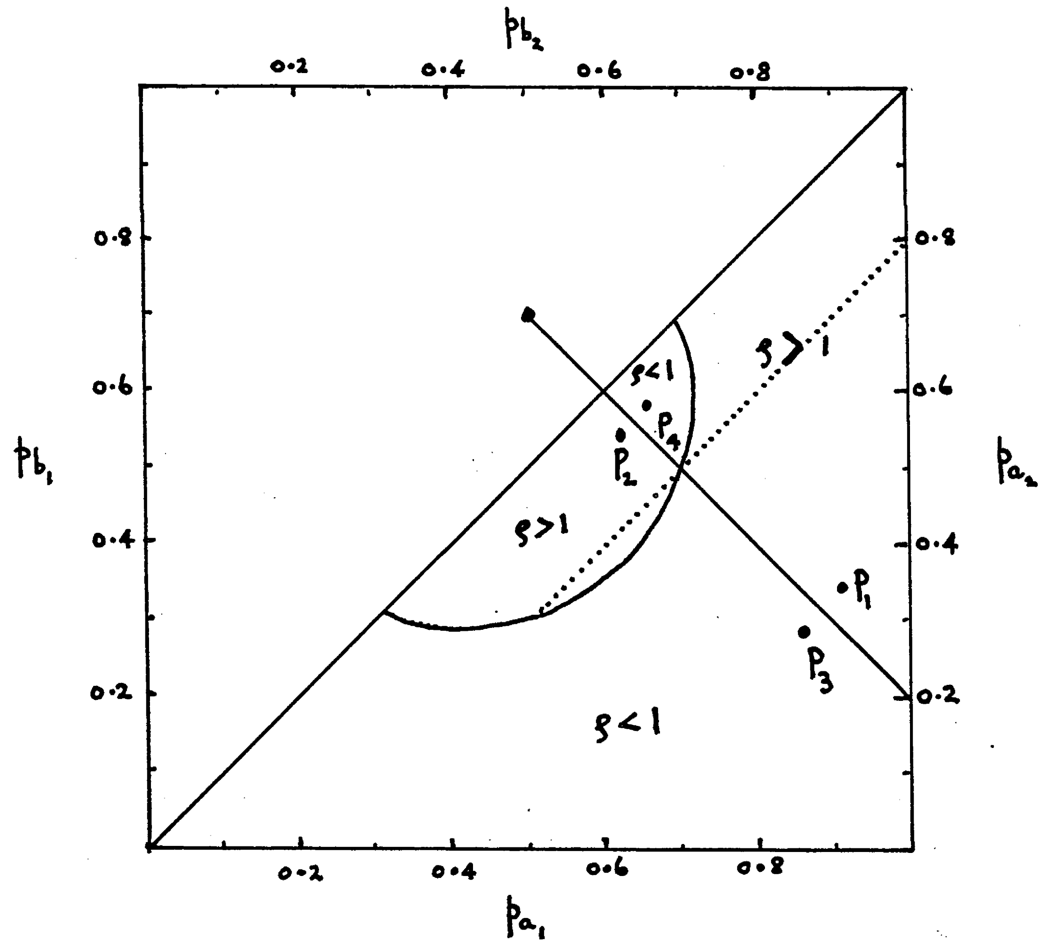


Figure 7.2.1b
A diagram showing the regions in which the asymptotic efficiency of W_n using the $M2$ module is greater or less than unity when $(p_{b1}, p_{b2}) = (0.5, 0.3)$.

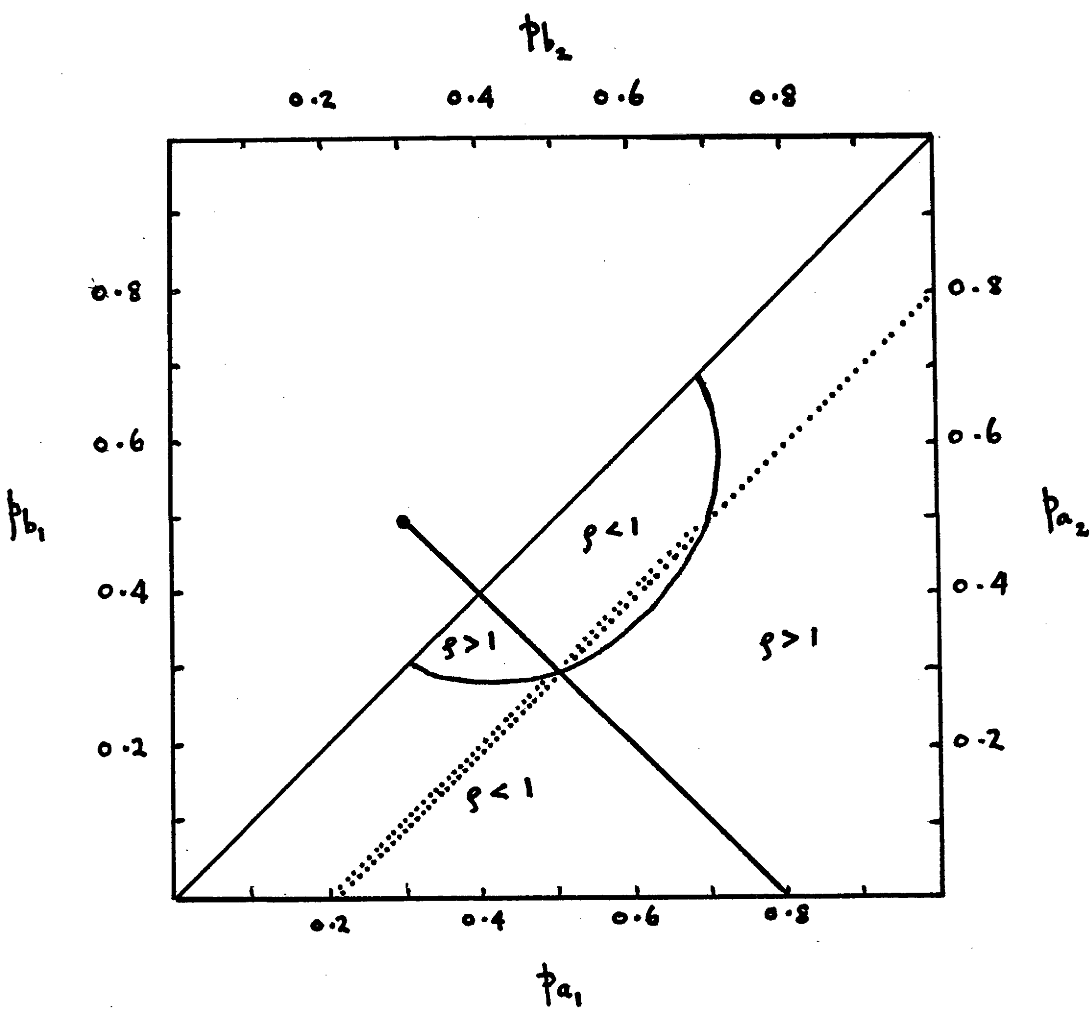
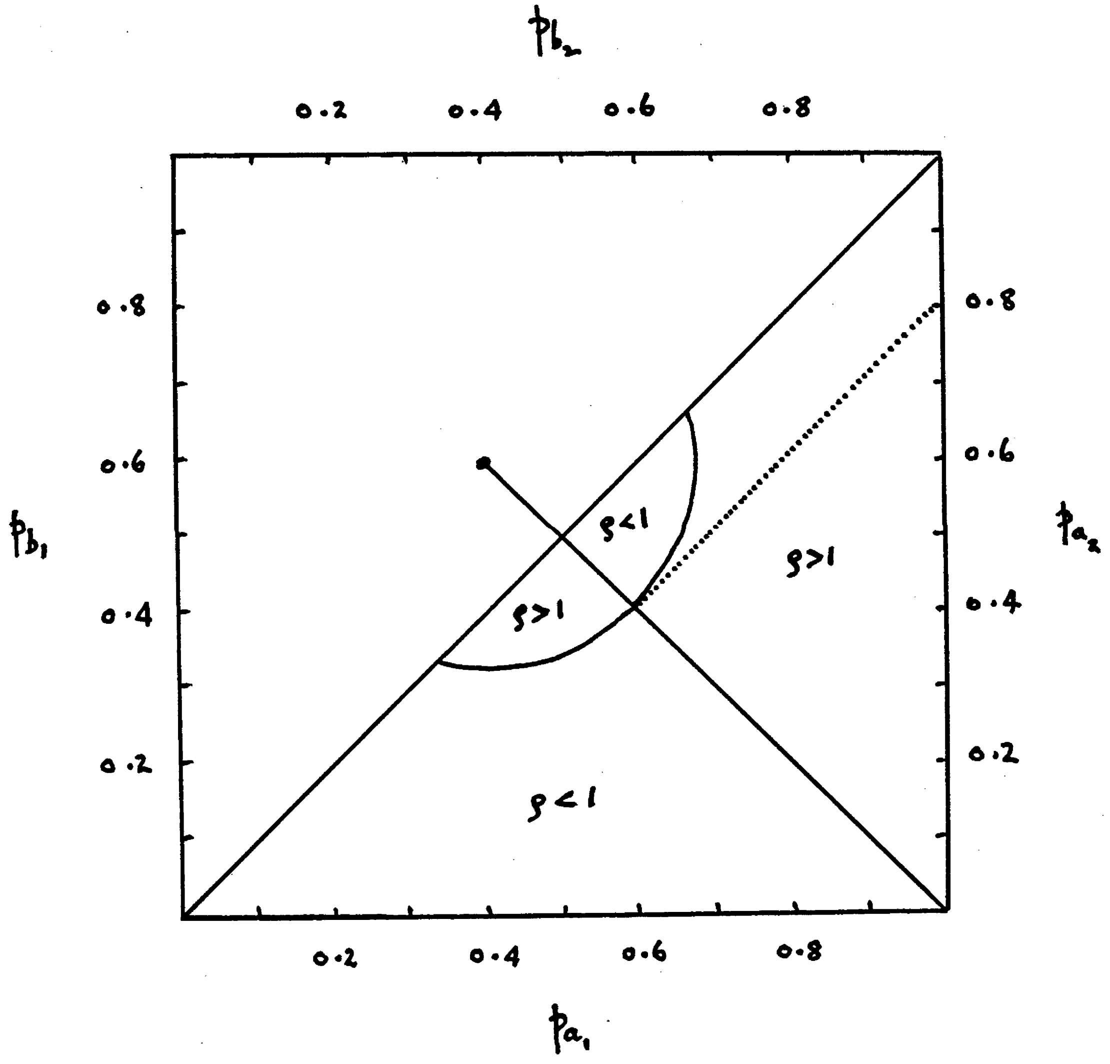


Figure 7.2.1c

A diagram showing the regions in which the asymptotic efficiency of W_n using the $M2$ module is greater or less than unity when $(p_{b1}, p_{b2}) = (0.6, 0.4)$.



As in the case of bipoints, only minute increases in efficiency above unity typically occur. It was also shown that, given any (p_{a1}, p_{a2}) and (p_{b1}, p_{b2}), the asymptotic efficiency of the $M2$ module (using the PL structure) is greater (less) than unity if and only if that of the associated module using the PW structure is less (greater) than unity.

It was noted in (iii) above that, if the differences δ_a and δ_b are equal and if $p_a + p_b > 1$, the asymptotic efficiency using the $M2$ module is always greater than unity. Such cases (δ's equal and $p_a + p_b > 1$) are displayed by the single-dotted lines in Figures 7.2.1a, b and c. Thus, in this restricted case there is increased asymptotic efficiency in using the $M2$ module in preference to the $M1$ module. We now consider the cases in which $\delta_a \neq \delta_b$ and $p_a + p_b > 1$. In this case, if there is only a little difference between p_a and p_b and a big difference between δ_a and δ_b, the following conclusions can be seen from Figure 7.2.1a.

$(i)'$ If the δ corresponding to $\max(p_a, p_b)$ is the greater of the two δ's, the asymptotic efficiency of the $M2$ module is greater than unity (see points P_1 and P_2 in Figure 7.2.1a for example).

$(ii)'$ If the δ corresponding to $\max(p_a, p_b)$ is the smaller of the two δ's, the asymptotic efficiency of the $M2$ module is less than unity (see points P_3 and P_4 in Figure 7.2.1a for example).

Thus, since it is unlikely that there would be *a priori* knowledge about the relative sizes of the δ's, it appears that, in most cases, no overall advantage can be taken of the PL or PW mechanisms.

Corresponding conclusions to those above can be made from Figure 7.2.1b. For example, if the differences δ_a and δ_b are equal and if $p_a + p_b < 1$, the asymptotic efficiency using the $M2$ module is always less than unity (see the double-dotted line in Figure 7.2.1b for example) and hence the corresponding PW module has asymptotic efficiency greater than unity in this case.

Thus, in general, it is unlikely that the PL or PW rule can be used in practice to increase asymptotic efficiency, as, not only do we need to know whether $p_a + p_b$ is greater or less than unity, but we also need information on the relative sizes of $|p_{a1} - p_{a2}|$ and $|p_{b1} - p_{b2}|$. However, if we are prepared to assume that $|p_{a1} - p_{a2}|$ equals $|p_{b1} - p_{b2}|$, we have a direct extension of the asymptotic super-efficient bipoints system to this quadri-points case: The $M2$-type module employing the PL_{team} (PW_{team}) structure used in conjunction with W_n systems has asymptotic efficiency greater than unity when $p_a + p_b > 1$ ($p_a + p_b < 1$).

7.2.2. A module using single points.

The $M2$ module used generalized point-pairs. However, a corresponding PL module using points rather than point-pairs, and denoted by $M3$, can be formed. This is done by *using the PL mechanism between teams but alternating the points within each team.* Figure 7.2.2 is a realization using this exchange mechanism which, though complex, can be converted to a 1-step dependent process (in a similar manner to that of Example 6.3.3) by noting the following. If, at each point being played, a record is kept of:-

(i) the current point to be played (a_1, a_2, b_1 or b_2), and

(iia) if the current point to be played is an a-point (a_1 or a_2), the next type of b-point to be played (b_1 or b_2), or

(iib) if the current point to be played is a b-point (b_1 or b_2), the next type of a-point to be played (a_1 or a_2),

then there are 8 possible states. Now suppose the notation (x_i, y_j) is used to represent the state in which x_i is the current point to be played ($x_i = a_1, a_2, b_1$ or b_2) and y_j is the next point of the other type ($y_j = b_1$ or b_2 if x_i is a_1 or a_2; or $y_j = a_1$ or a_2 if x_i is b_1 or b_2) to be played. By starting the process in state (a_1, b_1) and observing the process (only) after every two points, it can be verified that only 4 states are in fact possible ($(a_1, b_1), (a_2, b_2), (b_1, a_1)$ or (b_2, a_2)) and that a return to state (a_1, b_1) is not possible at the intervening (non-observed) stages. A branching

Figure 7.2.2

A realization based on the $M3$ module.

Type of Point	. .	a_1	b_1	a_2	a_1	a_2	b_2	b_1	b_2	a_1	b_1	.
Winner	. .	A	B	B	B	A	A	A	B	A	.	.

process was devised on a computer, starting with state (a_1, b_1) and observing the process after every two points until a return to state (a_1, b_1) occurred, thus forming the $M3$ module. It was verified that the value of $E(D)/E(S)$ is again given by equation (7.2.1) and e^{θ_0} is the square root (since points rather than point-pairs are now involved) of its value for the $M2$ module, as may have been anticipated. *Thus, from the point of view of asymptotic efficiency, $M2$ and $M3$ are equivalent. However, the advantage of the $M3$ approach is that it is suitable for extension to multi-points.*

7.2.3. A module suitable in a restricted domain.

A further module is available for the specific case in which, using the notation of Section 7.2.1, δ_a and δ_b are known to be equal. Thus $p_{a1} = p_a + \delta$, $p_{a2} = p_a - \delta$, $p_{b1} = p_b + \delta$ and $p_{b2} = p_b - \delta$, and we are comparing two teams constrained to lie somewhere along two lines parallel and equidistant from the main diagonal (see Figure 7.2.3). If none of the probabilities p_{a1}, p_{a2}, p_{b1}, p_{b2} are zero, then the module, $M4$, given by

$$\left\{ \begin{array}{l} \text{play } a_1\text{-points until } A \text{ wins; play } b_1\text{-points until } B \text{ wins;} \\ \text{play } a_2\text{-points until } A \text{ wins; play } b_2\text{-points until } B \text{ wins;} \end{array} \right\}$$

is "fair" and thus is available for use in this restricted case. Noting that this module is in essence a set of four geometric distributions, then, using the same notation as earlier, it can be shown that

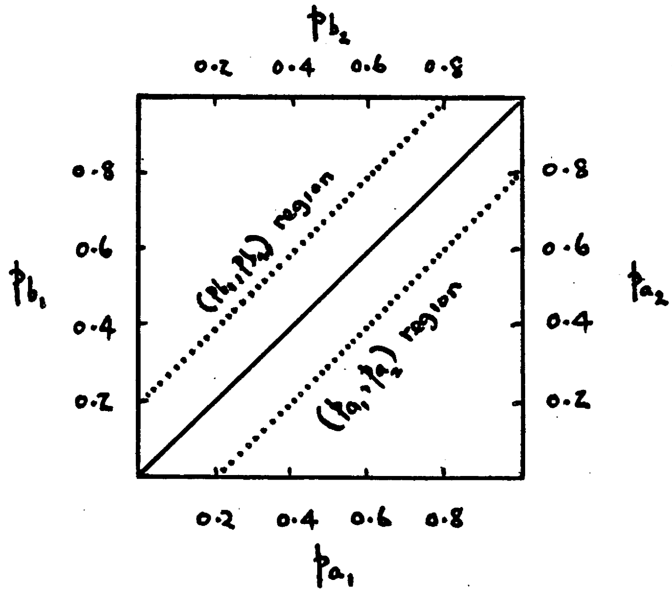
$$E(S) = \frac{1}{p_{b1}} + \frac{1}{p_{b2}} - \frac{1}{p_{a1}} - \frac{1}{p_{a2}}$$

$$E(D) = \frac{1}{p_{b1}} + \frac{1}{p_{b2}} + \frac{1}{p_{a1}} + \frac{1}{p_{a2}}$$

and $$f^*(\theta) = \prod_{i=1}^{2} \frac{p_{a_i} p_{b_i}}{(1 - q_{a_i} e^{\theta})(1 - q_{b_i} e^{-\theta})} .$$

Figure 7.2.3

A diagram showing a restricted region in which the module $M4$ is allowable (fair).



Again there appears to be no explicit formula for the solution of the equation $f^*(\theta) = 1$. However, by solving the equation numerically, the asymptotic efficiencies given in Table 7.2.1 were obtained. It was noted that:-

(i) When $\delta = 0$ the asymptotic efficiency was equal to that of the biformat $W_{2n}PL$ as $n \to \infty$.

Table 7.2.1

The asymptotic efficiency of the $M4$ module for some values of (p_{a1}, p_{a2}) and (p_{b1}, p_{b2}) when $\delta = 0.1$.

(p_{a1}, p_{a2}),	(p_{b1}, p_{b2})	ρ	(p_{a1}, p_{a2}),	(p_{b1}, p_{b2})	ρ
(.8,.6),	(.7,.5)	0.920	(.61,.41),	(.301,.101)	0.997
(.9,.7),	(.7,.5)	0.888	(.6,.4),	(.301,.101)	1.003
(1,.8),	(.7,.5)	0.763	(.5,.3),	(.301,.101)	1.076
			(.401,.201),	(.301,.101)	1.190
(.7,.5),	(.6,.4)	0.949	(.33,.13),	(.301,.101)	1.328
(.8,.6),	(.6,.4)	0.928	(.302,.102),	(.301,.101)	1.410
(.9,.7),	(.6,.4)	0.891	(.3011,.1011),	(.301,.101)	1.413
			(.30101,.10101),	(.301,.101)	1.417
(.6,.4),	(.5,.3)	0.981			
(.7,.5),	(.5,.3)	0.957	(.63,.43),	(.221,.021)	0.966
(.8,.6),	(.5,.3)	0.931	(.53,.33),	(.221,.021)	1.080
			(.43,.23),	(.221,.021)	1.273
(.5,.3),	(.4,.2)	1.037	(.33,.13),	(.221,.021)	1.694
(.6,.4),	(.4,.2)	0.995	(.23,.03),	(.221,.021)	3.669
(.7,.5),	(.4,.2)	0.960	(.2211,.0211),	(.221,.021)	4.318
(.59,.39),	(.4,.2)	0.998			
(.58,.38),	(.4,.2)	1.002	and when $\delta = 0$,		
(.4001,.2001),	(.4,.2)	1.096	(.7,.7),	(.6,.6)	1.002

(ii) Asymptotic efficiencies greater than 1 were obtained for small (not large as in the case of the $M2$ module) values of p_{a1}, p_{a2}, p_{b1}, and p_{b2}. The $M4$ module is thus not directly relevant to tennis.

(iii) For $|p_{a2} - p_{b2}| = \epsilon > 0$, the asymptotic efficiency, $\rho \to \infty$ as $\min(p_{a2}, p_{b2}) \to 0$ (cf. in bipoints, the largest asymptotic efficiency obtainable is 2 (Miles (to appear))). The explanation of this fact is that, as $\min(p_{a2}, p_{b2}) \to 0$, the $M4$ module is "dominated" by playing a very large number of those points corresponding to $\min(p_{a2}, p_{b2})$. If these points are a_2 (b_2) points, the weaker team must be A (B) and the better team wins all (but 1) of this very large number of a_2 (b_2) points.

(iv) For $|p_{a1} - p_{b1}| = \epsilon > 0$, the asymptotic efficiency of the corresponding PW module, $\rho \to \infty$ as $\max(p_{a1}, p_{b1}) \to 1$. A corresponding explanation to that in (iii) above applies.

7.3 Multi-points

The $M3$ module of Section 7.2.2 can be extended to handle multi-points by using the PL mechanism between teams, but rotating or cycling the points within each team. Using the branching process approach of Section 7.2.2, the case $N = 3$ was analysed, and Table 7.3.1 shows the asymptotic efficiency of W_n (n large) using this extension of the $M3$ module for several values of the parameters (p_{a1}, p_{a2}, p_{a3}) and (p_{b1}, p_{b2}, p_{b3}). (These efficiencies are relative to the natural extension of the $M1$ module defined in Section 7.2.1.) It was verified that equation (7.2.1) was true for this case ($N = 3$) and thus this extension of the $M3$ module to the case in which $N = 3$ is "fair". A close look at Table 7.3.1 reveals that corresponding conclusions to those of (i), (ii) and (iii) on pages 7.2 and 7.3 (for the case $N = 2$) can be made. However it appears that no useful or practical conclusions can in general be made for this case in which $N = 3$.

It also appears that it is not possible to find an extension (when $N > 2$) of the $M4$ module which is in general "fair".

7.4 An Extension of Unipoints

Consider the game of squash which is typically considered to be a unipoints game. Now let us suppose that there is a right service box and left service box effect and that player A's probability of winning points beginning in such a way is p_1 and p_2 respectively. Extending to N such points, player A is defined to be the better player if $\sum_{i=1}^{N} p_i > \frac{N}{2}$ and we now wish to test

$$H_0 : \sum_{i=1}^{N} p_i > \frac{N}{2}$$

$$\text{against} \qquad H_1 : \sum_{i=1}^{N} p_i < \frac{N}{2} .$$

7.4.1. The case in which $N = 2$.

Here we have two types of points corresponding to p_1 and p_2. Recall that in bipoints, if we had *a priori* information that $p_a + p_b > 1$ ($p_a + p_b < 1$) then the

Table 7.3.1
The asymptotic efficiency of the W_n (n large) system using the extension of the $M3$ module to the case in which $N = 3$.

(p_{a1}, p_{a2}, p_{a3}),	(p_{b1}, p_{b2}, p_{b3})	ρ
(.7,.7,.7),	(.6,.6,.6)	1.002
(.7,.7,.7),	(.7,.6,.5)	1.001
(.8,.7,.6),	(.6,.6,.6)	1.003
(.8,.7,.6),	(.7,.6,.5)	1.002
(.9,.7,.5),	(.6,.6,.6)	1.007
(.7,.7,.7),	(.8,.6,.4)	0.997
(.9,.7,.5),	(.8,.6,.4)	1.003
(.9,.6,.6),	(.6,.6,.6)	1.006
(.7,.7,.7),	(.8,.5,.5)	0.999
(.9,.6,.6),	(.8,.5,.5)	1.003
(.8,.8,.5),	(.6,.6,.6)	1.006
(.7,.7,.7),	(.7,.7,.4)	0.998
(.8,.8,.5),	(.7,.7,.4)	1.002
(.9,.9,.3),	(.6,.6,.6)	1.020
(.7,.7,.7),	(.8,.8,.2)	0.953
(.9,.9,.3),	(.8,.8,.2)	0.975
(1,.55,.55),	(.6,.6,.6)	1.012
(.7,.7,.7),	(.9,.45,.45)	0.994
(1,.55,.55),	(.9,.45,.45)	1.006

PL (PW) rule could be used to produce a super-efficient W_n system. Now, in this case, suppose it is known *a priori* that (say) $p_1 > \frac{1}{2}$ and $p_2 < \frac{1}{2}$. Then, the above hypotheses when $N = 2$ are equivalent to

$$H_0 : p_1 > q_2$$

against $$H_1 : p_1 < q_2$$

where $q_2 = 1 - p_2$ and, since $p_1 + q_2 > 1$, our knowledge of optimal biformats

(Chapters 2 to 5) tells us that the module

$$\left\{\begin{array}{l}\text{play type 1 points until player } A \text{ wins one;}\\ \text{play type 2 points until player } B \text{ wins one}\end{array}\right\}$$

used in association with W_n is optimal with efficiency greater than unity.

7.4.2. The case in which $N = 3$.

A result of the reasonable requirement that modules should be "fair" is that there appears to be no viable alternative to the "standard point-triplet module"

$$\{\text{a type 1 point; a type 2 point; a type 3 point}\}$$

in general. Suppose (and this is a highly specific case) we knew that $p_3 = \frac{1}{2}$ and (say) $p_1 > \frac{1}{2}$ and $p_2 < \frac{1}{2}$. Then the optimal procedure is to play no type 3 points at all and use the optimal test of Section 7.4.1 for the case in which $N = 2$.

This lack of a viable alternative to the "standard module" appears to apply in general when N is odd.

7.4.3. The case in which $N = 4$.

Again, in very limited circumstances, super-efficient systems can be produced. Suppose, for example, it is known that (say) p_1 and p_2 are each greater than 0.5 (or $p_1 + p_2 > 1$) whilst p_3 and p_4 are each less than 0.5 (or $p_3 + p_4 < 1$). Then, we wish to test

$$H_0 : p_1 + p_2 > q_3 + q_4$$

against $\quad H_1 : p_1 + p_2 < q_3 + q_4$

where $q_i = 1 - p_i$. Making use of Section 7.2.2, we can see that the module

$$\left\{\begin{array}{l}\text{If team } A \text{ loses a type 1 or type 2 point, play another type 1 or type 2}\\ \quad\text{point alternating. Else play a type 3 or type 4 point alternating;}\\ \text{If team } B \text{ loses a type 3 or type 4 point, play another type 3 or type 4}\\ \quad\text{point alternating. Else play a type 1 or type 2 point alternating.}\end{array}\right\}$$

used in conjunction with W_n (n large), is super-efficient if it is known that $|p_1 - p_2|$ equals $|p_4 - p_3|$. Suppose, however, that it is known that (say) $p_1 > \frac{1}{2}$, $p_2 > \frac{1}{2}$, $p_3 > \frac{1}{2}$ and $p_4 < \frac{1}{2}$ (and that the pairwise sums of the p_i's cannot be arranged such that one sum is greater than 1 and the other sum is less than 1, as above), then, as in Section 7.4.2, no viable fair super-efficient system appears to be available.

7.5 Summary

Quadri-points arose quite naturally when considering tennis doubles. Another example of quadri-points is tennis singles with a service court (deuce or advantage side) effect. An example of multi-points with $N = 4$ is tennis doubles with a service court effect or tennis singles with both a service court and end effect.

Quadri-points have been studied in Section 7.2 using the module approach and it has been shown that, if $|p_{a_1} - p_{a_2}|$ equals $|p_{b_1} - p_{b_2}|$, the PL (PW) rule used in conjunction with W_n systems has an asymptotic efficiency greater than unity if $\overline{p}_a + \overline{p}_b > 1$ ($\overline{p}_a + \overline{p}_b < 1$). This has been demonstrated using two different (but equivalent (in terms of efficiency)) modules, one using points and the other using generalized point-pairs. However, the restriction $|p_{a_1} - p_{a_2}| = |p_{b_1} - p_{b_2}|$ is rather limiting and means that this result is of little practical value. A third module based on the PL rule and fair only when $|p_{a_1} - p_{a_2}| = |p_{b_1} - p_{b_2}|$ was considered in Section 7.2.3, and although it had some interesting properties (see page 7.8), was again of little or no practical relevance to the sports scoring situation.

A module approach to multi-points was seen to be possible in Section 7.3 and the particular case in which $N = 3$ was considered. Although efficiencies could be evaluated, it again appeared that no useful or practical conclusions were available for the multi-points case in which $N \geq 3$.

Section 7.4 showed how some of the results in bipoints and quadri-points could be used (in a very limited way) in an extension of unipoints.

This chapter has shown that a module approach (using modules based on the PL and PW rules) is possible in the multi-points situation and some results were achieved. However, the practical relevance of the results was unfortunately seen to be very minimal.

CHAPTER 8
AN ANALYSIS OF TENNIS

8.1 Introduction

In recent years, in almost all tennis matches played, there has been a change from "classical" or "advantage" sets (in which a player must win at least 6 games and win by a margin of 2 games) to the new "tie-breaker" sets (in which, if the games' score reaches 6–6, the 12-point tie-breaker game is played). Sections 8.2 to 8.8 (see also Pollard (1983a)) compare the best-of-three "classical" sets with the best-of-three "tie-breaker" sets, the scoring system now used in most tournaments. Section 8.9 proposes a new method for scoring in tennis and is part of another paper (Pollard, 1987). This new method of scoring has a smaller variance of duration.

In earlier work, Kemeny & Snell (1960) modelled a single game of tennis using a Markov chain. Hsi & Burych (1971), taking a bipoints approach, evaluated the probability that one player wins a set of classical tennis. Carter & Crews (1974), taking a unipoints approach, found the expected number of points in a game, the expected number of games in a set and the expected number of sets in a match. The product of these three values does not, of course, give the expected number of points in a match. Miles (1984), setting up the bipoints approach, found the expected duration of a set of tennis.

Throughout this chapter, it is assumed that player A has a constant probability, p_a, of winning any point while he is serving and that player B has a constant probability, p_b, of winning any point on his service. It is also assumed that the outcome of any point, game or set is independent of the outcome of any other point, game or set. These assumptions are based on Pollard (1980). In analysing 35 matches (5503 points) in the 1977 N.S.W. Mens Open Grass Championships, he found that player A, the winner of the match, on average won 71% of points on his service whereas his opponent won 62% on his service. He applied a Wald-Wolfowitz runs test for randomness in the sequences of wins and losses in points, one-step dependency tests to see whether the outcome on a point was dependent on the outcome of the previous point, and state dependency tests to see whether the outcome on a point was dependent on whether the player was ahead or behind at that stage. These analyses made use of the normal approximation to the difference between two sample proportions and of sign tests. His conclusions were in agreement with the above assumptions. In this chapter, it is shown that if $p_a > p_b$, the probability that player A wins a classical set or match is greater than the corresponding probability for a tie-breaker set or match. Also, the mean and variance of the number of points in a tie-breaker set or match is always less than that for the classical version. This mean and variance for the classical version is also shown to increase rapidly as both p_a and p_b increase, whereas they remain relatively stable for the tie-breaker version.

8.2 The Analysis of a Single Game of Tennis

Suppose player A is serving in a game and that he has a probability p_a ($q_a = 1 - p_a$) of winning (losing) any point in the game. The game is won by the player who scores 4 points before the other player scores more than 2 points. If the points' score reaches 3 points each (3/3 or "deuce") the player who first wins 2 more points than his opponent wins the game. Given player A is serving, suppose P_a denotes the probability that player A wins one of his own service games. Then,

$$P_a = \sum_{i=0}^{2} \binom{3+i}{i} {p_a}^4 {q_a}^i + 20{p_a}^3{q_a}^3 \sum_{i=0}^{\infty} {p_a}^2 (2p_a q_a)^i \, .$$

Table 8.2.1

Some characteristics of a game of tennis

	p_a		
	0.5	0.6	0.7
P_a	0.50	0.74	0.90
μ_a	6.75	6.48	5.83
σ_a^2	7.69	6.71	4.40

By noting the individual elements within the above summations, the conditional distribution of the duration of a service game of player A given that he won that game can be determined. This conditional distribution has mean $\mu_{a,1}$ and variance $\sigma_{a,1}^2$. The conditional distribution given that he lost the game can also be generated (mean $\mu_{a,0}$ and variance $\sigma^2_{a,0}$). Table 8.2.1 shows the unconditional mean and variance of duration (μ_a and σ^2_a) for three values of p_a.

8.3 The Mean and Variance of the Duration of a Set of Classical Tennis

In all the equations of Sections 8.3 to 8.8, it is assumed that player A serves in the first game of the first set. Service alternates from one game to the next. A classical set of tennis is completed when a player wins 6 games before his opponent wins more than 4 games. If the games' score reaches 5 games each (5/5), the winner of the set is the first player to win 2 more games than his opponent.

Suppose player A (B) has a probability P_a (P_b), as derived in Section 8.2, of winning any of his service games. Then, the probability that player A wins the set by 6 games to k games ($k = 0, 2$, or 4), $P_{6/k}$, is given by

$$P_{6/k} = \sum_{i=(6-k)/2}^{(6+k)/2} P(k,i)$$

where

$$P(k,i) = Q_b b\left(\frac{6+k}{2},\, i,\, P_a\right) b\left(\frac{4+k}{2},\, 5-i,\, Q_b\right) .$$

Here $P(k,i)$ denotes the probability that player A wins by 6 games to k games and wins exactly i of his own service games. Also, the binomial notation, $b(n,r,p) = \binom{n}{r} p^r (1-p)^{n-r}$, is used and $Q_b = 1 - P_b$. Equations can also be written down for the cases in which $k = 1$ or 3 and also when player A loses the set $k/6$.

The probability that the games' score reaches 5 games each, $P_{5/5}$, is given by

$$P_{5/5} = \sum_{i=0}^{5} P_{5/5}(i)$$

where

$$P_{5/5}(i) = b(5, i, P_a)\, b(5, 5-i, Q_b)$$

and $P_{5/5}(i)$ denotes the probability that the games' score reaches 5 games each and player A wins exactly i of his own service games. Suppose P_{ad} denotes the probability that player A wins the set given that the games' score reached 5 games each. Then

$$P_{ad} = \sum_{j=0}^{\infty} P_j ,$$

where

$$P_j = P_a Q_b (P_a P_b + Q_a Q_b)^j .$$

Here P_j is the probability that an advantage set lasts for $(2j+2)$ games beyond 5/5 and that player A wins this set.

The person who serves in the first game of the second set of a tennis match is determined by the server in the first set and whether the number of games in the first set is odd or even. For this reason, much of the remainder of this section deals with sets that last an odd or an even number of games. The probability that player A wins the set in an odd (even) number of games, P_{od} (P_{ev}), can be written down.

The mean and variance of the duration of a set given player A won the set in an odd number of games, $\mu_{a,od}$ and $\sigma^2_{a,od}$ are given by

$$\begin{aligned}
\mu_{a,od} &= \sum_{k=1,3}\sum_{i=(7-k)/2}^{(7+k)/2} \frac{P(k,i)}{P_{od}}\Big\{i\mu_{a,1} + \left(\frac{7+k}{2}-i\right)\mu_{a,0} \\
&\qquad + (6-i)\mu_{b,0} + \left(\frac{k-7}{2}+i\right)\mu_{b,1}\Big\} \\
\sigma^2_{a,od} &= \sum_{k=1,3}\sum_{i=(7-k)/2}^{(7+k)/2} \frac{P(k,i)}{P_{od}}\Big\{i\sigma^2_{a,1} + \left(\frac{7+k}{2}-i\right)\sigma^2_{a,0} \\
&\qquad + (6-i)\sigma^2_{b,0} + \left(\frac{k-7}{2}+i\right)\sigma^2_{b,1} \\
&\qquad + \left(i\mu_{a,1} + \left(\frac{7+k}{2}-i\right)\mu_{a,0} + (6-i)\mu_{b,0}\right. \\
&\qquad \left. + \left(\frac{k-7}{2}+i\right)\mu_{b,1}\right)^2\Big\} - \mu^2_{a,od}
\end{aligned} \tag{8.3.1}$$

where $\mu_{i,j}$ and $\sigma^2_{i,j}$ ($i = a$ or b, $j = 0$ or 1) were obtained in Section 8.2.

Expressions for $\mu_{b,od}$ and $\sigma^2_{b,od}$, for the case in which player B won the set, can be correspondingly evaluated.

An advantage set which lasts an even number of games may result in a score such as, say, 10 games to 8 games. The duration of such sets can be analysed by summing the duration up to the score 5/5 and the duration beyond the score 5/5. The expected value (variance) of the duration of a set up to the score 5/5, given that the score 5/5 was reached, $\mu_{5/5}$ ($\sigma^2_{5/5}$), is given by

$$\begin{aligned}
\mu_{5/5} &= \sum_{i=0}^{5} \frac{P_{5/5}(i)}{P_{5/5}}\{i(\mu_{a,1}+\mu_{b,1}) + (5-i)(\mu_{a,0}+\mu_{b,0})\} \\
\sigma^2_{5/5} &= \sum_{i=0}^{5} \frac{P_{5/5}(i)}{P_{5/5}}\left\{\begin{array}{l} i(\sigma^2_{a,1}+\sigma^2_{b,1}) + (5-i)(\sigma^2_{a,0}+\sigma^2_{b,0}) \\ +\big(i(\mu_{a,1}+\mu_{b,1}) + (5-i)(\mu_{a,0}+\mu_{b,0})\big)^2 \end{array}\right\} - \mu^2_{5/5}\,.
\end{aligned} \tag{8.3.2}$$

The expected value and variance of the duration of two consecutive games which players A and B share, $\mu_{1/1}$ and $\sigma^2_{1/1}$, are given by

$$\begin{aligned}\mu_{1/1} &= \{P_aP_b(\mu_{a,1}+\mu_{b,1}) + Q_aQ_b(\mu_{a,0}+\mu_{b,0})\}/(P_aP_b+Q_aQ_b)\\ \sigma^2_{1/1} &= \left\{\frac{P_aP_b}{P_aP_b+Q_aQ_b}(\sigma^2_{a,1}+\sigma^2_{b,1}+(\mu_{a,1}+\mu_{b,1})^2)\right.\\ &\quad\left.+\frac{Q_aQ_b}{P_aP_b+Q_aQ_b}(\sigma^2_{a,0}+\sigma^2_{b,0}+(\mu_{a,0}+\mu_{b,0})^2)\right\} - \mu^2_{1/1}\,.\end{aligned} \tag{8.3.3}$$

The expected value and variance of the duration of a set beyond the score 5/5, given 5/5 is reached and player A wins the set, $\mu_{a,ad}$ and $\sigma^2_{a,ad}$, are given by

$$\begin{aligned}\mu_{a,ad} &= \frac{1}{P_{ad}}\sum_{j=0}^{\infty} jP_j\mu_{1/1} + (\mu_{a,1}+\mu_{b,0})\\ \sigma^2_{a,ad} &= \frac{1}{P_{ad}}\sum_{j=0}^{\infty} P_j\{j\sigma^2_{1/1} + (\mu_{a,1}+\mu_{b,0}+j\mu_{1/1})^2\}\\ &\quad + \sigma^2_{a,1} + \sigma^2_{b,0} - \mu^2_{a,ad}\,.\end{aligned} \tag{8.3.4}$$

Hence, the expected value and variance of the duration of a set given player A wins the set in an even number of games, $\mu_{a,ev}$ and $\sigma^2_{a,ev}$, are given by

$$\begin{aligned}\mu_{a,ev} &= \left\{\sum_{k=0,2,4}\sum_{i=(6-k)/2}^{(6+k)/2}\frac{P(k,i)}{P_{ev}}\right.\\ &\quad\times\left\{i\mu_{a,1} + \left(\frac{6+k}{2}-i\right)\mu_{a,0} + (6-i)\mu_{b,0} + \left(\frac{k-6}{2}+i\right)\mu_{b,1}\right\}\\ &\quad\left.+\frac{P_{5/5}P_{ad}}{P_{ev}}(\mu_{5/5}+\mu_{a,ad})\right\}\\ \sigma^2_{a,ev} &= \left\{\sum_{k=0,2,4}\sum_{i=(6-k)/2}^{(6+k)/2}\frac{P(k,i)}{P_{ev}}\right.\\ &\quad\times\left\{i\sigma^2_{a,1} + \left(\frac{6+k}{2}-i\right)\sigma^2_{a,0} + (6-i)\sigma^2_{b,0} + \left(\frac{k-6}{2}+i\right)\sigma^2_{b,1}\right.\\ &\quad\left.+\left(i\mu_{a,1} + \left(\frac{6+k}{2}-i\right)\mu_{a,0} + (6-i)\mu_{b,0} + \left(\frac{k-6}{2}+i\right)\mu_{b,1}\right)^2\right\}\\ &\quad\left.+\frac{P_{5/5}P_{ad}}{P_{ev}}\{\sigma^2_{5/5}+\sigma^2_{a,ad}+(\mu_{5/5}+\mu_{a,ad})^2\} - \mu^2_{a,ev}\right\}.\end{aligned} \tag{8.3.5}$$

Corresponding expressions for $\mu_{b,ev}$ and $\sigma^2_{b,ev}$ can be obtained.

Table 8.3.1 gives the values of these expressions for the case in which $p_a = 0.7$ and $p_b = 0.6$. This table gives values for PA, the probability that player A wins the set, μ_a and σ_a^2, the mean and variance of the duration of a set won by player A, and μ and σ^2, the unconditional mean and variance.

All the values in Table 8.3.1 have been checked using a random walk procedure and, where comparisons are possible, agree with the few results available in earlier publications. The analysis and simulations of Miles (1984) produced corresponding results.

Table 8.3.1
Some characteristics of a classical set
assuming $p_a = 0.7$ and $p_b = 0.6$

	Server in the First Game	
	Player A	Player B
$P_{6/0}$	0.0135	0.0135
$P_{6/1}$	0.1055	0.0310
$P_{6/2}$	0.0867	0.1613
$P_{6/3}$	0.3113	0.0913
$P_{6/4}$	0.0867	0.3067
$P_{5/5}$	0.2751	0.2751
$Q_{6/4}$	0.0940	0.0266
$Q_{6/3}$	0.0105	0.0780
$Q_{6/2}$	0.0152	0.0082
$Q_{6/1}$	0.0011	0.0081
$Q_{6/0}$	0.0004	0.0004
P_{ad}	0.7653	0.7653
PA	0.8143	0.8143
$\mu_a(\sigma_a^2)$	65.11 (803.62)	67.56 (750.40)
$\mu_b(\sigma_b^2)$	74.10 (838.62)	71.80 (907.13)
$\mu(\sigma^2)$	66.78 (822.33)	68.35 (782.23)

8.4 The Distribution of the Duration of a set of Classical Tennis

Suppose $A_1(s)$ ($A_0(s)$) denotes the probability generating function (p.g.f.) of the duration of a game which player A served and won (lost). Similarly, suppose

$B_1(s)$ ($B_0(s)$) is the p.g.f. of the duration of a game which player B served and won (lost). These p.g.f. 's can be obtained from Section 8.2. Noting equation (8.3.1), the p.g.f. for the duration of a set which is won by player A in an odd number of games, $P_{od}(s)$, is given by

$$P_{od}(s) = \sum_{k=1,3} \sum_{i=(7-k)/2}^{(7+k)/2} P(k,i)A_1^i(s)A_0(s)^{((7+k)/2)-i} \times B_1(s)^{((k-7)/2)+i}B_0(s)^{6-i}/P_{od}\,. \tag{8.4.1}$$

A corresponding expression can be written down for $Q_{od}(s)$ for the case in which player A lost the set.

Suppose $C(s)$ is the p.g.f. for the duration of a set up to the score 5/5, given that 5/5 was reached. Suppose $D(s)$ is the p.g.f. for the duration of two consecutive games which players A and B share. Noting equations (8.3.2) and (8.3.3), it follows that

$$C(s) = \left[\sum_{i=0}^{5} P_{5/5}(i)A_1(s)^i A_0(s)^{5-i}B_1(s)^i B_0(s)^{5-i}\right] \Big/ P_{5/5}$$

and

$$D(s) = \left[P_a P_b A_1(s)B_1(s) + Q_a Q_b A_0(s)B_0(s)\right]/(P_a P_b + Q_a Q_b)\,.$$

Also suppose $E(s)$ is the p.g.f. for the duration of a set beyond the score 5/5, given 5/5 was reached and player A won the set. Then, noting equation (8.3.4),

$$E(s) = \left(\sum_{j=0}^{\infty} P_j D(s)^j A_1(s)B_0(s)\right) \Big/ P_{ad}\,.$$

Now, noting equation (8.3.5), the p.g.f. for the duration of a set given player A won the set in an even number of games, $P_{ev}(s)$, is given by

$$P_{ev}(s) = \frac{1}{P_{ev}} \left\{ \begin{array}{l} \displaystyle\sum_{k=0,2,4} \sum_{i=(6-k)/2}^{(6+k)/2} P(k,i)A_1(s)^i A_0(s)^{((6+k)/2)-i}B_1(s)^{((k-6)/2)+i} \\ \times B_0(s)^{6-i} + P_{5/5}P_{ad}C(s)E(s) \end{array} \right\}.$$

A corresponding expression can be written down for $Q_{ev}(s)$ for the case in which player A lost the set.

The p.g.f. for the duration of a set conditional on player A winning (losing) the set, $P(s)$ ($Q(s)$), can then be derived, as can the unconditional p.g.f. . It has been shown computationally that the distribution given by equation (8.4.1) produces a mean and variance in agreement with the corresponding results of Section 8.3.

8.5 The Mean and Variance of the Duration of a "Best-of-3 Sets" Match of Classical Tennis

Noting that, if and only if a classical set lasts an even number of games, the server at the beginning of the next set is the same player as the server at the beginning of the set just completed, it is possible, using an elementary "branching" diagram, to obtain "match" results.

8.6 An Analysis of the Tie-Breaker Game

In a tie-breaker set, the tie-breaker game is played if the games' score reaches 6 games each. The "12 point tie-breaker game" is played in the following manner. Player A serves the first point (since it has been assumed that he served in the first game of the set); Player B serves the next 2 points; player A serves the next 2 points; player B serves the next 2 points; etc. The first player to score at least 7 points and be at least 2 points ahead of his opponent is the winner of the tie-breaker game and set. Suppose, for $i = 0$ to 5, PT_i denotes the probability that player A wins the tie-breaker game by 7 points to i points. Then

$$
\begin{aligned}
PT_0 &= {p_a}^3{q_b}^4 \\
PT_1 &= 3{p_a}^3 q_a {q_b}^4 + 4{p_a}^4 p_b {q_b}^3 \\
PT_2 &= 6{p_a}^3{q_a}^2{q_b}^4 + 16{p_a}^4 q_a p_b {q_b}^3 + 6{p_a}^5{p_b}^2{q_b}^2 \\
PT_3 &= 10{p_a}^2{q_a}^3{q_b}^5 + 40{p_a}^3{q_a}^2{q_b}^4 p_b + 30{p_a}^4 q_a {p_b}^2{q_b}^3 + 4{p_a}^5{p_b}^3{q_b}^2 \\
PT_4 &= 5p_a{q_a}^4{q_b}^6 + 50{p_a}^2{q_a}^3 p_b {q_b}^5 + 100{p_a}^3{q_a}^2{p_b}^2{q_b}^4 \\
&\quad + 50p_a^4 q_a {p_b}^3{q_b}^3 + 5{p_a}^5{p_b}^4{q_b}^2 \\
PT_5 &= p_a{q_a}^5{q_b}^6 + 30{p_a}^2{q_a}^4 p_b {q_b}^5 + 150{p_a}^3{q_a}^3{p_b}^2{q_b}^4 + 200{p_a}^4{q_a}^2{p_b}^3{q_b}^3 \\
&\quad + 75{p_a}^5 q_a {p_b}^4{q_b}^2 + 6{p_a}^6{p_b}^5 q_b
\end{aligned}
$$

where p_a and p_b are defined in Section 8.1.

The probability that the points' score reaches 6 points each in the tie-breaker game, PTT, satisfies the equation,

$$PTT = \sum_{j=0}^{6} b(6, j, p_a)\, b(6, j, p_b)\,.$$

The probability, PT_i, that player A wins a tie-breaker game by $(i+2)$ points to i $(i = 6, 7, 8, \ldots)$ is given by

$$PT_i = PPT(p_a p_b + q_a q_b)^{i-6} p_a q_b \ .$$

Losing probabilities QT_i can be obtained correspondingly. Hence conditional and unconditional distributions of duration can be obtained. The above expressions were checked using random walk methods.

8.7 The Distribution of the Duration of a Tie-Breaker Set

The "tie-breaker game", being regarded as a *single* game, means that a tie-breaker set reaching 6/6 must produce an odd number of games (contrary to classical tennis). Making use of Section 8.6, Section 8.4 can thus be modified to give the distribution of the duration of a tie-breaker set. Table 8.7.1 gives some characteristics of a tie-breaker set and could be compared with Table 8.3.1.

Table 8.7.1
Some characteristics of a tie-breaker set
assuming $p_a = 0.7$ and $p_b = 0.6$

	Server in First Game	
	Player A	Player B
PA	0.7949	0.7949
$\mu_a(\sigma_a^2)$	59.54 (237.98)	62.04 (210.79)
$\mu_b(\sigma_b^2)$	69.18 (214.46)	67.13 (257.55)
$\mu(\sigma^2)$	61.52 (248.29)	63.08 (224.61)

8.8 A Comparison of a Classical and Tie-Breaker Match of Tennis

(i) Table 8.8.1a shows that the probability that the better player wins a match of the best-of-three sets is smaller for the tie-breaker system, the reduction in probability typically being small.

(ii) It can be seen from Table 8.8.1b that the expected duration of a match is reduced for the tie-breaker version (at the expense of a (typically) slightly

Table 8.8.1a

The probability player A wins a best-of-three sets match

p_b	p_a	Probability Classical	Tie-breaker	Difference
0.5	0.5	0.5000	0.5000	0.0000
0.5	0.6	0.9147	0.9078	0.0069
0.5	0.7	0.9959	0.9947	0.0012
0.6	0.6	0.5000	0.5000	0.0000
0.6	0.7	0.9093	0.8910	0.0183
0.6	0.8	0.9958	0.9899	0.0059
0.7	0.7	0.5000	0.5000	0.0000
0.7	0.8	0.9329	0.8632	0.0697
0.8	0.8	0.5000	0.5000	0.0000

Table 8.8.1b

The expected duration of a best-of-three sets match

p_b	p_a	Expected Duration Classical	Tie-breaker	Difference
0.5	0.5	169.3	164.6	4.7
0.5	0.6	144.7	142.1	2.6
0.5	0.7	113.1	112.3	0.8
0.6	0.6	173.3	164.0	9.3
0.6	0.7	155.7	145.1	10.6
0.6	0.8	129.4	120.9	8.5
0.7	0.7	216.0	165.5	50.5
0.7	0.8	238.4	154.0	84.4
0.8	0.8	447.0	172.0	275.0

smaller probability of the better player winning). This reduction in expected duration is large when both players have strong services (e.g. $p_a = p_b = 0.8$). It can also be seen that the expected duration is relatively constant for the tie-breaker version (as the parameters p_a and p_b change).

(iii) Table 8.8.1c shows that the variance of duration of a match is always smaller for the tie-breaker version. This smaller variance is relatively constant as the parameters change whereas the variance of the classical version increases rapidly as both p_a and p_b increase.

(iv) The conclusions in (ii) and (iii) explain the long matches often previously observed in classical tennis between two players each with very effective services on fast grass-court surfaces.

Table 8.8.1c

The variance of duration of a best-of-three sets match

p_b	p_a	Variance Classical	Variance Tie-breaker
0.5	0.5	2264.4	1767.3
0.5	0.6	1915.5	1592.3
0.5	0.7	827.5	730.2
0.6	0.6	2736.0	1701.0
0.6	0.7	2947.4	1551.3
0.6	0.8	1956.5	884.4
0.7	0.7	8866.5	1630.9
0.7	0.8	16490.3	1599.4
0.8	0.8	55021.3	1508.2

(v) Carter and Crews (1974, p.134) claimed that "the maximum expected set length is 67.70 points" for classical tennis. However, it can be seen that the expected set length can be made arbitrarily large by selecting p_a and p_b close enough to 1.

(vi) In the 32 "best-of-3 tie-breaker sets" matches analysed by Pollard (1980), player A (the "winner") won 71% of points on his service and his opponent averaged 62% on his. These 32 matches averaged 149.7 points and had a variance of 1734. Noting that p_a and p_b vary from match to match, it is reasonable to believe that this data is consistent with Tables 8.8.1b and 8.8.1c.

It can be seen that the (conditioning) methodology used in Sections 8.2 to 8.4 could be used to obtain corresponding expressions for other "nested- type" scoring systems. For example, tennis is 3-nested (game - set - match) whereas squash is only 2-nested (game - match).

8.9 A Tennis Scoring System With Smaller Variance of Duration

We have seen in the previous section that the variance of duration of a match is reduced by using tie-breaker sets rather than advantage sets. Probably the main reason for this change from advantage to tie-breaker sets was to prevent the possibility of very long sets which would delay the scheduled starting times of later matches. Another disadvantage of such long sets was that it was sometimes the physically stronger player rather than the better player who finally won and, in any case, the eventual winner was often too tired to play well on the following day. However, the tie-breaker change has not satisfactorily solved the delay problem. As an example of such delays, during one day of the 1984 Australian Indoor Championships, the play scheduled during the day session unfortunately carried over until one hour after the scheduled start of the evening session. Many people who had bought tickets for the evening session were inconvenienced as they had to wait until the spectators from the day session vacated their seats. The cause of the problem was the occurrence of several 3-set matches, some involving tie-breakers. No doubt similarly large delays (and many smaller ones) have occurred in other tournaments. Is it possible to take the present best-of-three tie-breaker sets scoring system, modify it by as little as possible (administrators and players may be prepared to accept slight modifications!) to produce a new system with a similar expected duration and probability that each player wins but *having a smaller variance of duration?* Such a new scoring system would typically produce a match of a more predictable duration (leading to smaller delays) which would be advantageous to players, administrators and spectators. The answer to the above question is "yes" and the proposed scoring system is a best-of-five half-sets which operate in a way outlined in the next paragraph.

Under the proposed scoring system games are played as at present. However, instead of these games being the components that make up sets, it is recommended that they make up *half-sets* which operate in the following way. A player wins a half-set when the game score reaches 4–0, 4–1 or 4–2 in that player's favour; if the game score reaches 3–3, the half-set has resulted in a draw. The next half-set is then commenced. A half-unit is awarded to a player as soon as that player's game score reaches 3 games and another half unit if he reaches 4 games. The match is over as soon as one player's score reaches a total of three units and that player is declared the winner. However, if after five half-sets have been played the unit score is 2.5 to 2.5, the tie-breaker game (as currently in operation) is played to determine the winner. The first server in the first half-set is determined by a toss, whereas in every other half-set the first server is the person who received in the last game of the previous half-set. Thus, each half-set carries one unit which is split in the event of a draw (3–3) and a player requires a total of 3 units in order to win the match.

Note that occasionally a half-set is not completed. For example, when the overall leader obtains a lead in the 4th or 5th half-set of 3–0, 3–1 or 3–2, the match will stop and the overall leader declared the winner if the overall leader's unit score has just reached 3. This added complication means that it is impractical to use the exact methods of the earlier sections of this chapter, as this would be far too time-consuming for the end results. For this reason the proposed scoring system was analysed by simulation using a Fortran program. It was assumed that player A (B) had a constant probability p_a (p_b) of winning any point when serving and that the outcome of any point was independent of the outcome of any other point. For each of several pairs of the parameters p_a and p_b, 10,000 simulated matches were carried out and the results are given in Tables 8.9.1, 8.9.2 and 8.9.3.

It was established by simulating matches of best-of-three tie-breaker sets and comparing the results with the known exact values available from Section 8.8 that sample sizes of 10,000 matches were sufficient to produce quite accurate results. A comparison of Tables 8.8.1c and 8.9.1 shows that the variance of the duration of a match is reduced under the proposed method of scoring. This reduction is greatest when the values of both p_a and p_b are large (say both 0.8). This corresponds to the case in which both players have 'strong' services or the match is a doubles match between two strong service-volley combinations. These have indeed most often been the cases which have given rise to the delays experienced at tournaments and hence the proposed method of scoring makes the biggest improvements for those cases causing the worst problems. Tables 8.8.1a, 8.8.1b and 8.9.1 also show that the proportions of matches won by player A and the mean durations are quite similar under the present (best-of-three tie-breaker sets) and the proposed methods of scoring. A comparison of these tables also indicates that, under the

Table 8.9.1

The proportion of matches won by player A and the mean and variance of the duration of a best-of-five half-sets match.

(Simulation results)

p_b	p_a	Proportion	Mean	Variance
0.5	0.5	0.5059	168.0	1055.2
0.5	0.6	0.9043	148.0	1185.6
0.5	0.7	0.9940	119.1	789.7
0.6	0.6	0.5001	166.6	873.9
0.6	0.7	0.8869	150.3	895.6
0.6	0.8	0.9896	129.8	688.9
0.7	0.7	0.5069	163.2	498.0
0.7	0.8	0.8449	153.8	446.7
0.8	0.8	0.5018	157.5	175.5

proposed system of scoring, the average duration varies to a lesser extent as the parameters p_a and p_b change. This is a favourable property of the proposed system as, in practice, parameters do vary from match to match.

Table 8.9.2 gives the observed distribution of duration for the simulation of the present and proposed methods of scoring when $p_a = 0.7$ and $p_b = 0.6$ (an appropriate pair of values for p_a and p_b in a typical men's singles match (Pollard, 1980, 1983a)). This table shows that under the proposed scoring system not only is the variance reduced, but the distribution of duration (when $p_a = 0.7$, $p_b = 0.6$) is reasonably symmetric whereas, for the present system it is clearly positively skewed (giving rise to some very long matches).

Table 8.9.3 presents data relevant to strong services in singles matches or strong service-volley combinations in doubles matches ($p_a = p_b = 0.75$); it clearly indicates that delay problems must be expected in tournaments which use the present tie-breaker system. The very effective truncation (by using the proposed scoring system) of such potentially long matches is clearly identified in this table. As it is matches with such parameters that have often caused the underlying delay problems, the resolution of these problems by using the proposed method of scoring

Table 8.9.2
The observed distribution of duration
when $p_a = 0.7$ and $p_b = 0.6$.
(Simulated results)

Number of points played	Frequency	
	Best-of-three tie-breaker sets	Best-of-five half-sets
60–69	3	1
70–79	32	32
80–89	209	161
90–99	644	365
100–109	1066	430
110–119	1247	606
120–129	1169	977
130–139	1051	1206
140–149	789	1121
150–159	610	994
160–169	554	1113
170–179	502	1152
180–189	522	929
190–199	438	546
200–209	381	244
210–219	300	95
220–229	213	24
230–239	130	3
240–249	83	1
≥ 250	57	0
Total	10,000	10,000
Prop (A)	0.8950	0.8869
Mean	144.7	150.3
Variance	1537	895.6

Table 8.9.3
The observed distribution of duration
when $p_a = p_b = 0.75$. (Simulated results)

Number of points played	Frequency: Best-of-three tie-breaker sets	Frequency: Best-of-five half-sets
60–69	0	0
70–79	1	1
80–89	20	7
90–99	157	14
100–109	341	28
110–119	526	219
120–129	861	481
130–139	896	446
140–149	872	930
150–159	919	2090
160–169	664	2651
170–179	592	1916
180–189	602	921
190–199	790	239
200–209	814	45
210–219	709	10
220–229	550	0
230–239	382	2
240–249	213	0
≥ 250	91	0
Total	10,000	10,000
Prop (A)	0.4977	0.4975
Mean	169.3	160.6
Variance	1585	315.7

is most evident. The negatively skewed distribution of duration for such matches is a desirable property of the proposed scoring system (or for any scoring system!) since potentially long matches are effectively truncated and one can be almost certain that matches will not take more than a certain number of points. Also, from the individual player's point of view, it can be seen from Tables 8.9.2 and 8.9.3 that a player will, under the proposed system, only rarely be required to play a long, tiring match; also a match is more likely to be a test of skill rather than endurance.

What is the reason for the improvement in variance under the proposed system? The essential reason for the big variance in the best-of-three sets case is that (about) 68% of matches take only 2 sets and, in the remaining (about) 32% of matches, a third set (which typically takes another (say) 60 points) is required, usually when the first two sets have already taken longer than average (often involving at least one tie-breaker). The above approximate 68/32% breakdown can be inferred from Table 8.7.1. Also in the data examined by Pollard (1980), 16 out of the 32 best-of-three set matches required a third set. By comparison, when using best-of-five half-sets, 4 or 5 half-sets are required about 90% of the time, the difference in duration here being one half-set or only (say) 35 points. In only about 10% of matches will one player win by 3 half-sets to 0 half-sets because of the high probability of one of the first three half-sets being a draw.

An analysis of the best-of-five half-sets proposal together with an extension of it in which "sudden-death" games are used (here, if deuce is reached, only one more point is played to determine the winner of the game) is contained in a paper by Pollard (1987). This extension produced a further reduction in the variance.

I am indebted to Dr Ken Noble for carrying out the simulations referred to in this section as well as those in Example 9.4.1 and Section 9.5.2.

CHAPTER 9
SOME CONSIDERATIONS IN THE DESIGN OF SCORING SYSTEMS

9.1 Introduction

This chapter very briefly *outlines* some of the problems in designing a scoring system and presents a few relevant results. An extended discussion of this problem is currently being carried out by my supervisor, Dr Roger Miles, and will appear in a monograph (Miles (to appear)).

The designer of a scoring system should firstly determine, for the particular sport being considered, the appropriate number of points that should, on average, be played in a match (μ). A match which is too short will give the poorer player too great a probability of winning whilst a match which is too long has other disadvantages such as (a) the ultimate winner may be determined by stamina rather than skill, (b) it may become too boring for the spectators (c) its outcome may become too predictable, ...

Having determined the appropriate size of μ, the designer would ideally prefer a system with a small variance of duration since a planned program of matches can occur more predictably for all persons concerned (players, officials and spectators). Unfortunately all of the W_n systems discussed in Chapters 1 to 5 have very large variances. Thus the designer needs to make a compromise between ρ, the efficiency of a system and σ^2, its variance of duration since a system with decreased σ^2 typically also has decreased ρ. Miles (1984, p.106) indicated, for the bipoints case, how a truncation of W_n(point-pairs) could be used to reduce σ^2 quite dramatically whilst the drop in ρ was not too marked. His truncation procedure involved a hybrid system based on the W_n and B_{2m-1} systems. An alternative variance reduction procedure is given in Section 9.3 for the unipoints case. Miles (1984, p.107) had conjectured that, for a "given" μ, B_{2n-1} ($n = 1, 2, \ldots$) is the class of uniformats minimizing σ^2. However, it is shown in Section 9.3 that it is possible to create uniformats with considerably smaller σ^2 than the B_{2n-1} family. This is achieved, however, at the expense of efficiency.

Nested scoring systems, although typically having reduced efficiency, provide a useful structure in which either player can more easily overcome a period of poor play. For example, in best-of-three sets tennis, the better player could lose the first (say) 24 points during a period of poor play and yet quite possibly recover to win the match. Such a recovery would be nigh impossible if one of the very efficient scoring systems was used. Section 9.4 investigates variance and efficiency for nested

B_{2n-1} systems and gives an appropriate recommendation for such systems. In Section 9.5 a further recommendation is made for B_{2n-1} systems in which each of the possible $2n - 1$ units could result in a draw. Since many scoring systems used in practice are based on or resemble to some extent the B_{2n-1} system (e.g. badminton and international squash), it would be reasonable to believe that the recommendations for B_{2n-1} systems made in Sections 9.4 and 9.5 also apply to such related systems. Section 9.5 also gives a second alternative method of scoring in tennis.

In Section 9.2, the $P - D - I$ diagram is proposed as a useful tool (for the person designing a scoring system) for displaying the relevant characteristics of a system under consideration (such as the differing importances of the points (and hence the efficiency of the system), the entropy (defined in Section 9.2) and the variance of duration).

One further comment concerns the distribution of duration rather than just its mean and variance. In Section 8.9, it was recommended that, other things being equal, a scoring system with a negatively skewed distribution of duration is preferable as potentially long matches are thus more effectively truncated.

9.2 The $P - D - I$ Diagram as a Useful Tool for the Designer of Scoring Systems

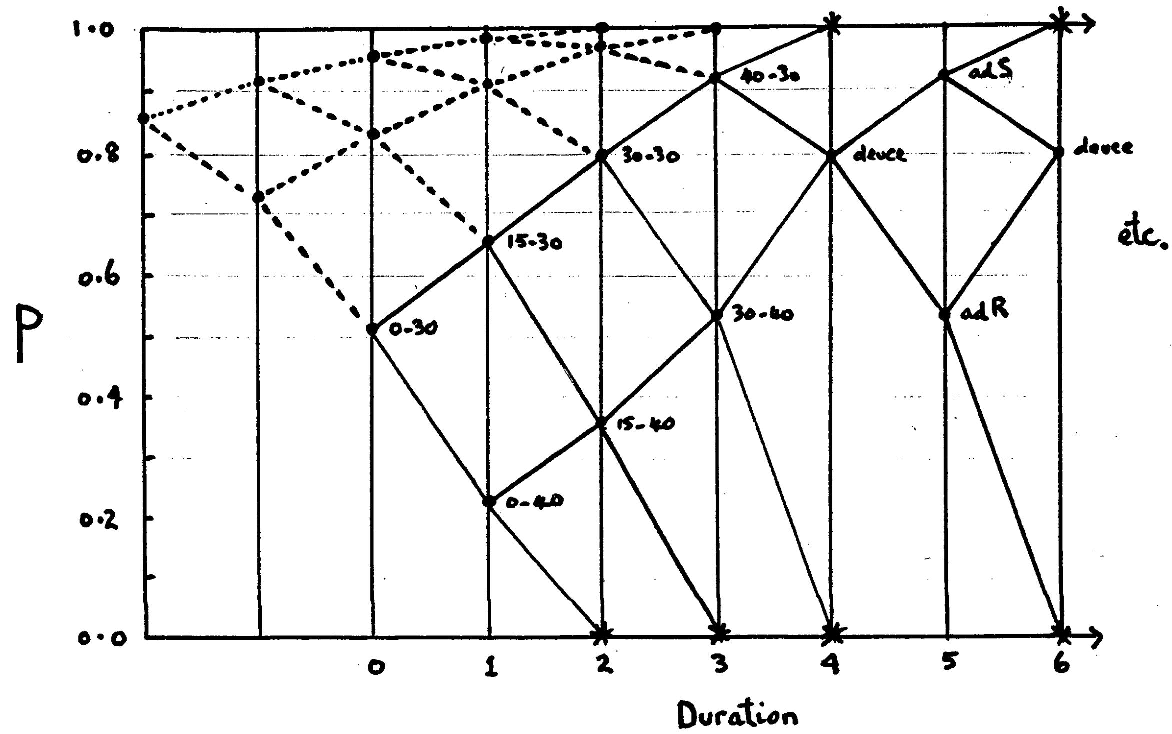
The $P - D - I$ diagram is now presented as a useful tool for highlighting the characteristics that the designer of a scoring system needs to consider. An example is taken from Miles (1984, p.104–107) who recommended "0–30 tennis" for cases in which serving was a distinct advantage ($v = \frac{1}{2}(p_a + p_b - 1)$; $.1 \leq v < .24$). He noted the increased entropy, H, defined by

$$H = -(P \ln P + Q \ln Q)$$

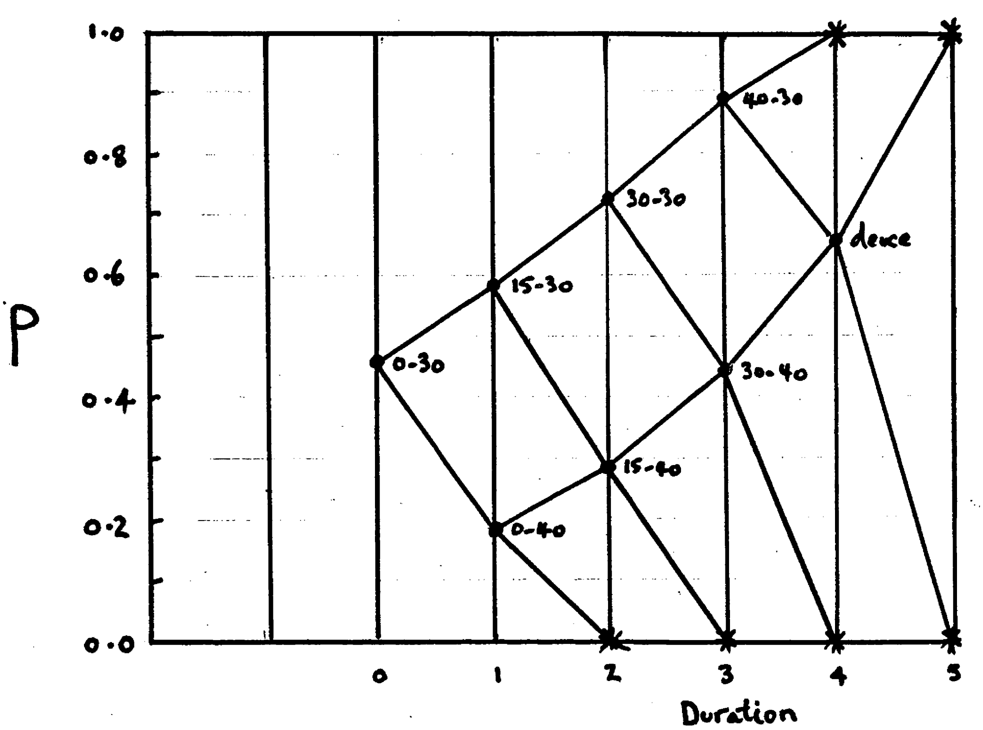
(where P is player A's probability of winning and $Q = 1 - P$) and the increased efficiency of 0–30 tennis in such situations. These two facts can be seen in the $P-D-I$ diagram for 0–30 tennis when p_a or p_b is equal to (say) $\frac{2}{3}$ (see Figure 9.2.1(a)). In this diagram it can be seen that the relatively unimportant points (0–0), (0–15), (15–0), (15–15), (30–15), (30–0), (40–0) and (40–15) are deleted from the game leaving points which are "relatively equal" in importance, thus increasing efficiency (recall equation (4.5.2)). Also, when p_a or p_b equals (say) $\frac{2}{3}$, the entropy at (0–30) when $P = 0.51$ is much greater than at (0–0) when $P = 0.86$. Thus, the entropy and the efficiency of a proposed scoring system can be appreciated

Figure 9.2.1
The $P-D-I$ diagram for "0–30 tennis" when p_a or p_b equals $\frac{2}{3}$.

(a) advantage games



(b) sudden-death games



(at least in an approximate sense) by observing the $P-D-I$ diagram of it for an appropriate parameter set. The variance of duration of a proposed system can also be appreciated by noting the spread of the starred (*) or absorption points in the $P-D-I$ diagram of it. For example, a comparison of Figures 9.2.1(a) and 9.2.1(b) shows that, when sudden death games (in which only one further point is played if deuce is reached) are used (instead of advantage games) together with 0–30 tennis, the variance of duration is clearly decreased with little effect on entropy or efficiency when p_a or p_b equals (say) $\frac{2}{3}$. This example indicates that the $P-D-I$ diagram is a useful tool for displaying importances (and hence efficiency), entropy and variance of duration.

9.3 Variance Reduction

Considering unipoints for example, suppose a scoring system with a small variance of duration is required. This can be achieved if the earliest and latest possible finishing durations, d_1 and d_2 respectively, are "close together". The structure of the $P-D-I$ diagram indicates that one way this can be achieved is by making the $(d_1+1)^{\text{st}}$ point and later ones increasingly important. Clearly the d_2^{th} point has an importance of unity. This is effectively how the hybrid system (Miles (1984, p.106)) discussed in Section 9.1 operates. A uniformat with a very small variance of duration is now devised. This uniformat, which operates in the following way, must have a duration of n or $n+1$ points.

(a) the first two points are worth 1 "unit" each;

(b) the m^{th} point $(3 \leq m \leq n)$ is worth

  (i) 2^{m-2} units if one player is ahead of the other or

  (ii) 2^{m-1} units if the players are equal;

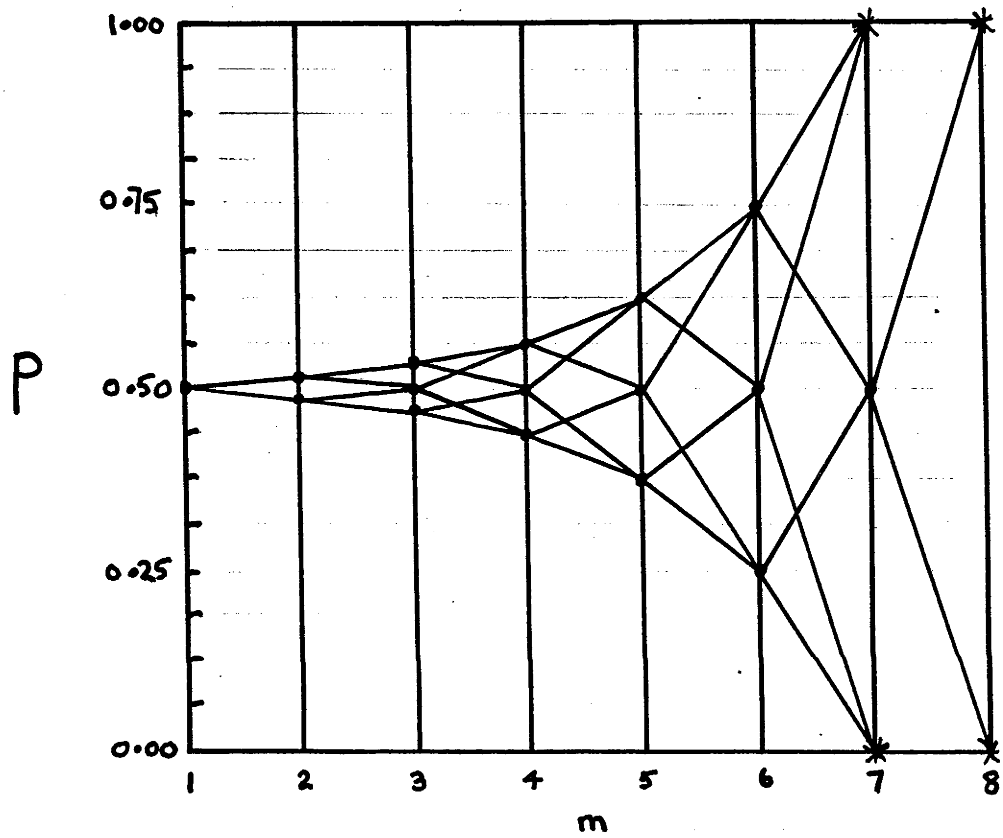
(c) the $(n+1)^{\text{st}}$ is worth 2^{n-1} units.

The winner is the first player to get 2^{n-1} units ahead of his opponent ($W_{2^{n-1}}$ units). Figure 9.3.1 gives an example of the $P-D-I$ diagram for the case in which $n=6$ and $p=\frac{1}{2}$. This particular system is clearly of no real practical value as although its variance of duration is very small (even for a large μ (n large)), its efficiency is also clearly very low and decreases to zero as $n \to \infty$. This example does indicate, however, that by making points increasingly more important beyond a certain point, the variance of duration can be very effectively reduced. The approach in this section recognizes or accepts *a priori* that points need not be equally important for two equal players whereas, in all other uniformats (except W_n (see

Section 1.2)), this fact is not immediately evident and is effectively disguised within the structure of the uniformat.

Figure 9.3.1

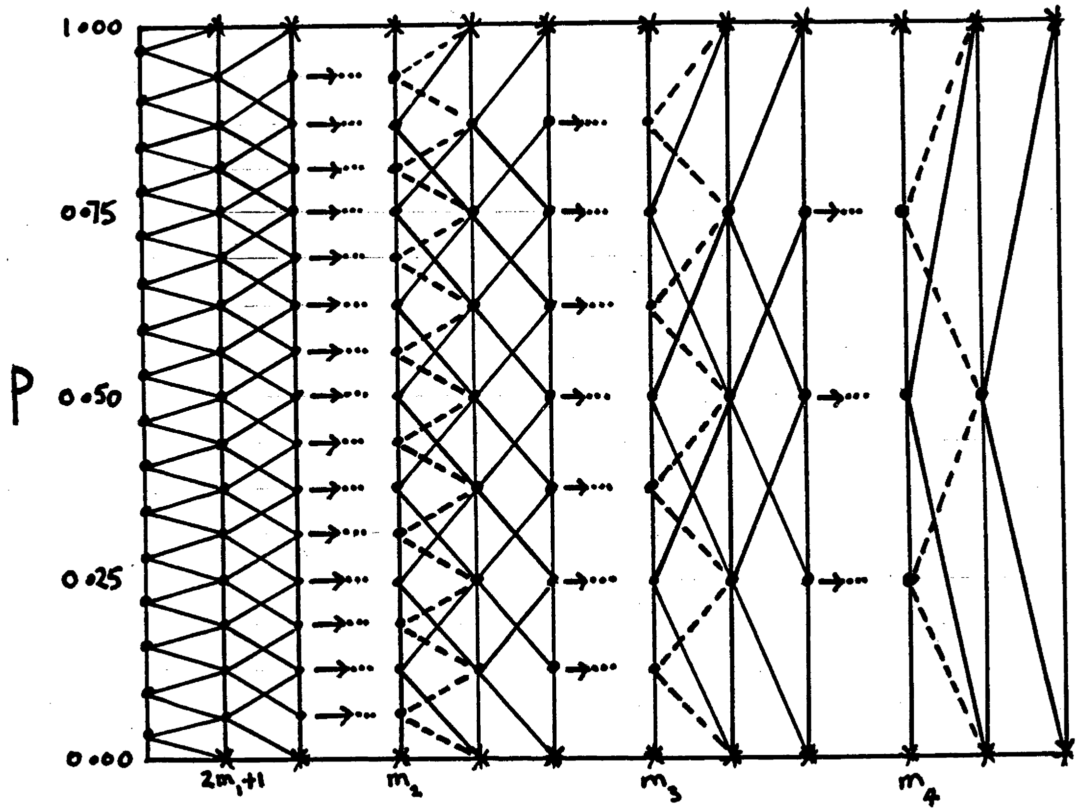
The $P - D - I$ diagram of a uniformat with a very small variance of duration ($p = 0.5$).



An alternative approach to variance reduction is to take the optimally efficient uniformat W_{2^n} and modify it so that at various stages (the $(2m_1+1)^{\text{st}}$ point, the m_2^{th} or $(m_2+1)^{\text{st}}$ point, the m_3^{th} or $(m_3+1)^{\text{st}}$ point, ...) points are made (say) twice as important as their current value, until at the final stage the point has an importance of unity ($2m_1+1 < m_2 < m_3 < \ldots$). Figure 9.3.2 gives an example for the case in which $n = 4$. The values of $m_1, m_2, m_3, \ldots$ can be selected in order to satisfy the designer's requirements. It can be seen that when all the m_i are small and "close together" the reduction in variance is greatest (as is the decrease in efficiency) whereas, when all the m_i are very large, there is little reduction in variance or efficiency on that of the W_{2^n} uniformat.

Figure 9.3.2

Sections of the $P - D - I$ diagram for a method of reducing the variance of duration of the W_{16} uniformat ($p = 0.5$).



9.4 Nested B_{2n-1} Systems

We have seen that nesting is a useful device which allows either player to overcome a period of poor play (either a "slow" start or a "slump" during the match). In this section we will firstly consider the effect of nesting on B_{2n-1} uniformats (see page 1.3). As the cases of greatest interest are when p lies in the range 0.4 to 0.6, we will focus our attention on the case in which $p = 0.5$.

Firstly we note that, for the uniformat B_{2n-1} at $p = 0.5$,

$$\mu_W = \mu_L = \mu$$

$$\text{and} \qquad \sigma_W^2 = \sigma_L^2 = \sigma^2$$

where μ_W (μ_L) is the expected duration conditional on player A winning (losing), μ is the unconditional expected duration and the σ^2's are the corresponding variances of duration. Now consider the uniformat B_{2n_1-1} (B_{2n_2-1}) at $p = 0.5$. The methods of Section 8.3 can be used to show that, for B_{2n_1-1} (B_{2n_2-1}) at $p = 0.5$, the mean and variance of duration, μ and σ^2, are given by

Table 9.4.1

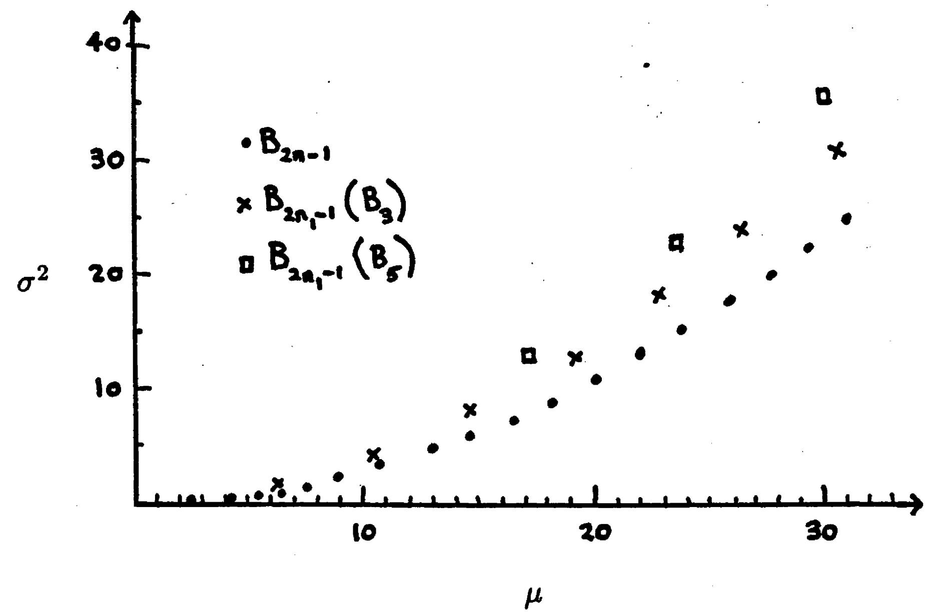
The expected value and variance of the duration and the limit value of the efficiency of "best-of-..." uniformats at $p = 0.5$.

Uniformat	Expected Duration	Variance of Duration	$\lim_{\delta\to 0} \rho(\frac{1}{2}+\delta)$
B_3	2.50	0.25	0.900
B_5	4.13	0.61	0.852
B_7	5.81	1.11	0.823
B_9	7.54	1.77	0.803
$B_3(B_3)$	6.25	2.19	0.810
$B_5(B_3)$	10.31	4.84	0.767
$B_7(B_3)$	14.53	8.39	0.741
$B_9(B_3)$	18.85	12.94	0.723
$B_5(B_5)$	17.02	12.88	0.726
$B_7(B_5)$	23.98	22.44	0.702
$B_9(B_5)$	31.10	34.70	0.685
$B_7(B_7)$	33.79	43.98	0.678
$B_3(B_5)$	10.31	5.78	0.767
$B_3(B_7)$	14.53	11.22	0.741
$B_3(B_9)$	18.85	18.63	0.723
$B_5(B_7)$	23.98	25.17	0.702
$B_5(B_9)$	31.10	41.93	0.685
$B_3\big(B_3(B_3)\big)$	15.63	15.23	0.729
$B_5\big(B_3(B_3)\big)$	25.78	32.83	0.690
$B_3\big(B_5(B_3)\big)$	25.78	38.69	0.690
$B_3\big(B_3(B_5)\big)$	25.78	41.03	0.690

$$\begin{aligned} &\mu = \mu_1\mu_2 \\ \text{and} \qquad &\sigma^2 = \mu_1\sigma_2^2 + \mu_2^2\sigma_1^2 \end{aligned} \tag{9.4.1}$$

where B_{2n_i-1} has a duration with mean μ_i and variance σ_i^2 at $p = 0.5$ ($i = 1, 2$). Table 9.4.1 gives values for μ and σ^2 for some such nested systems. It can be seen from Figure 9.4.1 that, for B_{2n_1-1} (B_{2n_2-1}) at $p = 0.5$, the variance for a "given" μ increases as n_2 increases and n_1 decreases (provided $n_1 \geq 2$). Thus, for example, B_7 (B_3) has a smaller variance at $p = 0.5$ than does B_3 (B_7). Corresponding

Figure 9.4.1
The variance of duration, σ^2, for a "given" expected duration, μ, for "best-of-..." uniformats when $p = 0.5$.



asymptotic results also apply. Uppuluri and Blot (1974) have shown that for B_{2n-1} (n large),

$$\mu = 2n - 2\sqrt{\frac{n}{\pi}}$$

$$\text{and} \qquad \sigma^2 \doteq 2n - \frac{4n}{\pi} - 2\sqrt{\frac{n}{\pi}}$$

at $p = 0.5$. Thus, applying equations (9.4.1), we have that, for B_{2n_1-1} (B_{2n_2-1}) and n_1 and n_2 large, μ and σ^2 at $p = 0.5$ are approximately given by

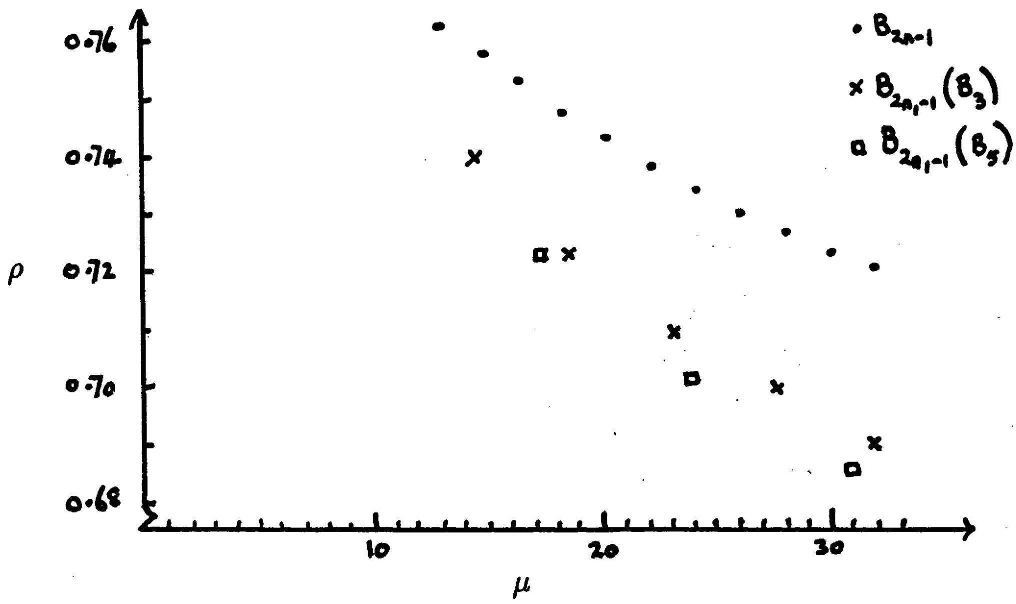
$$\mu \doteq 4n_1n_2$$

$$\text{and} \qquad \sigma^2 \doteq \left(1 - \frac{2}{\pi}\right)(4n_1n_2)(1 + 2n_2)\,. \tag{9.4.2}$$

Thus, Figure 9.4.1 and equations (9.4.2) show that, when designing B_{2n_1-1} (B_{2n_2-1}) systems and variance of duration is an important concern (for a "given" μ), it is preferable (in terms of variance) that n_2 be no larger than necessary.

Figure 9.4.2

The limit value of the efficiency, ρ, for a "given" expected duration, μ, for "best-of-..." uniformats as $p \to 0.5$.



The efficiency of these systems is now considered. Table 9.4.1 gives values for the limit of the efficiency at $0.5+\delta$ as $\delta \to 0$. These values can be obtained computationally by letting $\delta \to 0$ in the formula for efficiency (or by using the $P-D-I$ diagram and equation (4.5.2)) and by using the result that efficiency under nesting is multiplicative (Miles (to appear)). Figure 9.4.2 gives graphs, for these systems, for the limit of the efficiency at $0.5+\delta$ as $\delta \to 0$ against the expected duration at $p = 0.5$. It follows from Section 1.2 (page 3) that the limiting efficiency of B_{2n-1} at $p = 0.5$ tends to 0.25 as $n \to \infty$. Thus, it can be seen that the above rule for variance that "n_2 be no larger than necessary" also acts, for a "given" μ, in the interests of greater limiting efficiency.

Our overall conclusion is that, if efficiency or variance of duration (near $p = 0.5$*) were the only functions of interest, nested* B_{2n-1} *systems would not be used. However, nested systems are required in order to allow either player to overcome a period of poor play.*

An example of two scoring systems based on the tie-breaker game in tennis is now given.

Example 9.4.1. Using the notation TTB_{13} to represent a "truncated tie-breaker game of duration at most 13 points" (operating like a best-of-13 points system), consider the two systems

$$\text{(i)}\quad B_3\big(B_7(TTB_{13})\big)$$
$$\text{and}\qquad \text{(ii)}\quad B_7\big(B_3(TTB_{13})\big)\ .$$

The purpose of truncating the tie-breaker was to make the expected durations of these two systems comparable to those of the present "best-of-3 tie-breaker sets" scoring system. Monte Carlo experiments were carried out for these two systems and Table 9.4.2 gives results in which each statistic is based on 10,000 simulated matches. The first server in each match was alternated and the first server of each truncated tie-breaker game within each match was also alternated. Table 9.4.2 shows that the variance of duration of system (i) is (typically) greater than that of system (ii), in agreement with the earlier results of this section. When p_a and p_b are large (e.g. $(p_a, p_b) = (0.8, 0.7)$), a comparison of Table 9.4.2 with Tables 8.8.1a and 8.8.1b shows that these two systems (i) and (ii) are more efficient than the present tennis scoring system. (This is essentially due to the fact that, when p_a and p_b are large, a tie-breaker game is a more efficient unit than one game of each of player A's and player B's service.) These two systems also have smaller variances of duration than the present tennis scoring system, but, unfortunately, the variances are still relatively large. *It appears that large variances will typically occur when B_3 is used on the outer nesting, $B_3\big(\cdot(\cdot)\big)$, and the inner units must result in a win or a loss. Section 9.5 will show that this variance of duration can in some cases be reduced by using units which can also result in a draw.*

9.5 B_{2n-1} Systems When Draws are Possible

9.5.1 Introduction

When designing systems based on B_{2n-1} units and a draw is possible for a unit, the results of Section 8.9 suggest the following rule if a major concern is to keep variance as small as possible. *In the event of a unit being drawn, no attempt to break the tie should be carried out but rather $\frac{1}{2}$ a unit awarded to each player (rather than no unit awarded to either player).* Of course, the overall outcome (after all $2n-1$ units have been played) could possibly be a draw in which case a form of tie-breaker would finally be required. The following section can be viewed as an example of the above rule.

Table 9.4.2
The proportion of matches won by the better player and the mean and the variance of duration of systems (i) and (ii).
(Simulation results.)

p_b	p_a	system	Proportion	mean	variance
0.5	0.5	(i)	0.4975	161.1	1383.6
0.5	0.5	(ii)	0.5043	160.6	1019.5
0.5	0.6	(i)	0.8879	142.4	1351.4
0.5	0.6	(ii)	0.8873	144.3	1133.0
0.5	0.7	(i)	0.9928	113.6	700.5
0.5	0.7	(ii)	0.9920	115.5	678.0
0.6	0.6	(i)	0.4952	161.7	1394.5
0.6	0.6	(ii)	0.5022	161.8	1021.6
0.6	0.7	(i)	0.8949	142.5	1392.9
0.6	0.7	(ii)	0.9000	144.4	1129.7
0.6	0.8	(i)	0.9958	111.7	614.7
0.6	0.8	(ii)	0.9959	113.6	637.2
0.7	0.7	(i)	0.4953	163.3	1399.7
0.7	0.7	(ii)	0.5041	163.8	1024.5
0.7	0.8	(i)	0.9156	141.9	1372.0
0.7	0.8	(ii)	0.9207	143.9	1159.9
0.8	0.8	(i)	0.5024	167.8	1483.4
0.8	0.8	(ii)	0.5000	168.6	1094.4

9.5.2 An alternative scoring system for tennis

The tennis scoring system proposed in Section 8.9 used the B_5 system on the outer nesting. Some players believe that it is preferable to use B_3 (not B_5) on the outer nesting. Their reasoning is that the situation of being down 1–0 in a B_3 system is far better than being down 2–0 in a B_5 system. The problem of Section 8.9 is now reconsidered with the added constraint that B_3 must be used on the outer nesting. Is it possible to take the present best-of-three tie-breaker sets scoring system, modify it by as little as possible to produce a new system with a similar expected duration and probability that each player wins but *having a smaller variance of duration*?

The scoring system considered in this section is as follows. Games are played as at present. These games make up sets which operate in the following way. A player wins the set when the games score reaches 6–0, 6–1, 6–2, 6–3 or 6–4 in that player's favour; if the game score reaches 5–5, the set has resulted in a draw. The next set is then commenced if necessary. A half-unit is awarded to a player as soon as that player's game score reaches 5 games and another half unit if he reaches 6 games. The match is over as soon as one player's score reaches a total of 2 units and that player is declared the winner. However, if after three sets have been played, the unit score is 1.5 to 1.5, a *24 point tie-breaker game* is played to determine the winner. (In a 24 point tie-breaker game the rules for service exchange are the same as in the present 12 point tie-breaker game; however, the winner is now the first player to score at least 13 points and be at least 2 points ahead of his opponent.) The first server in the first set is determined by a toss, whereas in every other set the first server is the person who received in the last game of the previous set.

Table 9.5.1

The proportion of matches won by player A and the mean and variance of the duration of the best-of-three sets system described in Section 9.5.2.

(Simulation results.)

p_b	p_a	Proportion	Mean	Variance
0.5	0.5	0.4983	169.2	1517.7
0.5	0.6	0.9085	146.9	1584.3
0.5	0.7	0.9949	115.8	910.5
0.6	0.6	0.4938	168.0	1315.5
0.6	0.7	0.8929	149.3	1271.6
0.6	0.8	0.9906	127.2	993.4
0.7	0.7	0.5035	165.7	803.3
0.7	0.8	0.8621	155.5	715.6
0.8	0.8	0.4999	163.7	307.0

Monte Carlo experiments were carried out for this scoring system and Table 9.5.1 gives results in which each statistic is based on 10,000 simulated matches.

The server in each match was alternated. A comparison of Table 9.5.1 with Tables 8.8.1a and 8.8.1b shows that this scoring system has values for μ and P which are similar to those of the present best-of-three tie-breaker sets. However, a comparison of Tables 9.5.1 and 8.8.1c shows that this scoring system has a smaller variance when p_a and p_b are both large. It can be seen however that *the reduction in variance for this system is not as great as was achieved by the system proposed in Section 8.9. This would appear to be the price paid for having B_3 rather than B_5 on the outer nesting.* It was noted that the system described in this section also had a distribution of duration which was negatively skewed when p_a and p_b were both greater than 0.7.

9.6 Summary

This chapter has recommended that, for a scoring system of a given expected duration, there are three main characteristics that should be considered — the variance of duration, the efficiency and the level of nesting required (if at all). In many cases the designer must make a trade-off between these characteristics. Although this chapter has been exploratory in nature, and was not intended to be a definitive treatment of the design of scoring systems, some relevant examples have been given. A thorough treatment of this topic may have to be to some extent subjective since, when comparing two proposed scoring systems, one of them may be preferable in terms of (say) two of the above-mentioned characteristics whilst the other scoring system may be preferable in terms of the remaining characteristic. Thus, two persons may not agree on which is the preferable of a set of systems.

The examples given in this chapter and in Section 8.9 have demonstrated a few main ideas.

1. Scoring systems with small or very small variances can be produced, although it appears to be at the expense of decreased efficiency. One method of variance reduction was by making points increasingly more important beyond a certain stage (Section 9.3). Another method which could sometimes be used (e.g. in bipoints) was to use units which have a reasonable probability of resulting in a draw (Section 9.5.2 (when p_a and p_b are both large)).

2. Section 9.4 demonstrated that nesting is carried out at the expense of efficiency and variance of duration. However, it was suggested that nesting is required in order that either player could more easily overcome a period of poor play. The theory and example in Section 9.4 indicated that the B_{2n_1-1} (B_{2n_2-1}) system has a smaller variance of duration than the B_{2n_2-1} (B_{2n_1-1}) system when

$n_1 > n_2$. The use of B_3 on the outer nesting, $B_3\big(\cdot(\cdot)\big)$, which, from the player's viewpoint, maximizes the period of poor play possible, results in a scoring system with a large variance of duration in most cases. (A situation in which the variance may not be large is when the units, $\big(\cdot(\cdot)\big)$, of the above nested system, have a reasonable probability of resulting in a draw.)

3. Section 9.2 demonstrated that the $P - D - I$ diagram can be quite useful when analysing the components that make up a scoring system (e.g. a game of tennis, etc.). Simulation was also seen to be a very useful tool since exact analyses of scoring systems are typically quite cumbersome (e.g. Chapter 8). Judicious use of these two tools is recommended in further work.

4. Sections 8.9 and 9.5.2 indicated that, other things being equal, durations with negatively skewed distributions are preferable since potentially long matches are then quite effectively truncated.

Considerably more research can obviously be carried out on this topic of design. In fact, there would appear to be almost no limit to the number of variations in scoring systems that one might consider. My supervisor, Dr Roger Miles, is at present carrying out further investigations.

CHAPTER 10
THE IMPORTANCES OF POINTS AND SETS OF POINTS WITHIN A SCORING SYSTEM

10.1 Introduction

Morris (1977) defined the importance of a point and established an approximate relationship between the increased probability of winning a point or set of points and the increased probability of winning a match. A corresponding exact relationship is established in Section 10.2; a relationship which is generalized to the case in which p (or p_a and p_b) can take many different values and any arbitrary scoring system can be used including those that are not "fair". In Section 10.3 an alternative exact approach is taken to that of Section 10.2 and some results are given in Section 10.4. A numerical matrix/Markov chain approach is also used and it is seen that interactions between states occur; interactions that can be positive or negative. In Section 10.5 a different model to that of Morris is investigated; a model in which a player is assumed to be able to increase the p-value on a *fixed* number of points within a single realization of a match. Some possibly unexpected results are found for this model.

10.2 A Relationship Between the Increase in the Probability of Winning Each Point and the Increase in the Overall Probability of Winning.

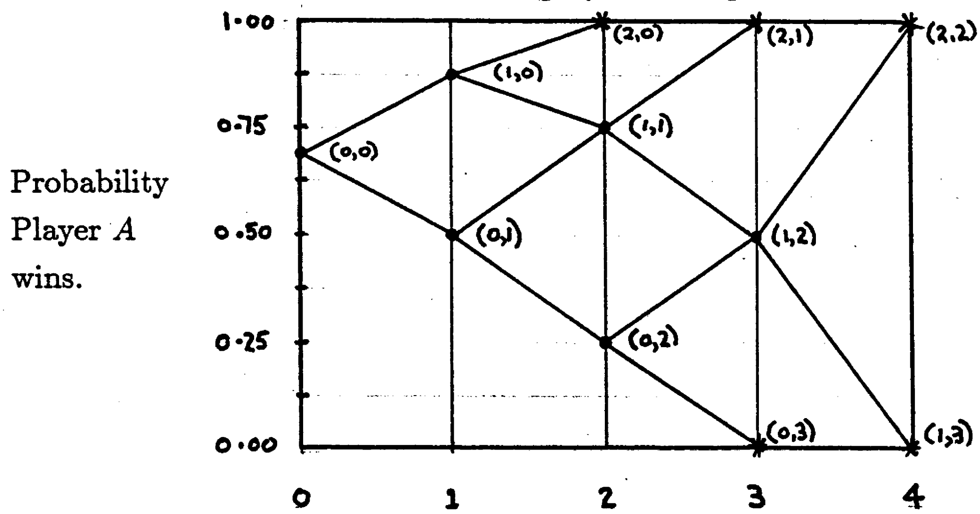
Consider an extension of "unipoints" in the following way. Suppose $0.5 + d_i$ is the probability that player A wins a point in state i. When all $d_i = 0$ (i.e. the players are equal), suppose I_i is the importance of (non-absorbing) state i and P is the overall probability that player A wins (the scoring system may not necessarily be fair (i.e. P may not equal 0.5)). For any given set of values for d_i ($-0.5 < d_i < 0.5$, all i), suppose P' is the overall probability that player A wins and n_i is the expected number of times that non-absorbing state i is visited. Then, we have

$$P'(1-P) - Q'P = \sum_{\text{all } i} n_i\big((\tfrac{1}{2} + d_i)(\tfrac{1}{2}I_i) - (\tfrac{1}{2} - d_i)(\tfrac{1}{2}I_i)\big)$$

where $Q' = 1 - P'$, since both sides of this equation represent, for this given set of values for d_i, the expected value of the sum of the steps in the $P - D - I$ diagram from start of play until absorption. Defining ΔP by $\Delta P = P' - P$, it follows that

$$\Delta P = \sum_{\text{all } i} n_i I_i d_i \,. \tag{10.2.1}$$

Figure 10.2.1.
The $P-D-I$ diagram for the unfair unipoints scoring system B_5^3
when both players are equal.



Number of points played.

Thus we have a generalization of the corresponding equation in Section 4.3.

Example 10.2.1.

Consider the unfair "unipoints" scoring system in which player A must win 2 points before player B wins 3 points, denoted by (say) B_5^3. Suppose state (i,j) represents the state in which player A has won i points ($i = 0,1,2$) and lost j points ($j = 0,1,2,3$).The $P-D-I$ diagram for this scoring system for the case of two equal players is shown in Figure 10.2.1. Suppose $p_{ij} = 0.5 + d_{ij}$ represents the probability that player A wins the point in state (i,j) and $q_{ij} = 1-p_{ij}$. Then, assuming points are independent, ΔP as defined above is given by

$$\Delta P = p_{00}p_{10} + p_{00}q_{10}p_{11} + q_{00}p_{01}p_{11} + p_{00}q_{10}q_{11}p_{12} + \\ q_{00}p_{01}q_{11}p_{12} + q_{00}q_{01}p_{02}p_{12} - \frac{11}{16}$$

and the r.h.s. of equation (10.2.1) is given by

$$\sum_{i=0}^{1}\sum_{j=0}^{2} n_{ij}I_{ij}d_{ij} = 1.\frac{3}{8}.d_{00} + p_{00}.\frac{1}{4}.d_{10} + q_{00}.\frac{1}{2}.d_{01} + (p_{00}q_{10} + q_{00}p_{01}).\frac{1}{2}.d_{11} \\ + q_{00}q_{01}.\frac{1}{2}.d_{02} + (p_{00}q_{10}q_{11} + q_{00}p_{01}q_{11} + q_{00}q_{01}p_{02}).1.d_{12}\ .$$

It can be shown that these two expressions are equal, thus verifying equation (10.2.1) for this example.

Now consider an extension of "bipoints" in the following way. Suppose player A's (B's) probability of winning a point in non-absorbing state a, i (b, j) is $p+d_{a,i}$ $(p-d_{b,j})$. Again the $d_{a,i}$ and $d_{b,j}$ need not be small. When all $d_{a,i}$ and $d_{b,j}$ are zero (i.e. the players are equal), suppose $I_{a,i}$ and $I_{b,j}$ are the importances of the a- and b- states and P is overall probability that player A wins. Again, the game need not be fair (i.e. P may not equal 0.5). For any given set of values for $d_{a,i}$ and $d_{b,j}$ (subject to the constraints that all $p+d_{a,i}$ and $p-d_{b,j}$ are legitimate probabilities), suppose P' is the overall probability that player A wins and $n_{a,i}$ $(n_{b,j})$ is the expected number of times non-absorbing state a, i (b, j) is visited. Again defining $\Delta P = P' - P$, we have, as before,

$$P'(1-P) - Q'P = \sum_{\text{all } i} n_{a,i}\{(p+d_{a,i})(qI_{a,i}) + (q-d_{a,i})(-pI_{a,i})\}$$
$$+ \sum_{\text{all } j} n_{b,j}\{(q+d_{b,j})(pI_{b,j}) + (p-d_{b,j})(-qI_{b,j})\}$$

where $q = 1-p$ and $Q' = 1-P'$. Hence,

$$\Delta P = \sum_{\text{all } i} n_{a,i} I_{a,i} d_{a,i} + \sum_{\text{all } j} n_{b,j} I_{b,j} d_{b,j} \ . \tag{10.2.2}$$

Example 10.2.2.

Consider the unfair "bipoints" scoring system B_3 in which play starts with an a-point and the point-types alternate. Suppose a similar notation is used to that of the previous paragraph and Example 10.2.1. Thus, for example, $p_{a11} = p + d_{a11}$ $(p_{b10} = p - d_{b10})$ is player A's (B's) probability of winning a point when the score is (1,1) ((1,0)). Figure 10.2.2 gives the $P-D-I$ diagram for this scoring system for the case in which $p = \frac{2}{3}$ and the players are equal (i.e. all d_{aij} and d_{bij} are zero). Thus, assuming points are independent, ΔP as defined above is given by

$$\Delta P = p_{a00}q_{b10} + (p_{a00}p_{b10} + q_{a00}q_{b01})p_{a11} - (pq + p^3 + pq^2)$$

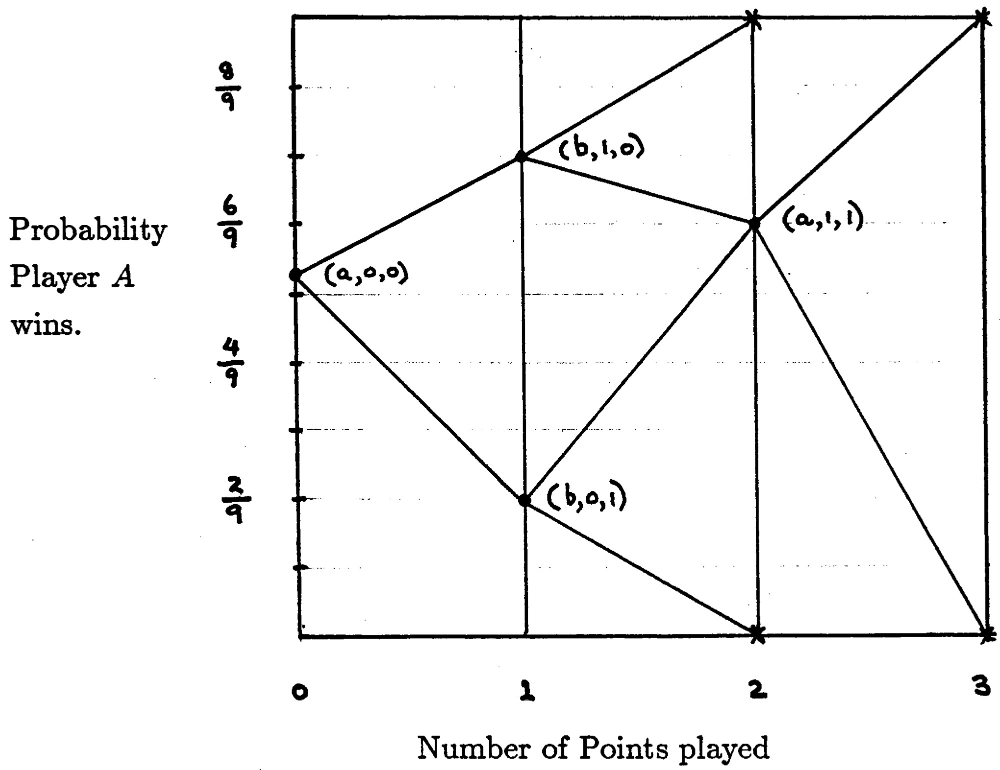
and the r.h.s. of equation (10.2.2) is given by

$$\sum_{\substack{\text{all}\\ \text{non-absorbing}\\ \text{states}}} n.I.d. = 1(1-2pq)(p_{a00} - p) + p_{a00}q(p - p_{b10}) + q_{a00}p(p - p_{b01})$$
$$+ (p_{a00}p_{b10} + q_{a00}q_{b01}).1.(p_{a11} - p)$$

since $I_{a00} = 1 - 2pq$, $I_{b10} = q$, $I_{b01} = p$ and $I_{a11} = 1$. Again, it can be shown that these two expressions are equal, verifying equation (10.2.2) for this example.

Figure 10.2.2

The $P-D-I$ diagram for the (unfair) bipoints scoring system B_3AL^a when the two players are equal and $p=\frac{2}{3}$.



10.3 An Alternative Approach to the Problem of Section 10.2.

Consider a scoring system with n states and suppose player A has probability p_i ($q_i = 1 - p_i$) of winning (losing) the point corresponding to non-absorbing state i. The outcomes of the points (or states) are assumed to be independent and suppose the probability that player A wins overall is $P = P\ (p_1, p_2, \ldots, p_n) = P(\mathbf{p})$. The (primary) "gain", $G_i(\mathbf{p}, \delta_i)$, player A obtains by increasing p_i from p_i to $p_i + \delta_i$ is defined by

$$G_i(\mathbf{p}, \delta_i) = P(p_1, \ldots, p_i + \delta_i, \ldots, p_n) - P(p_1, \ldots, p_i, \ldots, p_n)\ .$$

Now if the scoring system is such that state i can be visited at most once (i.e. P is linear in p_i), then using the Taylor's series expansion, we have

$$G_i(\mathbf{p}, \delta_i) = \frac{\delta P}{\delta p_i}.\delta_i\ . \tag{10.3.1}$$

(Note that if state i cannot be revisited and if all $p_j = p$ except p_i which is given by $p+\delta_i$, then it can be seen that the time-importance of state i as defined by Morris (1977) is equal to $\frac{\delta P}{\delta p_i}$ (evaluated when all $p_j = p$). Recall that the time-importance (Section 1.3) of a state is equal to the expected number of times that that state is entered in a single realization of the game multiplied by the importance of that state (each evaluated when all $p_j = p$).) It can be seen however, that if state i can be revisited, the higher partial differential terms in the Taylor's series expansion are necessary in equation (10.3.1). The "effort", $E_i(\mathbf{p}, \delta_i)$, required to obtain this "gain", $G_i(\mathbf{p}, \delta_i)$, is defined to be the (long-run) proportion of time the system is in state i. Thus, for the case in which state i cannot be revisited, $E_i(\mathbf{p}, \delta_i)$ is the probability that a realization of the game passes through state i divided by the expected duration when the p-values are p_j except p_i which equals $p_i + \delta_i$. The "return rate", $R_i(\mathbf{p}, \delta_i)$, is defined to be the ratio $G_i(\mathbf{p}, \delta_i)/E_i(\mathbf{p}, \delta_i)$. Thus, if state i cannot be revisited and if all $p_j = p$ except p_i which equals $p + \delta_i$, then it can be seen that

$$R_i(\mathbf{p}, \delta_i) = \mu_i I_i \delta_i$$

where I_i is the importance of state i when all $p_j = p$ and μ_i is the expected duration when all $p_j = p$ except p_i which equals $p + \delta_i$. Thus, in the approach of this section, the return rate of state i plays a corresponding role to that of the importance of state i (for the case in which state i cannot be revisited).

The (secondary) "gain", $G_{ij}(\mathbf{p}, \delta_i, \delta_j)$, player A obtains by increasing the p-values from p_i to $p_i + \delta_i$ and from p_j to $p_j + \delta_j$ on states i and j respectively, is defined by

$$\begin{aligned} G_{ij}(\mathbf{p}, \delta_i, \delta_j) = {} & P(p_1, \ldots, p_i + \delta_i, \ldots, p_j + \delta_j, \ldots, p_n) \\ & - P(p_1, \ldots, p_i, \ldots, p_j, \ldots, p_n) \, . \end{aligned}$$

If states i and j cannot be revisitied, then, using the Taylor's series expansion, we have

$$G_{ij}(\mathbf{p}, \delta_i, \delta_j) = \frac{\delta P}{\delta p_i}\delta_i + \frac{\delta P}{\delta p_j}\delta_j + \frac{\delta^2 P}{\delta p_i \delta p_j}\delta_i \delta_j \, . \qquad (10.3.2)$$

Again, if states i and j can be revisited, the higher partial differential terms are necessary. The (secondary) "effort", $E_{ij}(\mathbf{p}, \delta_i, \delta_j)$, is defined to be the (long-run) proportion of time the system is in state i or state j. For the case in which $\delta_i = \delta_j = \delta$, the (secondary) "return rate", $R_{ij}(\mathbf{p}, \delta, \delta)$ is defined to be the ratio $G_{ij}(\mathbf{p}, \delta, \delta)/E_{ij}(\mathbf{p}, \delta, \delta)$. Tertiary and higher-order effects can be correspondingly defined.

Some examples of the approach of this section and that of Section 10.2 are now given.

Example 10.3.1. The B_3 system.

Using p_{ij} to denote the probability that player A wins the point when the score is i and j points to players A and B respectively, we have

$$P(p_{00}, p_{10}, p_{01}, p_{11}) = p_{00}p_{10} + p_{00}q_{10}p_{11} + q_{00}p_{01}p_{11} \tag{10.3.3}$$

where $q_{ij} = 1 - p_{ij}$. Now, taking for example the case in which $p_{00} = p_{01} = p$, $p_{10} = p + \delta_{10}$ and $p_{11} = p + \delta_{11}$ we have, using equations (10.3.3) and (10.3.2), that the probability player A wins is given by

$$P = (p^2 + 2p^2q) + (p - p^2)\delta_{10} + 2pq\delta_{11} - p\delta_{10}\delta_{11}$$

where $q = 1 - p$. The last term in this equation indicates that the two states (1,0) and (1,1) have a negative interaction term. (States (0,1) and (1,1), for example, can be shown to have a positive interaction term.) Alternatively, using equation (10.2.1) and noting Figure 4.2.1, another expression for P is

$$\begin{aligned} P &= \tfrac{1}{2} + \sum_{\text{all } i} n_i I_i d_i \\ &= \tfrac{1}{2} + 1.\tfrac{1}{2}(p - \tfrac{1}{2}) + p.\tfrac{1}{2}.(p + \delta_{10} - \tfrac{1}{2}) + (1-p).\tfrac{1}{2}(p - \tfrac{1}{2}) \\ &\quad + \{(1-p)p + p(1 - p - \delta_{10})\}.1.(p + \delta_{11} - \tfrac{1}{2}) \,. \end{aligned}$$

These two expressions for P can be shown to be equal. It can be seen that the first expression for P makes use of differences in the p-values from p whereas the second expression makes use of the differences from 0.5. *It can also be seen that, in general, the second and higher order effects in the method of this section are "taken into account" by the n_i-values of the method of Section 10.2.*

A third and convenient numerical approach is to use absorbing Markov chain methods (Kemeny and Snell, 1960, p.43–49). Using their notation for the "primary" case in which p is increased to $p + \delta$ only for state (1,0) (row 2 of Table 10.3.1), we have that the transition matrix, $\mathbf{M}$, for this B_3 system, is given by

$$\mathbf{M} = \begin{array}{c} \text{win} \\ \text{loss} \\ (0,0) \\ (1,0) \\ (0,1) \\ (1,1) \end{array} \left[\begin{array}{cc|cccc} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & p & q & 0 \\ p+\delta & 0 & 0 & 0 & 0 & q-\delta \\ 0 & q & 0 & 0 & 0 & p \\ p & q & 0 & 0 & 0 & 0 \end{array}\right] = \left[\begin{array}{c|c} \mathbf{1} & \mathbf{0} \\ \hline \mathbf{R} & \mathbf{Q} \end{array}\right] .$$

Table 10.3.1
Some characteristics of the uniformat B_3 when $\mathbf{p} = \mathbf{0.47}$ and $\delta = 0.1$.

States	Gain	Effort	Return Rate
(0,0)	.04982	.39933	.12476
(1,0)	.02491	.19174	.12991
(0,1)	.02491	.20775	.11991
(1,1)	.04982	.19942	.24982
(0,0),(1,0)	.08003	.64155	.12475
(0,0),(0,1)	.07003	.56140	.12474
(0,0),(1,1)	.10024	.60067	.16688
(1,0),(0,1)	.04982	.39933	.12476
(1,0),(1,1)	.07003	.37582	.18634
(0,1),(1,1)	.08003	.42380	.18884
(0,0),(1,0),(0,1)	.10024	.80315	.12481
(0,0),(1,0),(1,1)	.12475	.82429	.15134
(0,0),(0,1),(1,1)	.12475	.77622	.16071
(1,0),(0,1),(1,1)	.10024	.60067	.16688

The mean number of times the game is in state (1,0) for example (given that the game started in (0,0)) is given by the second element of the first row of the matrix $\mathbf{N}$ defined by $\mathbf{N} = (\mathbf{1} - \mathbf{Q})^{-1}$. The average number of points played is the sum of the elements in this first row of $\mathbf{N}$. The probability that player A wins the game is given by the first element in the first row of the matrix $\mathbf{B}$ defined by $\mathbf{B} = \mathbf{NR}$.

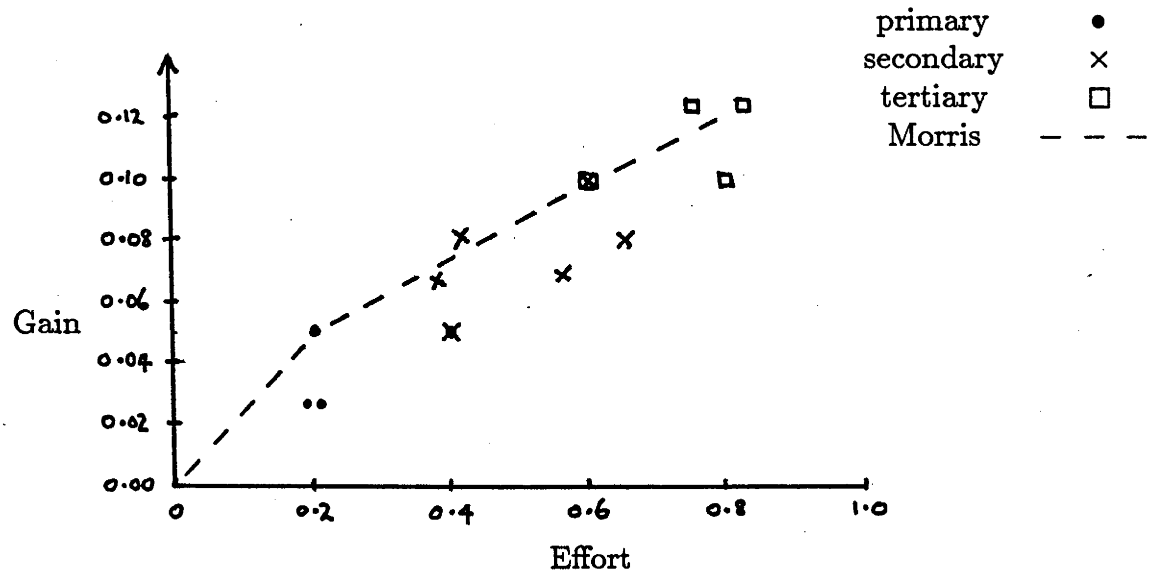
Any of the above three methods can be used and Table 10.3.1 gives the results for the case in which $\mathbf{p} = \mathbf{0.47}$ and $\delta = 0.1$. This table has some interesting features. For example,

(a) the two states (0,1) and (1,1) together have the highest secondary return rate even though state (0,1) has, by itself, the lowest primary return rate.

(b) the pair of states (1,0) and (0,1) have the same (secondary) return rate as does the individual (primary) state (0,0). Also, the triple (1,0), (0,1), (1,1) has the same return rate as the pair (0,0), (1,1). Figure 10.3.1, which shows the gain received for the effort made, can be used to determine the selection of the state(s) on which to optimally allocate the (given) extra effort. The

Figure 10.3.1

A Diagrammatical Representation of Table 10.3.1.



procedure of Morris (1977) would however recommend the selection of the states in their order of importance, which may not quite be optimal for a given effort. It can be seen that the above procedure also gives the player a wider selection of combinations of points on which to "try harder".

Example 10.3.2. The W_2 system.

Using p_i $(i = -1, 0, 1)$ to denote the probability that player A wins the point when player A has won i points more than player B, we have

$$P(p_{-1}, p_0, p_1) = \sum_{i=0}^{\infty} (q_0 p_{-1} + p_0 q_1)^i \, p_0 p_1$$

where $q_i = 1 - p_i$. A difference between this example and the previous one is that a return to the states -1, 0 or 1 is possible and hence the higher partial derivatives exist so that equations (10.3.1) and (10.3.2) would need modifications. These are omitted here. However, the third approach outlined in Example 10.3.1 is quite suitable for this example. The transition matrix, $\mathbf{M}$, is given by

$$\mathbf{M} = \begin{array}{r} \text{win}\\ \text{loss}\\ -1\\ 0\\ 1 \end{array} \left[\begin{array}{cc|ccc} 1 & 0 & 0 & 0 & 0\\ 0 & 1 & 0 & 0 & 0\\ \hline 0 & q & 0 & p & 0\\ 0 & 0 & q-\delta & 0 & p+\delta\\ p & 0 & 0 & q & 0 \end{array}\right] = \left[\begin{array}{c|c} \mathbf{1} & \mathbf{0}\\ \hline \mathbf{R} & \mathbf{Q} \end{array}\right]$$

where the above matrix applies to the (primary) case in which p is increased to $p+\delta$ for state 0. Table 10.3.2 gives the results for this example when $\mathbf{p} = \mathbf{0.47}$ and $\delta =$

0.1. It can be seen for example that states -1 and $+1$ have a positive interaction term and that states -1 and 0, the two states with the lowest primary return rates, produce, as a pair, a greater return rate than either of them individually and the largest of all the secondary return rates.

Table 10.3.2

Some characteristics of the uniformat W_2 when $\mathbf{p} = \mathbf{0.47}$ and $\delta = 0.1$.

States	Gain	Effort	Return Rate
-1	.05199	.26500	.19617
0	.10012	.50000	.20025
1	.04794	.23500	.20400
-1,0	.15144	.71500	.21180
-1,1	.10012	.50000	.20025
0,1	.14752	.78500	.18792

Example 10.3.3. A Nested Scoring System.

Consider a scoring system for a "match" with m "game-states", and suppose player A has probability p_i ($q_i = 1 - p_i$) of winning (losing) the ith game-state. The probability that player A wins the match, P, is given by

$$P = P(p_1, p_2, \ldots, p_m) .$$

Suppose the ith game-state has n "point-states" and player A has probability r_j^i ($s_j^i = 1 - r_j^i$) of winning (losing) the jth point-state within the ith game-state. The probability that player A wins the ith game-state, p_i, is given by

$$p_i = p_i(r_1^i, r_2^i, \ldots, r_n^i) .$$

Then, for example,

$$\frac{\partial P}{\partial r_j^i} = \frac{\partial P}{\partial p_i} \frac{\partial p_i}{\partial r_j^i}$$

and

$$\frac{\partial^2 P}{\partial r^i_j \partial r^k_l} = \frac{\partial^2 P}{\partial p_i \partial p_k} \frac{\partial p_i}{\partial r^i_j} \frac{\partial p_k}{\partial r^k_l} .$$

Primary, secondary and higher order gains for the point-states can be related to these and higher order partial derivatives.

Nested scoring systems can be conveniently handled using the third approach of Example 10.3.1. For example, consider a "match" consisting of the best-of-3 "games", each "game" being the "best-of-3" "points" (denote by $B_3(B_3)$). Using the notation (a, b, c, d) to represent the state in which player A has won a games and player B b games and that in the current game being played player A has scored c points and player B d points, the transition matrix, $\mathbf{M}$, for this match, is given by the following matrix.

$$\mathbf{M} = \begin{array}{l} \text{win} \\ \text{loss} \\ (0,0,0,0) \\ (0,0,1,0) \\ (0,0,0,1) \\ (0,0,1,1) \\ (1,0,0,0) \\ (1,0,1,0) \\ (1,0,0,1) \\ (1,0,1,1) \\ (0,1,0,0) \\ (0,1,1,0) \\ (0,1,0,1) \\ (0,1,1,1) \\ (1,1,0,0) \\ (1,1,1,0) \\ (1,1,0,1) \\ (1,1,1,1) \end{array} \left[\begin{array}{cc|cccc|cccc|cccc|cccc}
1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & 1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\hline
\cdot & \cdot & \cdot & p & q & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & q & p & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & p & \cdot & \cdot & \cdot & \cdot & q & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & p & \cdot & \cdot & \cdot & q & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\hline
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & p & q & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
p & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & q & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & p & \cdot & \cdot & \cdot & \cdot & q & \cdot & \cdot & \cdot \\
p & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & q & \cdot & \cdot & \cdot \\
\hline
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & p & q & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & q & p & \cdot & \cdot & \cdot \\
\cdot & q & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & p & \cdot & \cdot & \cdot & \cdot \\
\cdot & q & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & p & \cdot & \cdot & \cdot \\
\hline
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & p & q & \cdot \\
p & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & q \\
\cdot & q & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & p \\
p & q & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot
\end{array} \right]$$

(Zeros are represented by dots.)

Table 10.3.3 lists the gains, efforts and return rates for the third point in each game when $\mathbf{p} = \mathbf{0.47}$ and $\delta = 0.1$. It can be seen for example that, even though state (0,1,1,1) has a lower primary return rate than state (0,0,1,1), it produces a secondary return rate with state (1,1,1,1) greater than does (0,0,1,1). Hence, secondary (and higher order effects) are again involved.

Table 10.3.3
Some characteristics of the nested uniformat $B_3(B_3)$
when $\mathbf{p = 0.47}$ and $\delta = 0.1$.

State(s)	Gain	Effort	Return Rate
(1,1,1,1)=w	.02471	.03963	.62354
(1,0,1,1)=x	.01235	.03669	.33671
(0,0,1,1)=y	.02471	.07976	.30981
(0,1,1,1)=z	.01235	.04307	.28683
w and x	.03593	.07485	.48005
w and y	.04964	.11967	.41481
w and z	.03842	.08442	.45506
x and y	.03842	.12124	.31685
x and z	.02471	.07976	.30980
y and z	.03593	.11808	.30432

Example 10.3.4. An analysis of a single game of tennis.

Consider a single game of tennis in which player A has a probability 0.50 of winning a point and 0.596 (0.404) of winning a point on the half most (least) important points. It can be shown that, in such a game, the points at 30–40 (or "advantage receiver"), 30–30 (or "deuce"), 40–30 (or "advantage server"), 15–30, 15–15 or 30–15 are the half most important points in the sense that these states have a greater (primary) return rate than all the other states and they occur (almost exactly) half the time (in the long run). Thus, player A is "lifting (lowering)" his play from 0.50 by 0.096 on the half most (least) important points and it might thus be argued that this takes no extra overall effort. The third method of Example 10.3.1 can be used to show that player A has a probability of 0.556 of winning such a game, which is somewhat larger than 0.5. Thus, a player can obtain considerable gain by selecting the correct states on which to "lift and lower the p-values" even though this may require no overall extra effort in the long run. It can be shown that the first-order methods of Morris (1977) give a value of 0.566 rather than 0.556. The above exact procedure however takes into account second order, third order, ..., fifteenth order (there are 15 possible states) interactions.

10.4 Some Comments and Results.

Player A's gain is player B's "loss" and hence, when playing the uniformat B_3 for example,

$$G_{(1,0)}(\mathbf{p},\delta) = -G_{(0,1)}(\mathbf{q},-\delta)$$

where $\mathbf{q} = \mathbf{1} - \mathbf{p}$ and every element in the vector $\mathbf{p}$ is equal to p and every element in the vector $\mathbf{1}$ is equal to 1. A corresponding statement applies to the primary return rates and the higher order gains and return rates for any uniformat.

If player A has the choice of increasing p by δ on two states or by 2δ on the state with the highest (primary) gain or return rate, the choice with the greatest overall gain or return rate is typically the second of these. However, as a counter example, when playing B_3 and $\mathbf{p} = \mathbf{0.47}$, it has been shown that

$$G_{(0,0),(1,1)}(\mathbf{p},0.06,0.06) > \max_{\substack{i=(0,0),(1,0)\\(0,1)\text{ or}\\(1,1)}} \{G_i(\mathbf{p},0.12)\}\ .$$

Also, for example, when playing W_2 and $\mathbf{p} = \mathbf{0.47}$,

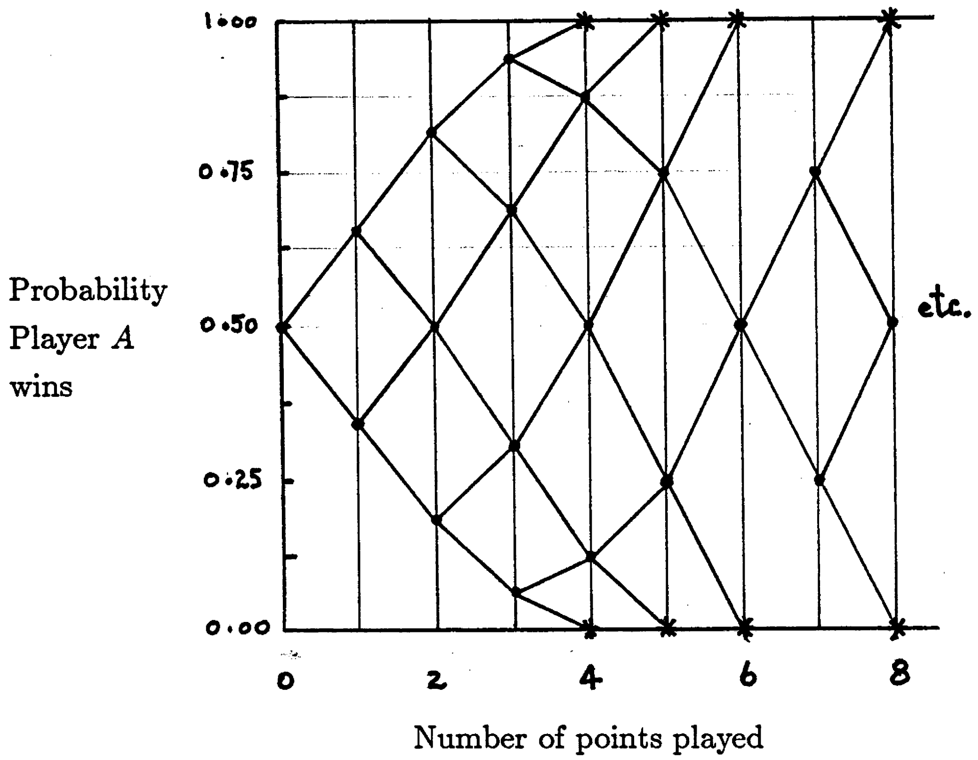
$$2R_{-1,0}(\mathbf{p},0.06,0.06) > \max_{i=-1,0\text{ or }1} \{R_i(\mathbf{p},0.12)\}\ .$$

10.5 A Different Model

We have considered an approach to the situation in which, *in the long run,* player A would increase p, the probability of winning a point to $p+\delta$ on a given percentage, $P_g\%$, of the total points played. This approach implies that in some matches player A will increase p less than $P_g\%$ of the time and in other matches greater than $P_g\%$ of the time. We now consider an alternative model in which, at some stage in the match (and it could be at the start), player A is able to increase p on exactly l points between that stage and the end of the match. Which l points should be selected?

Figure 10.5.1

The $P-D-I$ diagram for a game of tennis when $p=0.5$.



For simplicity we will take a unipoints approach. At any stage, t, of a match, the possible states (that the game could be in) can be ordered from that state from which player A has the lowest to that state from which player A has the highest probability of an overall win. (States such as (say) 30-all and deuce in tennis need to be considered as one state.) Suppose this ordering is $S_1, S_2, \ldots, S_{n_t}$ where n_t is the number of non-absorbing states at stage t. We now restrict to scoring systems which have the following property. For all t, a loss to player A in state S_i or a win to player A in state S_{i-1} at stage t ($i = 2, 3, \ldots, n_t$) leads to the same non-absorbing state at stage $t+1$. Thus n_{t+1} equals n_t, $n_t + 1$ or $n_t - 1$. Many scoring systems have this property (e.g. a game of tennis (see Figure 10.5.1), a set of classical or tie-breaker tennis (the state being the games' score), a game of table tennis, a game of North American squash, B_{2n-1} (see Figure 10.5.2 for B_9), W_n or W_m^n (see Figure 10.5.3 for W_4^2)). A few scoring systems do not have this property (e.g. international squash; R_n in which the winner is the first player to obtain a "run" of winning n points in a row (see Figure 10.5.4 for R_3)).

Figure 10.5.2
The $P-D-I$ diagram for B_9 when $p=0.5$.

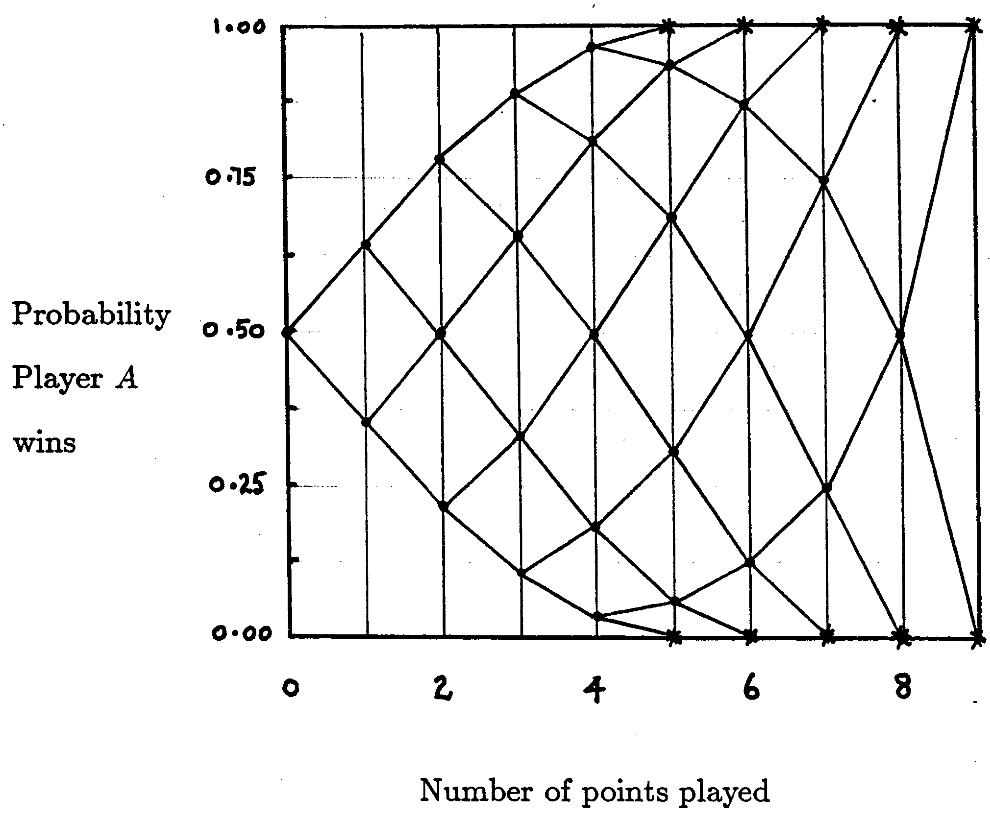


Figure 10.5.3
The $P-D-I$ diagram for W_4^2 when $p=0.5$.

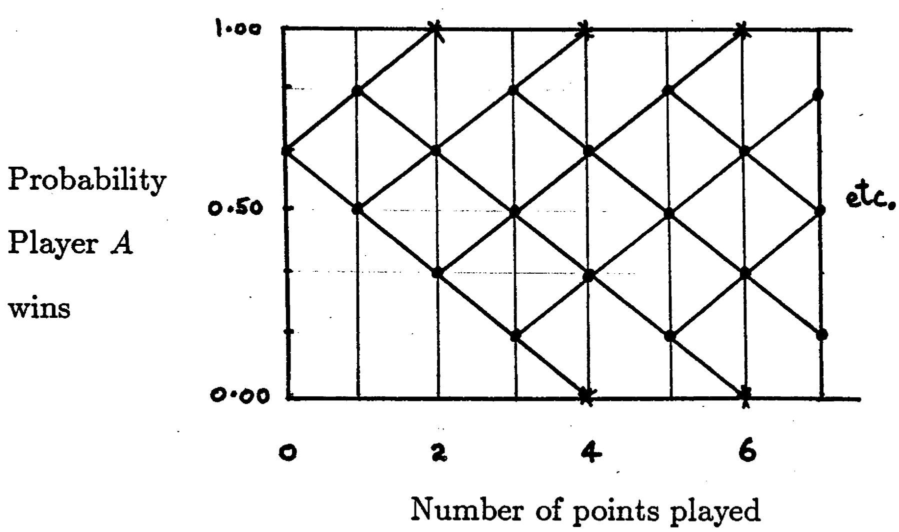


Figure 10.5.4

The $P - D - I$ diagram for R_3 when $p = 0.5$.

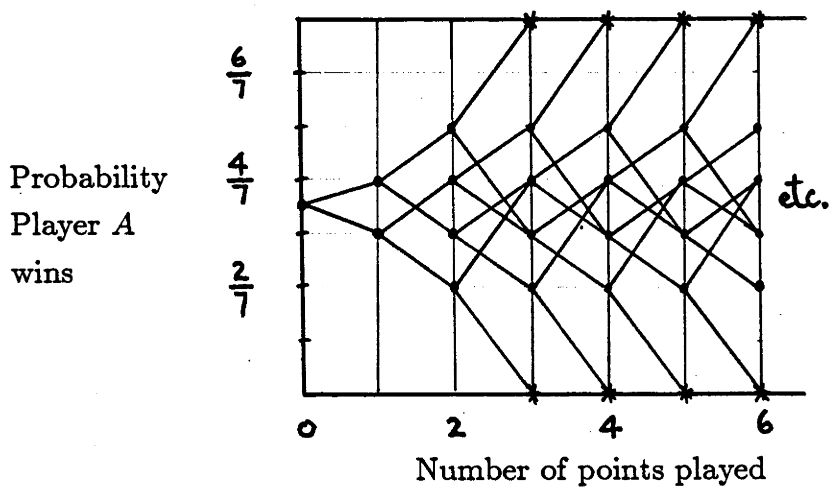
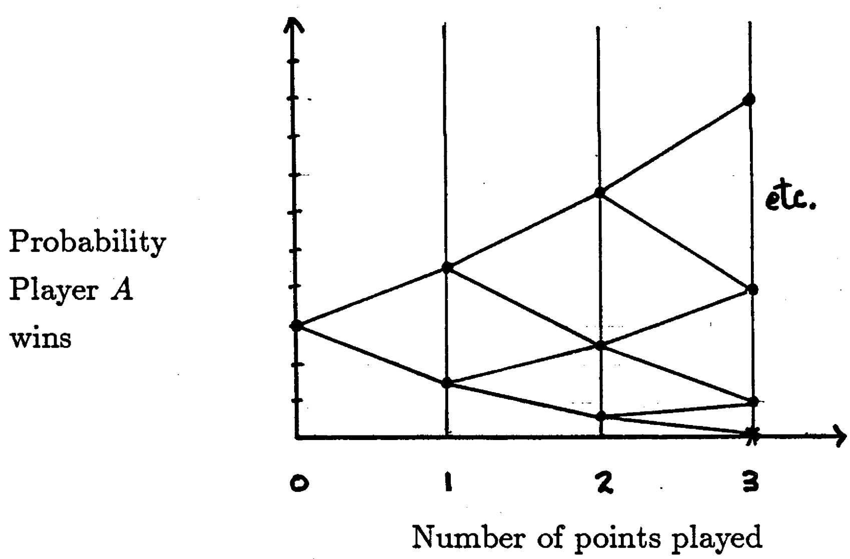


Figure 10.5.5

An example of a section of a $P - D - I$ diagram with a minimum of 3 points to an absorbing barrier (when $p = 0.5$).



We now suppose that at some stage there is a minimum of m points to an absorbing barrier (see Figure 10.5.5 for an example in which $m = 3$) and that player A can increase p by δ on exactly l ($l \leq m$) of these points. Then it can be seen that it does not matter which l points out of the next m points are selected since the distribution of the outcome over the $m+1$ possible states m points later is unaffected by the ordering of the p-values. Such a policy, however, may not be an optimal one, as the following example demonstrates.

Example 10.5.1.

Consider the rather impractical uniformat W_{32} units which was represented diagramatically in Figure 9.3.1 and discussed in Section 9.3. For this uniformat, which does not satisfy the property mentioned in the second paragraph of this section, it can be shown, using dynamic programming methods that, when (p, δ, l) is equal to (say) (0.3,0.5,3) and the current state is +4 or −4 units, the optimal decision for player A is to *not* increase p by δ even though there is a minimum of only 3 points to an absorbing barrier in each case. This might be seen as rather strange but the reason for this fact is that it is preferable for player A to save this increase in p by δ for a (possible) subsequent more (or much more) important point.

We now consider two particular scoring systems which possess the property mentioned in the second paragraph of this section.

(i) B_{2n-1}. Suppose that at some stage player A (B) needs to wins a (b) more points in order to win. Then the maximum number of points remaining is $a+b-1$. Supposing player A's probability of winning each such point is $p_1, p_2, \ldots, p_{a+b-1}$, it can be seen that player A's probability of an overall win is not dependent on the ordering of these p-values. (This is a generalization of the results of Kingston (1976) and Anderson (1977).) Thus, if player A can increase p by δ on exactly l ($l \leq a+b-1$) of the remaining points, player A's probability of an overall win is unaffected by which l are selected. (Note however that by deferring the increase in p until the last l points, the strategy with the greatest return rate is produced since often player A will not need to increase p on all (or possibly any!) of the l points. Note also that the expected duration (from that stage until the end) is however affected by the ordering of the p-values.)

(ii) W_n and $W_{n_2}^{n_1}$. Suppose that at some stage player A's losing boundary is a minimum of m points away and l is the number of points on which player A can lift p by δ. Since points are equally important for W_n (or $W_{n_2}^{n_1}$) when $p = 0.5$, it can be seen from equation (10.2.1) that an optimal policy is to make use of all of the l possible increases in p prior to a possible loss. Thus, if $l < m$, an optimal policy is to increase or not increase p on the very next point and then reassess the situation given the new values of l and m after the present point has been played. Alternatively, if $l \geq m$, p must be increased on the very next point and then the situation reassessed as above. It can be seen that such a policy requires the p-value to be increased if $l > m$ or as soon as l becomes equal to m at any stage.

10.6 Summary

Following on from the first order methods of Morris (1977), Section 10.2 gave an exact relationship between the increase in the probability of winning each point (each point could have a different probability!) and the increase in the overall probability of winning. An alternative exact approach to this problem was then given in Section 10.3. It was shown that two states could have a positive or negative interaction and that, for more than two states, there were higher order interactions. Thus, as was seen in Figure 10.3.1 for example, when considering increasing the p-values on certain states, the procedure of introducing these states in the order of their importances (as in the paper by Morris (1977)) might not be strictly optimal due to these second and higher order effects. Also, when allowance was made for these effects, different selections of states producing the same gains or return rates were sometimes possible. The matrix/Markov chain approach of Section 10.3 was seen to be a useful method which took account of these second and higher order effects. This approach could in fact be used to analyse any specific scoring system. As was seen in Example 10.3.4 and Section 1.3, a player could increase his probability of winning a match by lifting the p-value on the more important points and correspondingly lowering the p-value on the less important points. It could be argued that, in the long run, such a strategy takes no extra overall effort. It can be seen however that, in using such a strategy, some matches would require lifting the p-value more often than in others. Section 10.5 investigated an alternative model in which, at some stage in the match, player A was able to increase p on *exactly* l points (for some fixed l). It was noted that, for many unipoints scoring systems, if at the same stage in the match, m is the minimum possible number of points left to play and $l \leq m$, it does not matter which l out of the next m points are selected on which to increase p. It was also shown that, for the B_{2n-1} system in particular, it does not matter which l points are selected or the value of m. For the W_n or $W_{n_2}^{n_1}$ system however, the player must make use of all the l possible increases in p prior to a possible loss. This "exactly l points (for some fixed l)" model is however rather artificial since a player would typically wish to increase the p-value on more occasions in a longer game than in a shorter game.

CHAPTER 11
COUNTBACK METHODS AND RELATED ASPECTS OF TEAM PLAY

11.1 Introduction

In this chapter we consider a match between two teams A and B, each consisting of N players, in which the pairs of players $A_i\text{v}B_i$ $(i = 1, 2, \ldots, N)$ play individual matches. It is assumed that each pair of players uses the same scoring system, *all* of the N individual matches are first completed and *then* the results of these individual matches are combined (using a countback rule) in order to determine the winning team. A physical analogy would be the situation in which the N individual matches are played simultaneously on N parallel courts and no knowledge of how these individual matches are proceeding is exchanged between the courts until all individual matches have been completed when the countback rule finally goes into operation. It is recognized that there are situations in practice in which only one court is used, the individual matches are played sequentially with the order of play being determined by the team captains and the overall match finishing as soon as one team has an unbeatable lead. Such situations have multiple values for their expected duration (depending on the order of the possible N individual matches) and hence have multiple efficiencies, thus considerably complicating any mathematical analysis. For this reason, those situations in which all of the individual matches may not be completed are not considered in this chapter. The complete record of individual matches is considered necessary as it helps clubs and associations measure the individual performances of each of the players; measures which the players themselves desire and the clubs require for the selection of future teams.

For the team situation described in the previous paragraph, the following question naturally arises. How should the individual results be combined in order to determine the winning team? Typically, and taking (say) squash as an example, the winning team is the one with the greater number of individual player wins; if these are equal, it is the team with the greater number of games won; and if these are equal, it is the team with the greater number of points won; and if these are equal, it is a draw. Is there a better way than this "downward nested" method for determining the winning team? Also, are there preferable scoring systems that should be used for the individual matches? These questions are addressed in this chapter and we will restrict attention to the case in which each pair of players is playing unipoints. The constant probability that player A_i wins a point is denoted by p_i $(i = 1, 2, \ldots, N)$ and team A is the better team if $\sum_{i=1}^{N} p_i > \frac{N}{2}$.

11.2 Two Players Per Team

The case of two players per team is firstly considered. Many (single nested) scoring systems have the following property: *team A wins (loses) an individual match if and only if it wins (loses) more points in that match.* For example, it is true for the following single nested scoring systems: a game of tennis, a game of squash, a game of table-tennis, B_{2n-1}, W_n,.... It can be seen however that it is not necessarily true for the single nested scoring system R_n, in which the winner is the first player to obtain a run of n point-wins in a row. Also, it is not necessarily true for many double and triple nested scoring systems such as the double nested match of squash (rallies - games - match) and the triple nested match of tennis (points - games - sets - match). In the following examples we will restrict ourselves to scoring systems in which the above property is true. We will also assume that both individual matches use the same scoring system and that the outcomes of the two individual matches are independent.

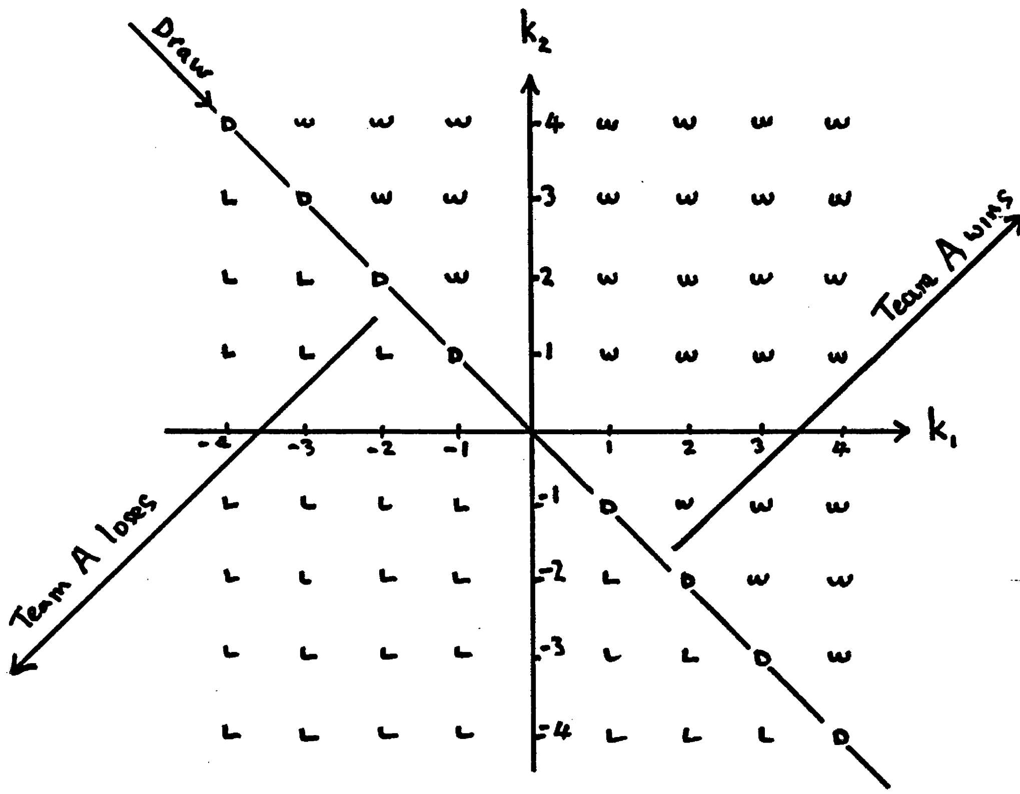
We will also restrict attention to the case in which the teams win and lose an individual match each and suppose the countback system breaks this tie by awarding the match to the team which wins the greater number of points in total. We now find the condition under which such a countback rule favours the better team. Suppose $f_{k_i}(p_i)$ is the probability that player A_i wins or loses with a point-difference of k_i points in player A_i's favour $(k_i = \ldots, -2, -1, 1, 2, \ldots;\ i = 1, 2)$. Figure 11.2.1 gives a diagrammatical representation of the possible outcomes. In this figure, quadrant 1 (3) represents the situation in which team A wins (loses) both individual matches and quadrant 2 (4) represents the case in which A_1 loses and A_2 wins (A_1 wins and A_2 loses).

Now let us assume without loss of generality that team A is the better team. Then, restricting attention to quadrants 4 and 2, the condition under which such a countback method favours team A is

$$\begin{aligned}&\sum_{k=1,2,\ldots}\ \sum_{j=k+1,k+2,\ldots}\left[f_j(p_1)f_{-k}(p_2)+f_j(p_2)f_{-k}(p_1)\right]\\ &\quad>\sum_{k=1,2,\ldots}\ \sum_{j=k+1,k+2,\ldots}\left[f_{-j}(p_2)f_k(p_1)+f_{-j}(p_1)f_k(p_2)\right].\end{aligned} \tag{11.2.1}$$

We now examine a few typical scoring systems to see whether this inequality is satisfied when team A is the better team (i.e. $p_1 + p_2 > 1$).

Figure 11.2.1
A diagrammatical representation of the possible outcomes in a match between two teams of two players each.



Example 11.2.1 Consider the B_{2n-1} system. From the distribution of duration, we have

$$f_j(p_i) = C_{n,j} p_i^n q_i^{n-j} \quad j = 1, 2, \ldots, n\,; \quad i = 1, 2$$

and

$$f_{-j}(p_i) = C_{n,j} q_i^n p_i^{n-j} \quad j = 1, 2, \ldots, n\,; \quad i = 1, 2$$

where $q_i = 1 - p_i$ and $C_{n,j}$ is a constant involving n and j. Now, since

$$\frac{f_j(p_i)}{f_k(p_i)} = \frac{C_{n,j}}{C_{n,k} q_i^{j-k}} \qquad (i = 1, 2)$$

and

$$\frac{f_{-j}(p_i)}{f_{-k}(p_i)} = \frac{C_{n,j}}{C_{n,k} p_i^{j-k}} \qquad (i = 1, 2)\,,$$

it can be shown that, for each of the three possible cases $(p_1 > \frac{1}{2}, p_2 > \frac{1}{2}), (p_1 > \frac{1}{2}, p_2 < \frac{1}{2})$ and $(p_1 < \frac{1}{2}, p_2 > \frac{1}{2})$,

$$\left.\begin{array}{l} f_j(p_1) f_{-k}(p_2) > f_{-j}(p_2) f_k(p_1) \\ f_j(p_2) f_{-k}(p_1) > f_{-j}(p_1) f_k(p_2) \end{array}\right\} \quad \begin{array}{l} k = 1, 2, \ldots, n-1 \\ j = k+1, \ldots, n \end{array},$$

provided $p_1 + p_2 > 1$, and hence inequality (11.2.1) is satisfied for this example.

Example 11.2.2 Consider two games of tennis. From the distribution of duration of a game of tennis (see, for example, Pollard (1983, p.497)), we have, for $i = 1, 2$,

$$\begin{aligned}
f_4(p_i) &= p_i^4 \\
f_3(p_i) &= 4p_i^4 q_i \\
f_2(p_i) &= 10p_i^4 q_i^2/(1 - 2p_i q_i) \\
f_{-4}(p_i) &= q_i^4 \\
f_{-3}(p_i) &= 4q_i^4 p_i \\
f_{-2}(p_i) &= 10q_i^4 p_i^2/(1 - 2p_i q_i)
\end{aligned}$$

where $q_i = 1 - p_i$. Unlike Example 11.2.1, each component on the left-side of the inequality (11.2.1) is not necessarily greater than the corresponding component on the right side when $p_1 + p_2 > 1$ (e.g. when $p_1 = 0.2$ and $p_2 = 0.9$, $f_3(p_1)f_{-2}(p_2) < f_{-3}(p_2)f_2(p_1)$). However, it has been shown numerically that,

$$f_j(p_1)f_{-k}(p_2) + f_j(p_2)f_{-k}(p_1) \gtrless f_{-j}(p_2)f_k(p_1) + f_{-j}(p_1)f_k(p_2)$$

when $p_1 + p_2 \gtrless 1$ and $(k, j) = (2, 3), (2, 4)$ and (3,4). Hence, inequality (11.2.1) is satisfied for two games of tennis.

Example 11.2.3. Consider two (first-to-9) games of international squash. Pollard (1983b), making use of an occupancy problem (see Feller (1957, p.36)), found expressions for the joint distribution of the score and the duration of a game of squash. More particularly, for the case in which player A serves first and has a probability p of winning any rally, he showed, for $j = 1, 2, \ldots, 8$,

$$\begin{aligned}
P(\text{player } A \text{ wins 9–0}) &= \frac{p^9}{(1 - pq)^9} \\
P(\text{player } A \text{ wins } 9 - j) &= \frac{\sum_{k=0}^{j-1} \binom{j-1}{k} \binom{9}{k+1} p^{10+k} q^{j+1+k}}{(1 - pq)^{9+j}} \\
P(\text{player } A \text{ loses 0–9}) &= \frac{q^{10}}{(1 - pq)^9} \\
P(\text{player } A \text{ loses } j - 9) &= \frac{\sum_{k=1}^{j} \sum_{l=0}^{k-1} \binom{k-1}{l} \binom{9}{l+1} q^{11+l} p^{j+l+1}}{(1 - pq)^{9+k}} + \frac{q^{10} p^j}{(1 - pq)^9}
\end{aligned}$$

where $q = 1 - p$. The first three expressions above are also given in a paper by Phillips (1978); the fourth is a different expression but equal to that obtained by Phillips.

As in Example 11.2.2, it has been shown numerically that, for the case in which team A serves first in one match and team B first in the other (a "fair" system), if $p_1 + p_2 \gtrless 1$, then

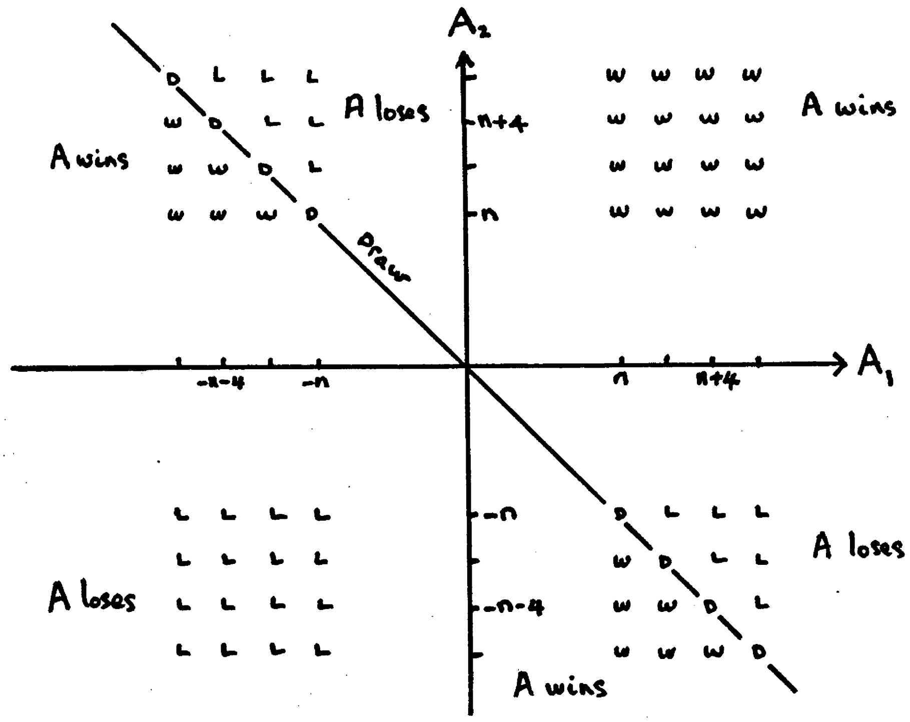
$$f_j(p_1)f_{-k}(p_2) + f_j(p_2)f_{-k}(p_1) \gtrless f_{-j}(p_2)f_k(p_1) + f_{-j}(p_1)f_k(p_2)$$

for $k = 1, 2, \ldots, 8$ and $j = k+1, k+2, \ldots, 9$ and hence inequality (11.2.1) is satisfied for two games of squash.

Thus, for the examples considered, it has been shown that, in the event of a tie between two teams of two players, the match should be awarded to that team which won more points. For those examples, this result is equivalent to awarding the match to the team which won more points overall (and ignoring individual results) since the property mentioned at the beginning of Section 11.2 is true in each case. It seems reasonable to believe that, for most (and if not all) scoring systems used in practice (even those in which the above-mentioned property is not true), such a countback rule would favour the better team.

Figure 11.2.2

A win-loss-draw diagram for the W_n system



We now consider the W_n system as a separate case. Here the previous countback system giving rise to inequality (11.2.1) is no longer applicable since, in the event of a tie, the two teams must have scored an equal number of points. Suppose we now break the tie by using *a countback rule in which the winning team is the one which won in the shorter duration.* Figure 11.2.1 is now no longer applicable and is replaced by Figure 11.2.2. Suppose $d_{n+k}(p_i)$ is the probability player A_i wins in $(n+k)$ points and $d_{-(n+k)}(p_i)$ is the probability player A_i loses in $(n+k)$ points $(k = 0, 2, 4, \ldots)$. Assuming, without loss of generality, that team A is the better team, we require, in the event of a tie, a higher probability of the match being awarded to team A. That is, we require

$$\begin{aligned}&\sum_{k=0,2,4,\ldots}\ \sum_{j=k+2,k+4,\ldots}\left\{d_{n+k}(p_1)d_{-(n+j)}(p_2)+d_{n+k}(p_2)d_{-(n+j)}(p_1)\right\}\\ &>\sum_{k=0,2,4,\ldots}\ \sum_{j=k+2,k+4,\ldots}\left\{d_{-(n+k)}(p_2)d_{n+j}(p_1)+d_{-(n+k)}(p_1)d_{n+j}(p_2)\right\}.\end{aligned}\tag{11.2.2}$$

Now, for $i = 1, 2$ and $l = j$ or k,

$$d_{n+2l}(p_i) = K_{n,l}p_i^{n+l}q_i^l$$

and

$$d_{-(n+2l)}(p_i) = K_{n,l}q_i^{n+l}p_i^l$$

where $K_{n,l}$ is a constant and $l = 0, 1, 2, \ldots$. Substituting the above into (11.2.2), the left side equals

$$\sum_{k=0,1,2,\ldots}\ \sum_{j=k+1,k+2,\ldots} K_{n,k}K_{n,j}(p_1q_1p_2q_2)^k\left\{(p_1q_2)^n(p_2q_2)^{j-k}+(p_2q_1)^n(p_1q_1)^{j-k}\right\}$$

and the right side equals

$$\sum_{k=0,1,2,\ldots}\ \sum_{j=k+1,k+2,\ldots} K_{n,k}K_{n,j}(p_1q_1p_2q_2)^k\left\{(p_1q_2)^n(p_1q_1)^{j-k}+(p_2q_1)^n(p_2q_2)^{j-k}\right\}.$$

Now, it can be shown that

$$p_1q_2 \gtrless p_2q_1 \quad \text{if and only if} \quad p_2q_2 \gtrless p_1q_1$$

provided $p_1+p_2 > 1$. It follows that the inequality (11.2.2) holds provided $p_1+p_2 > 1$ (except when $p_1 = p_2$ in which case (11.2.2) becomes an equality implying that there is no benefit in using this countback rule if the two individual matches are between the same pair of players $(p_1 = p_2)$. This is because we are assuming p_1 and p_2 are *constants.*).

Thus, with the W_n system, in the event of a tie, the team which won its individual match in the shorter duration should be awarded the match. Note that many scoring systems have a W_n component at the end of them. For example, a game of tennis, a classical set of tennis, a game of table-tennis, ... Thus, it seems quite reasonable to believe that the above rule would be appropriate for breaking a tie in cases such as (say) 21–23, 27–25 in table-tennis, (say) 4–2, 3–5 (points) in two games of tennis, or (say) 5–7, 11–9 in two classical sets of tennis. The same countback rule would also appear to be appropriate for a tie of (say) 6–9, 9–6 in two games of international squash, ...

11.3 Three Players Per Team

In Section 11.2 we noted that in the event of a tie between two teams of two players each, a countback involving the total number of points won by each team and/or the duration of each of the individual matches could be used to break the tie. However, when there are three players in each team a tie between the number of individual match wins by each team is no longer possible. Counting only individual matches, a team either wins 3–0 or 2–1 or loses 1–2 or 0–3. However, the scoring system could alternatively compare the total number of points won by each team rather than the number of individual match wins. This is one of the aspects that is investigated in this section. Three methods of scoring for the individual matches used in conjunction with three methods of countback are investigated. The three methods of scoring the individual matches that are considered are B_{2l-1}, W_m and P_n (in which, in each individual match, *exactly* n points are played) and the constants l, m, and n are selected in such a way that the expected durations of the systems are comparable (but not exactly equal). (It can be argued that, although P_n is not really a suitable scoring system for a single individual match (since often the final points need not be played as the game has already been won by one player), it has merit in the team situation as it predetermines that every player in the teams should play the same number of points. It also has merit in the individual round-robin and team round-robin competitions, for the same reason (although these two situations are not considered in this thesis).) The three methods of countback to be considered are:–

(i) the winning team is that team which won more individual matches. This is the usual method and is "downward nested".

(iia) the winning team is that team which won more points; in the event of a tie in the total number of points won, the winning team is that team which won more individual matches.

(iib) as in (iia) except that, in the event of a tie, the winning team is that team which won fewer individual matches.

Method (iib) might at first be seen as unusual, but it will become apparent that it has some advantages.

11.3.1 A comparison of the three countback systems for the case in which individual matches are B_{2l-1}.

For B_{2l-1}, supposing $P(i,j)$ is the probability that player A_i wins or loses with a points difference of j, we have

$$P(i,j) = \begin{cases} \binom{2l-1-j}{l-j} p_i^l q_i^{l-j} & j = 1,2,\ldots,l \\ \binom{2l-1+j}{l+j} q_i^l p_i^{l+j} & j = -1,-2,\ldots,-l \end{cases}$$

where $q_i = 1 - p_i$. Now supposing $P(i)$ is the probability player A_i wins, then $P(i) = \sum_{j=1}^{l} P(i,j)$ and the probability, P, that team A wins is

$$P(1)P(2) + P(1)P(3) + P(2)P(3) - 2P(1)P(2)P(3) \quad \text{under countback (i)}$$

$$\text{and} \sum_{j_1}\sum_{j_2}\sum_{\substack{j_3 \\ \text{s.t. } j_1+j_2+j_3>0}} P(1,j_1)P(2,j_2)P(3,j_3)$$

$$+ \sum_{j_1}\sum_{j_2}\sum_{\substack{j_3 \\ \text{s.t. } j_1+j_2+j_3=0 \\ \text{and } \prod_{i=1}^{3}(\text{sign}(j_i))<0}} P(1,j_1)P(2,j_2)P(3,j_3)$$

under countback (iia) and, for countback (iib), is the same as for countback (iia) except that the constraint $\prod_{i=1}^{3}(\text{sign}(j_i)) < 0$ in the second summation is replaced by $\prod_{i=1}^{3}(\text{sign}(j_i)) > 0$.

The case in which $l = 7$ (i.e. B_{13}) is now selected as a representative example. Two concepts are used to compare the three countback systems. Firstly, a "fair" system is one in which $P = 0.5$ when $\sum_{i=1}^{3} p_i = 1.5$. Secondly, following the work of Morris (1977), the magnification of a system is defined (in this chapter) as the increase in P when $\sum_{i=1}^{3} p_i$ is increased from 1.5 to $1.5 + \delta$. Ideally, the best of the three countback systems would be fair and would be the one which has the largest magnification. These three countback systems were compared numerically for the following four cases.

(a) Firstly, the very artificial case in which

$$p_1 = 0.5 + \delta_1$$

$$p_2 = 0.5 + \delta_1$$

$$p_3 = 0.5 + \delta_1$$

where δ_1 was given values 0 (0.01) 0.20. All three countback systems were fair; countback systems (iia) and (iib) had equal magnification and greater than that of countback (i). The reader is referred to the first and fifth rows of Table 11.3.1 for the specific cases in which $\delta_1 = 0$ and 0.05 respectively.

(b) Secondly, the rather artificial case in which

$$\begin{aligned} p_1 &= 0.5 + \delta_1 + 2\delta_2 \\ p_2 &= 0.5 + \delta_1 \\ p_3 &= 0.5 + \delta_1 - 2\delta_2 \end{aligned}$$

where δ_1 and δ_2 were given values 0(0.01)0.20 and 0(0.025)0.15 respectively. Two conclusions were reached. Firstly, that all three countback systems were fair and secondly, that countback (iib) had a slightly greater magnification than (iia) and also considerably greater than (i). The reader is referred to the second and sixth rows of Table 11.3.1 for the specific cases in which $(\delta_1, \delta_2) = (0, 0.1)$ and (0.05,0.1).

(c) Thirdly, a more practical case in which

$$\begin{aligned} p_1 &= 0.5 + \delta_1 + \delta_2 \\ p_2 &= 0.5 + \delta_1 + \delta_2 \\ p_3 &= 0.5 + \delta_1 - 2\delta_2 \end{aligned}$$

where δ_1 and δ_2 were given the values 0(0.01)0.20 and 0(0.025)0.25 respectively. For these cases, all three countback systems were not fair although countback (iib) was "less unfair" (i.e. P closer to 0.5 when $\delta_1 = 0$) than (iia) which in turn was less unfair than (i) ($P_{(i)} > P_{(iia)} > P_{(iib)} > 0.5$ when $\delta_1 = 0$ where subscripts are used to denote the countback system). Allowing for the different unfairnesses of the three systems, it was noted that (iib) had a slightly greater magnification than (iia) and a considerably greater magnification than (i). The reader is referred to the third and seventh rows of Table 11.3.1 for the cases in which $(\delta_1, \delta_2) = (0, 0.1)$ and (0.05,0.1).

(d) Finally, another practical case in which

$$\begin{aligned} p_1 &= 0.5 + \delta_1 + 2\delta_2 \\ p_2 &= 0.5 + \delta_1 - \delta_2 \\ p_3 &= 0.5 + \delta_1 - \delta_2 \end{aligned}$$

where δ_1 and δ_2 were given the values 0(0.01)0.20 and 0(0.025)0.15 respectively. Again, all three countback systems were not fair although countback (iib) was

Table 11.3.1
Some P and μ values for B_{13} using three
different countback systems (μ=expected duration).

P / μ		Countback (i)	Countback (iia)	Countback (iib)
(p_1, p_2, p_3)	(0.5,0.5,0.5)	0.500000 33.202148	0.500000 33.202148	0.500000 33.202148
	(0.7,0.5,0.3)	0.500000 30.670277	0.500000 30.670277	0.500000 30.670277
	(0.6,0.6,0.3)	0.616697 31.213568	0.567464 31.213568	0.537764 31.213568
	(0.7,0.4,0.4)	0.383303 31.213568	0.432536 31.213568	0.462236 31.213568
	(0.55,0.55,0.55)	0.709673 32.921612	0.720966 32.921612	0.720966 32.921612
	(0.75,0.55,0.35)	0.675133 30.537656	0.710482 30.537656	0.719034 30.537656
	(0.65,0.65,0.35)	0.787009 30.896937	0.771967 30.896937	0.752266 30.896937
	(0.75,0.45,0.45)	0.574455 31.212548	0.653491 31.212548	0.686131 31.212548

less unfair than (iia) which in turn was considerably less unfair than (i) (here $P_{(i)} < P_{(iia)} < P_{(iib)} < 0.5$ when $\delta_1 = 0$). Again it was noted that, except in the cases in which the unfairnesses of (i) and (iia) became considerable (i.e. when $\delta_1 = 0$ and δ_2 was large), (iib) had a slightly greater magnification than (iia) and a considerably greater magnification than (i). However, for each of these exceptions, $P_{(iib)} > P_{(iia)} > P_{(i)}$, demonstrating that the exceptions were caused by the relative distortions from fairness of the three systems. The reader is referred to the fourth and eighth rows of Table 11.3.1 for the cases in which $(\delta_1, \delta_2) = (0, 0.1)$

and (0.05,0.1).

Thus, for B_{13}, using fairness and magnification as our criteria, our conclusion is that countback (iib) is slightly preferable to (iia) and considerably preferable to (i). Other B_{2l-1} systems (e.g. B_3, B_5, B_7) were also considered and the same general conclusions reached. Thus, it seems reasonable to believe that the conclusions reached for B_{13} apply to all B_{2l-1} systems in general.

11.3.2 A comparison of B_{2l-1} with countback (iib) and P_n.

B_{13} and P_{11}, having comparable expected durations, are given as examples. It can be seen that when n is odd, P_n with countbacks (iia) and (iib) are equivalent since a tie in the number of points won by each team is not possible. Thus, the notation $P_n(ii)$ is used to represent the system P_n with countback (iia) or (iib) and similar notations used for other systems. For $P_n(ii)$, expressions for $P(i,j)$, the probability that player A_i wins or loses with a difference of j points ($j = -n, -n+2, \ldots, n-2, n$), can be written down by making use of the binomial distribution with parameters n and p_i. The countback systems were compared numerically for the four cases ((a), (b), (c) and (d)) of Section 11.3.1. Table 11.3.2 gives the results for the same parameter values as in Table 11.3.1 and the following general conclusions can be made:-

(a) $P_{11}(ii)$ is considerably less unfair than $P_{11}(i)$ and typically has a much greater magnification (see the first and second columns of Table 11.3.2).

(b) $P_{11}(ii)$ is considerably less unfair than $B_{13}(iib)$ (see the first to fourth rows of the second and third columns of Table 11.3.2).

(c) Typically, $P_{11}(ii)$ has a greater magnification than $B_{13}(iib)$ (see the second and third columns of Table 11.3.2). This is due, at least to some extent, to the decrease in μ (for the $B_{13}(iib)$ system as δ_2 increases) that can be seen, for example, in the fifth to eighth rows of the third column of Table 11.3.2. Also, a favourable property of the $P_{11}(ii)$ system is that the variation in P (as the δ_2 effect of Section 11.3.1 varies) is smaller than for the other systems considered (see the second column of Table 11.3.2).

From other examples considered, it appears reasonable to believe that the conclusions reached in this section for the specific example of the comparison of B_{13} and P_{11} also apply to the general case of the comparison of B_{2l-1} and P_n for values of l and n such that the expected durations are comparable. In general, for two such l and n, $P_n(ii)$ is considerably less unfair than $B_{2l-1}(iib)$ and typically has a slightly greater magnification. Thus, as the concept of fairness is essential

Table 11.3.2
Some P and μ values for $P_{11}(i)$, $P_{11}(ii)$ and $B_{13}(iib)$.

P / μ		$P_{11}(i)$	$P_{11}(ii)$	$B_{13}(iib)$
(p_1,p_2,p_3)	(0.5,0.5,0.5)	0.500000 33.000000	0.500000 33.000000	0.500000 33.202148
	(0.7,0.5,0.3)	0.500000 33.000000	0.500000 33.000000	0.500000 30.670277
	(0.6,0.6,0.3)	0.596818 33.000000	0.500452 33.000000	0.537764 31.213568
	(0.7,0.4,0.4)	0.403182 33.000000	0.499548 33.000000	0.462236 31.213568
	(0.55,0.55,0.55)	0.694966 33.000000	0.719138 33.000000	0.720966 32.921612
	(0.75,0.55,0.35)	0.667297 33.000000	0.730543 33.000000	0.719034 30.537656
	(0.65,0.65,0.35)	0.762379 33.000000	0.727787 33.000000	0.752266 30.896937
	(0.75,0.45,0.45)	0.583209 33.000000	0.727269 33.000000	0.686131 31.212548

to scoring systems, it could well be argued that $P_n(ii)$ is preferable to $B_{2l-1}(iib)$.

11.3.3 A comparison of W_m and P_n

For W_m, the three countback systems are identical and the probability player A_i wins is equal to $p_i^m/(p_i^m + q_i^m)$. W_4 and $P_{15}(ii)$ were taken as suitable specific examples for the comparison (since their expected durations were (at least to some extent) comparable) and were compared numerically for the four cases ((a), (b), (c) and (d)) of Section 11.3.1. Table 11.3.3a shows that W_4 has the unsatisfactory property that it is quite unfair in cases (c) and (d) when δ_2 is not small. Also, for W_4, the expected duration is quite dependent on δ_2 (and, in fact, δ_1). These two effects can be removed (to some extent) by selecting a small value for δ_2, so that the magnifications of the two systems can be reasonably compared. This is done in Table 11.3.3b in which $(\delta_1, \delta_2) = (0, 0.025)$ and (0.05,0.025), and it can be seen that the magnification of $P_{15}(ii)$ is less (but not by as much as may have been anticipated (by the optimal efficiency of W_m in unipoints)) than that of W_4.

Table 11.3.3a

Some P and μ values for W_4 and $P_{15}(ii)$

when $\delta_2 = 0.1$ and $\delta_1 = 0$ and 0.05.

P μ		$P_{15}(ii)$	W_4
(p_1, p_2, p_3)	(0.5,0.5,0.5)	0.500000 45.000000	0.500000 48.000000
	(0.7,0.5,0.3)	0.500000 45.000000	0.500000 34.694601
	(0.6,0.6,0.3)	0.500379 45.000000	0.706301 36.151424
	(0.7,0.4,0.4)	0.499621 45.000000	0.293699 36.151424
	(0.55,0.55,0.55)	0.750561 45.000000	0.771985 45.731535
	(0.75,0.55,0.35)	0.763040 45.000000	0.706484 34.314145
	(0.65,0.65,0.35)	0.759913 45.000000	0.862014 33.796266
	(0.75,0.45,0.45)	0.759585 45.000000	0.517931 38.292568

It seems reasonable to believe that these results also apply to a comparison of $P_n(ii)$ and W_m for general n and m such that the expected durations are comparable. Thus, our conclusions are:–

(a) $P_n(ii)$ is, in most practical cases, much less unfair than the comparable W_m system.

Table 11.3.3b
Some P and μ values for W_4 and $P_{15}(ii)$
when $\delta_2 = 0.025$ and $\delta_1 = 0$ and 0.05.

P μ		$P_{15}(ii)$	W_4
(p_1, p_2, p_3)	(0.5,0.5,0.5)	0.500000 45.000000	0.500000 48.000000
	(0.55,0.5,0.45)	0.500000 45.000000	0.500000 46.487690
	(0.525,0.525,0.45)	0.500005 45.000000	0.507211 46.849562
	(0.55,0.475,0.475)	0.499995 45.000000	0.492789 46.849562
	(0.55,0.55,0.55)	0.750561 45.000000	0.771985 45.731535
	(0.6,0.55,0.5)	0.751290 45.000000	0.762800 44.645907
	(0.575,0.575,0.5)	0.751110 45.000000	0.770144 44.815341
	(0.60,0.525,0.525)	0.751105 45.000000	0.759757 45.007779

(b) the magnification of $P_n(ii)$ is, surprisingly, only slightly less than that of the comparable W_m system.

Thus, since fairness is an important property that scoring systems should ideally possess, and since the variance of duration of the W_m system is known to be large whilst that of the P_n system is clearly zero, it could well be argued that the $P_n(ii)$ system is preferable to the comparable W_m system.

11.4 Summary

This chapter has considered a match between two teams A and B, each consisting of N players, in which pairs of players A_i v B_i ($i = 1, 2, \ldots, N$) played individual matches. Player A_i was assumed to have the (unipoints) probability p_i of winning a point against player B_i and team A was defined to be the better team if $\sum_{i=1}^{N} p_i > \frac{N}{2}$.

The case of 2 players per team was firstly considered in Section 11.2. Example 11.2.1 showed that, if each pair of players played according to the B_{2n-1} uniformat, then, in the event of team A and team B winning one individual match each, the countback rule which awards the overall match to the team which won the greater number of points in total, does in fact favour the better team. Examples 11.2.2 and 11.2.3 showed that the same result was true for the case in which each pair of players was playing a game of tennis or a game of squash. For each of these three examples the above rule was seen to be equivalent to a rule which awards the match to the team which won more points overall (and ignoring individual player results). A second rule was seen to be applicable to the case in which each individual match was played according to the W_n uniformat and team A and team B had won one individual match each. In this situation each team must have won the same number of points in total so that the above rule could not be used. In such a case, the appropriate rule was seen to be to award the overall match to that team which won its individual match in the shorter duration. It seems reasonable to believe that, when there are 2 players in each team, these two rules are useful countback rules for the many different scoring systems used in practice (and also for the cases in which the individual players are playing bipoints rather than unipoints).

The case of 3 players per team was considered in Section 11.3. The first rule of the previous paragraph (and not the second rule) could be extended to the case of 3 players per team and Section 11.3 considered the three situations in which the individual matches were B_{2l-1}, W_m or P_n (where, in P_n, each individual match consists of playing *exactly* n points). Two main countback rules were considered:— one based on individual matches won and a second based on total points won. For both B_{2l-1} and P_n, the countback rule based on points was seen to produce a fairer system and one which had greater magnification. A comparison of the two systems B_{2l-1} and P_n showed that the P_n system was fairer and had greater magnification than the comparable B_{2l-1} system. A comparison of the P_n and W_m systems showed that the P_n system was fairer and had a slightly less magnification than the comparable W_n system. Given the above favourable aspects of the P_n system and the fact that it has zero variance of duration as well as the fact that it is much more suited to the team situation than the individual situation, it certainly has merit and could well be used in practice.

In more general terms, a first main conclusion of this chapter was that it appears reasonable to believe that, for most scoring systems used in practice, *an "upward nested" countback system is preferable to a "downward nested" one.* Thus, in tennis for example, the upward ordering of points - games - sets - matches is undoubtedly preferable to the downward ordering of matches - sets - games - points when carrying out the countback. A second conclusion was that the P_n system has some merit in this team situation. These broad conclusions would appear to be true for teams of $2, 3, 4, \ldots$ players and for the bipoints cases as well as the unipoints cases.

CHAPTER 12
SCORING IN MULTIPLE CHOICE EXAMINATIONS

12.1 Introduction

An important requirement of sports scoring systems is that the weaker player should have a relatively small probability of winning. This is equivalent to requiring the better player's score (as defined by the scoring system) to have a relatively high probability of being greater than the poorer player's score at the completion of the match. Unlike sports scoring systems in which only two persons or two teams are usually involved, examinations typically involve many candidates. An important requirement of an examination is that the ranking of the total scores obtained by the candidates should correspond as closely as possible to the rankings of their respective abilities being measured. This requirement can thus be seen as an extension of the above requirement for sports scoring systems. An unfortunate characteristic of the multiple choice method of examination is that the instructions given to the candidates and the method of scoring typically used are such that an expected increase in score can often be obtained by sensible guessing. In this chapter we consider the case in which each question has a alternative answers, only one of which is correct.

Classically, the examinee is asked to place a tick in the box relating to the correct answer and leave the other boxes blank. $(a-1)$ marks are awarded for a correct tick and -1 for an incorrect one. No marks are awarded for unanswered questions. This scoring system is called Method A. Under Method A, an examinee who does not know the correct answer, but does know one or more boxes relating to incorrect answers, will in many cases guess (albeit intelligently), as, by doing so, he will have an expected increase in his total score. This guessing process introduces into his total score a random element which, of course, has nothing to do with his actual knowledge. This scoring system with $a = 5$ was used in the 1983 Australian Mathematics Competition, with about half of the 275,457 examinees answering all 30 questions and the remainder answering almost all questions. On average about 11 questions were correct. Thus, the experience with this competition suggests that the above assumption concerning guessing at random (when the correct answer is not known) is relevant.

Section 12.2, which is part of a published paper (Pollard (1985)), compares Method A with an alternative set of instructions and an alternative scoring system (Method B) which operates as follows. The examinee is asked to place one or more crosses ($\times$) in the box or boxes known to relate to incorrect answers and leave blank the correct box. If the box relating to the correct answer has been crossed, then $1-a$ marks are awarded for that question. On the other hand, if the box relating to the correct answer has not been crossed, one mark is awarded for each correctly positioned $\times$.

Section 12.3 gives a further method which is a modification of Method B, and one which, it is argued, removes random guessing completely.

12.2 A Comparison of Two Methods of Scoring

Under Method A, the examinee who does not know which box is correct, but does correctly know one or more boxes to contain incorrect answers, has an expected gain in marks by randomly placing a tick in one of the remaining boxes. Thus, under Method A, it can be argued that an examinee who knows i boxes ($i=(0),1,\ldots,a-1$) relating to incorrect answers will attempt to guess the correct response, thus incorporating a random element in his score.

Under Method B, it can be shown that an examinee who correctly knows that 1 or 2 or $\ldots a-1$ boxes represent incorrect answers would have a negative expected return from random guessing (and so would not guess). Thus, he will place an $\times$ in 1 or 2 or $\ldots a-1$ boxes respectively and accordingly obtain a mark of 1 or 2 or $\ldots a-1$. An examinee who from his knowledge cannot fill in any boxes may, however, randomly place (only) one $\times$ since (it can be shown that) this action has an expected score of 0 (and a variance of $a-1$). Thus, under Method B, the random element is removed or considerably reduced.

Consider the case in which $a=4$ and Method A scoring is used. An iterative procedure can be applied. Suppose the examinee knows i boxes (crosses) relating to incorrect answers over n questions ($i=0,1,2,\ldots,3n$). Assuming the crosses are randomly spread over the n questions, the probability, $P(j)$, that the examinee knows j crosses ($j=0,1,\ldots,\min(3,i)$) in the nth question (and $i-j$ over the remaining $n-1$ questions) is

$$P(j)=\frac{\binom{3n-3}{i-j}\binom{3}{j}}{\binom{3n}{i}}\,,\quad j=0,1,\ldots,\min(i,3)\,.$$

Suppose $D(i,k,n)$ is the probability an examinee obtains a total score of k ($k=-n,-n+4,-n+8,\ldots,3n$) (using Method A) on the n questions, given that he correctly knows i crosses over the n questions Then

$$D(i,k,n)=\sum_{j=0}^{\min(i,3)}P(j)\left(\frac{1}{4-j}D(i-j,k-3,n-1)+\frac{3-j}{4-j}D(i-j,k+1,n-1)\right)\,.$$

Using this equation, the distribution of an examinee's total score under Method A, given that he knows i crosses over the n questions, can be determined. A general equation can correspondingly be written down for Method B scoring. Table 12.2.1 gives the expected value and variance of the total score for various values of i when $n = 20$ questions. The linearity of an examinee's expected total score in his knowledge under Method B scoring and the considerably reduced variance (under Method B) of the score are noted.

Table 12.2.1
The expected value and variance of the examinee's total score for various values of i, the number of crosses known ($n = 20$, $a = 4$).

	Method A		Method B	
i	Mean	Variance	Mean	Variance
0	0.000	60.00	0	60.00
6	2,181	63.31	6	43.49
12	4.874	66.32	12	30.33
18	8.206	68.53	18	20.13
24	12.302	69.26	24	12.52
27	14.676	68.80	27	9.57
30	17.288	67.66	30	7.12
33	20.155	65.72	33	5.13
36	23.292	62.86	36	3.55
42	30.438	54.06	42	1.43
48	38.854	40.65	48	0.39
54	48.666	22.48	54	0.04
60	60.000	0.00	60	0.00

Computer programs were written for general n and the nine cases, $n = 10(10)30$ and $a = 3(1)5$, produced similar conclusions to those based on Table 12.2.1. For example, when $n = 10, 20$ or 30 questions and $a = 5$, the examinee, whose expected total score is half the maximum score (under Method B), has a variance under Method B scoring which is only about 5% of that under Method

A scoring. Thus, the advantages of Method B apply to those situations most common in practice.

The effect of slightly modifying Method A to consider the examinee who guesses only on those questions in which the expected gain is *strictly* positive was typically small.

It was noted from the distribution of scores that Method A has a considerably higher probability of passing an examinee who should fail and failing an examinee who should pass than does Method B.

12.3 Scoring to Remove Guessing

In this section a modification of Method B is analysed and it is shown that scoring systems can be devised in which random guessing can, under reasonable assumptions, be completely removed. We will consider the case in which $a = 4$. Under this modification, the examinee is asked to tick the box relating to the correct answer. If the correct answer is not known, the examinee is asked to leave the question blank or place a $\times$ in any box known to contain an incorrect answer. Each question must therefore be answered with a blank or a $\surd$ or one, two (or three) $\times$'s. The examinee is informed that, under the scoring system used, the expected return from randomly guessing (more than what is known) is negative under all circumstances.

An appropriate set of rewards and penalties for correct and incorrect responses is now devised. In this analysis it is assumed, without loss of generality, that box 4 contains the correct answer. A penalty, $-P_T$, is given for a question in which a $\surd$ is placed in box 1, 2 or 3 and 3 marks are given for a question in which the $\surd$ is placed in box 4. If a $\times$ has been placed in box 4, a penalty, $-P_C$, is given for that question. If box i is the only box to contain a $\times$, b_i marks are awarded for that question ($i = 1, 2, 3$). Alternatively, if boxes i and j are the only boxes containing $\times$'s , $b_i + b_j + b_{ij}$ marks are awarded ($i, j = 1, 2, 3;\ i \neq j$). Finally, if all boxes 1, 2 and 3 contain $\times$'s 3 (or $(b_1 + b_2 + b_3 + b_{12} + b_{13} + b_{23} + b_{123})$) marks are awarded.

12.3.1 The analysis. The expected marks obtained by examinees with different levels of partial knowledge on a single question are now examined.

(i) Consider firstly an examinee who knows nothing at all on a particular question. If this examinee is going to place one $\times$ at random in one of the 4 boxes, the expected marks obtained by such an action is $(b_1 + b_2 + b_3 - P_C)/4$. Alternatively, if this examinee is going to place two $\times$'s at random in two of the boxes, the expected marks obtained would now be

$$((b_1 + b_2 + b_{12}) + (b_1 + b_3 + b_{13}) + (b_2 + b_3 + b_{23}) - 3P_C)/6\,.$$

Also, if this examinee is going to place three $\times$'s at random, the expected marks obtained would be $(3 - 3P_C)/4$. Finally, if this examinee is going to place a $\surd$ at random in one of the boxes, the expected marks would be $(3 - 3P_T)/4$. All of the above expected values should be less than zero in order to discourage such guessing.

(ii) Now consider an examinee who correctly knows (say) box 1 to contain an incorrect answer and knows nothing else. If this examinee's plan is to place a $\times$ in box 1 and one other box selected at random, the expected mark obtained by such an action is $((b_1 + b_2 + b_{12}) + (b_1 + b_3 + b_{13}) - P_C)/3$. Alternatively, if this examinee intends to place a $\times$ in box 1 and two other boxes selected at random, the expected mark would now be $(3 - 2P_C)/3$. Finally, if this examinee is going to place a $\surd$ at random in one of boxes 2, 3 or 4, the expected mark would be $(3 - 2P_T)/3$. All of these expected values should be less than b_1 in order to discourage guessing.

(iii) A corresponding analysis to (ii) can be carried out for the case in which an examinee correctly knows box i to contain an incorrect answer and knows nothing else ($i = 2, 3$).

(iv) Now consider an examinee who correctly knows (say) box 1 and box 2 both to contain incorrect answers but knows nothing else. If this examinee's plan is to place $\times$'s in boxes 1 and 2 and one other box selected at random, the expected mark obtained by such an action would be $(3 - P_C)/2$. Alternatively, if this examinee is going to place a $\surd$ at random in one of boxes 3 or 4, the expected marks would now be $(3 - P_T)/2$. Both of these expected values should be less than $(b_1 + b_2 + b_{12})$ in order to discourage guessing.

(v) By an analogous argument, each of the expected marks in (iv) should also be less than both $(b_1 + b_3 + b_{13})$ and $(b_2 + b_3 + b_{23})$.

The above considerations suggest that the rewards $b_1, b_2, b_3, b_{12}, b_{13}, b_{23}, b_{123}$ and penalties P_C, P_T should satisfy the following inequalities.

(*i*)

$$b_1 + b_2 + b_3 < P_C$$

$$2(b_1 + b_2 + b_3) + b_{12} + b_{13} + b_{23} < 3P_C$$

$$1 < P_C$$
$$1 < P_T$$

(*ii*)
$$b_2 + b_3 + b_{12} + b_{13} - P_C < b_1$$
$$3 - 2P_C < 3b_1$$
$$3 - 2P_T < 3b_1$$

(*iii*)
$$b_1 + b_3 + b_{12} + b_{23} - P_C < b_2$$
$$3 - 2P_C < 3b_2$$
$$3 - 2P_T < 3b_2$$

$$b_1 + b_2 + b_{13} + b_{23} - P_C < b_3$$
$$3 - 2P_C < 3b_3$$
$$3 - 2P_T < 3b_3$$

(*iv*)
$$3 - P_C < 2(b_1 + b_2 + b_{12})$$
$$3 - P_T < 2(b_1 + b_2 + b_{12})$$

(*v*)
$$3 - P_C < 2(b_1 + b_3 + b_{13})$$
$$3 - P_T < 2(b_1 + b_3 + b_{13})$$

$$3 - P_C < 2(b_2 + b_3 + b_{23})$$
$$3 - P_T < 2(b_2 + b_3 + b_{23})\ .$$

There are many rewards and penalties which satisfy the above inequalities. For example,

(*a*) $b_1 = b_2 = b_3 = b_{12} = b_{13} = b_{23} = \frac{1}{3}$; $b_{123} = 1$; $P_C = 1.1$; $P_T = 1.5$;

(*b*) $b_1 = b_2 = b_3 = b_{12} = b_{13} = b_{23} = \frac{1}{2}$; $b_{123} = 0$; $P_C = 1.6$; $P_T = 1.5$;

(*c*) $b_1 = b_2 = b_3 = \frac{2}{3}$; $b_{12} = b_{13} = b_{23} = \frac{1}{3}$; $b_{123} = 0$; $P_C = 2.1$; $P_T = 1.5$;

(*d*) $b_1 = b_2 = b_3 = 0.9$; $b_{12} = b_{13} = b_{23} = 0.1$; $b_{123} = 0$; $P_C = 2.8$; $P_T = 1.5$.

The range of possibilities for the rewards and penalties can be seen to be quite wide (and,in fact, technically need not be the same for each question). It can be seen that P_C must increase as the b_i's increase. Also, it seems unwise to make P_C any larger than necessary as this would tend to deter the more timid examinee who fears the penalty for an incorrect $\times$. It can be seen that the examiner should not allow P_T to be too small or the weaker but less timid examinee will tend to use the $\surd$ option in the hope of obtaining a lucky pass. It seems reasonable to believe that P_T should be somewhere in the range 1.5–2.0 in order to strongly discourage the use of the $\surd$ option on questions in which the (weaker) examinee knows absolutely nothing.

Option (b) above is recommended with the modification that P_T be somewhere in the range 1.5–2.0. Note that, under (b), two $\times$'s in a question are worth three times the value of one $\times$. Similarly, three $\times$'s in a question are worth twice as much as two $\times$'s. These rewards relate well to the belief that the first $\times$ in a question is "easier" in some sense than the second which in turn is "easier" than the third. It is noted that one correctly positioned $\times$ in a question is, as it should be, not particularly well rewarded. Also, when only two $\times$'s have been correctly positioned, "half-marks" are awarded for the question, which seems appropriate as the question has been "narrowed down" to one out of two alternatives.

It may be possible to allocate values to the b_i's on an *a posteriori* basis with crosses or boxes which turned out to be "easier" worth less marks. Such an approach, however, has not been investigated and is not favoured as it seems reasonable that the examinee should be told the sizes of the rewards and penalties in advance.

12.3.2 Conclusions. It is possible to create multiple choice examinations in which it is argued random guessing will not occur due to the reward and penalty structure of the scoring system used. The examinee is instructed to tick ($\surd$) the box relating to the correct response or, if this is unknown, cross ($\times$) any boxes relating to responses known to be incorrect. Such actions are rewarded. If, however, a $\surd$ is placed in a box relating to an incorrect response or a $\times$ placed in a box relating to the correct response, penalties are involved. Many different reward and penalty structures are possible in which random guessing will only have a detrimental effect on the expected score. For the case of 4 alternatives, one such structure and the one recommended is: 3 marks are awarded for a correctly positioned $\surd$ (and no other entries) and $-P_T$ if the question contains an incorrectly positioned $\surd$ $(1.5 < P_T < 2.0)$. No marks are awarded for a blank. A penalty of 1.6 marks is incurred if the question contains an incorrectly positioned $\times$. 0.5, 1.5 or 3 marks are awarded for 1, 2 or 3 correctly positioned $\times$'s respectively within a question. This particular scoring system has several advantages over Method

B. Firstly, it is argued that random guessing is removed rather than considerably reduced. Secondly, rewards for $\times$'s need not be equal. This is a favourable property as the second and third crosses within a question are often seen as increasingly more "difficult" and worthy of greater rewards. Thirdly, a tick ($\surd$) is allowed as a response. This can be seen as properly encouraging the examinee to think "in a positive way" rather than "in several negative ways" about the question. Also, a $\surd$ is easier than the more cumbersome three crosses, $\times \times \times$. Fourthly, the penalty for an incorrectly positioned $\times$ has been reduced from 3 to 1.6, thus giving some encouragement to the more timid examinee who fears the penalty for such an error. One further aspect of the recommendation should be noted. The penalty for an incorrectly positioned $\surd$ should, I believe, be larger than that used in the classical scoring system (-1) and also larger than the minimum value required by the inequalities within this section. The effect of a larger value for this penalty (1.5 to 2 is recommended (perhaps (say) 1.6 to agree with the case of 3 $\times$'s one of which is incorrectly placed)) is to frustrate the (very) weak but less timid examinee who realizes that the hope of passing rests with lucky guessing using the $\surd$ option.

BIBLIOGRAPHY

ANDERSON, C.L. (1977) Note on the advantage of first serve. *Journal of Combinatorial Theory, Series A,* **23**, p.363.

APSIMON, H.G. (1951) The luck of the toss in squash rackets. *The Mathematical Gazette,* **35**, 193–194.

APSIMON, H.G. (1957) Squash chances. *The Mathematical Gazette,* **41**, 136–137.

BELLMAN, R.E. (1957) *Dynamic Programming.* Princeton: Princeton University Press.

CARTER, W.H. and CREWS, S.L. (1974) An analysis of the game of tennis. *The American Statistician,* **28(4)**, 130–134.

CLARKE, S. (1979) Tie point strategy in American and international squash and badminton. *Research Quarterly,* **50(4)**, 729–734.

CLARKE, S. and NORMAN, J. (1978) What chance playing up to ten? *The Squash Player International,* **7**, 50–51.

CLARKE, S. and NORMAN, J. (1979) Comparison of North American and international squash scoring systems — analytical results. *Research Quarterly,* **50(4)**, 723–728.

COX, D.R. and MILLER, H.D. (1965) *The Theory of Stochastic Processes.* London: Chapman and Hall.

CROUCHER, J.S. (1981) An analysis of the first 100 years of Wimbledon tennis finals. *Teaching Statistics,* **3(3)**, 72–75.

DIBLEY, C.S. and WALLIS, W.D. (1981) The effect of starting position in jai-alai. *Congressus Numerantium,* Vol **32**, Winnipeg, Canada.

FELLER, W. (1957) *An Introduction to Probability Theory and its Applications,* Vol.**1**, 2nd ed. New York: Wiley.

FISCHER, G. (1980) Exercise in probability and statistics, or the probability of winning at tennis. *Am. J. Phys.,* **48(1)**, 14–19.

GALE, D. (1971) Optimal strategy for serving in tennis. *Mathematics Magazine,* **44**, 197–199.

GEORGE, S.L. (1973) Optimal strategy in tennis: a simple probabilistic model. *Applied Statistics,* **22**, 97–104.

GITTINS, J.C. (1979) Bandit processes and dynamic allocation indices. *J. R. Statist. Soc. B.,* **41(2)**, 148–177.

HANNAN, E.L. (1976) An analysis of different serving strategies in tennis. From *Management Science in Sports,* R.E. Machol, S.P. Ladany and D.G.

Morrison, eds., Amsterdam: North-Holland Publishing Company.

HSI, B.P. and BURYCH, D.M. (1971) Games of two players. *Applied Statistics,* **20**, 86–92.

KEMENY, J.G. and SNELL, J.L. (1960) *Finite Markov Chains.* Princeton, New Jersey: D. Van Nostrand.

KING, H.A. and BAKER, J.A.W. (1979) Statistical analysis of service and match-play strategies in tennis. *Canadian J. App. Sports Science,* **4(4)**, 298–301.

KINGSTON, J.G. (1976) Comparison of scoring systems in two-sided competitions. *Journal of Combinatorial Theory, Series A,* **20**, 357–362.

MILES, R.E. (1980) Inefficiency in the scoring system of tennis. Fifth Australian Statistical Conference, Sydney, August, 1980.

MILES, R.E. (1982) Why tennis scoring should be changed. Tennis — Australia - Asia - the Pacific, April 1982, 54–55.

MILES, R.E. (1983) Sports scoring systems from a sequential analysis viewpoint. *Bull. I.S.I.,* **44**, Madrid, Sept. 1983.

MILES, R.E. (1984) Symmetric sequential analysis: the efficiencies of sports scoring systems (with particular reference to those of tennis). *J. R. Statist. Soc. B,* **46(1)**, 93–108.

MILES, R.E. (to appear) Symmetric sequential analysis: a theory of sports scoring systems.

MORAN, P.A.P. (1968) *An Introduction to Probability Theory.* Oxford: Clarendon Press.

MORRIS, C. (1972) The statistics of tennis. *Tennis West,* Nov., 10–11.

MORRIS, C. (1973) The statistics of tennis. *Tennis West,* March, 12–13.

MORRIS, C. (1977) The most important points in tennis. In: *Optimal Strategies in Sports,* (S.P. Ladany and R.E. Machol, eds.) pp. 131–140. Amsterdam: North-Holland. (Vol. 5 in *Studies in Management Science and Systems.*)

PHILLIPS, M.J. (1978) Sums of random variables having the modified geometric distribution with application to two-person games. *Adv. Appl. Prob.,* **10**, 647–665.

POLLARD, G.H. (1980) A statistical investigation of squash and tennis. M.Sc. Thesis, Australian National University, Canberra.

POLLARD, G.H. (1983a) An analysis of classical and tie-breaker tennis. *Austral. J. Statist.,* **25(3)**, 496–505.

POLLARD, G.H. (1983b) An analysis of the international system of scoring in squash. (Unpublished paper, enclosed.)

POLLARD, G.H. (1985a) A statistical investigation of squash. *Research Quarterly for Exercise and Sport,* **56(2)**, 144–150.

POLLARD, G.H. (1985b) Scoring in multiple choice examinations. *Math. Scientist,* **10**, 93–97.

POLLARD, G.H. (1987, to appear) A new tennis scoring system. *Research Quarterly for Exercise and Sport.*

REDINGTON, F. (1972) Usurpers. *Journal of the Institute of Actuaries Students Society,* **20(3)**, 353–354.

RENICK, J. (1976) Optimal strategy at decision points in singles squash. *Research Quarterly,* **47**, 562–568.

RENICK, J. (1977) Tie point strategy in badminton and international squash. *Research Quarterly,* **48**, 492–498.

ROHATGI, V.K. (1976) *An Introduction to Probability Theory and Mathematical Statistics.* New York: Wiley.

SCHUTZ, R.W. (1970) A mathematical model for evaluating scoring systems with specific reference to tennis. *Research Quarterly,* **41**, 552–561.

SCHUTZ R.W. and KINSEY, W. (1977) Comparison of North American and international squash scoring systems — a computer simulation. *Research Quarterly,* **48**, 248–251.

SIMON, R. (1977) Adaptive treatment assignment methods and clinical trials. *Biometrics,* **33**, 743–749.

SOBEL, M. and WEISS, G.H. (1970) Play-the-winner sampling for selecting the better of two binomial populations. *Biometrika,* **57(2)**, 357–365.

UPPULURI, V.R.R. and BLOT, W.J. (1974) Asymptotic properties of the number of replications of a paired comparison. *J. Appl. Prob.,* **11**, 43–52.

WALD, A. (1947) *Sequential Analysis.* New York: Wiley.

WALD, A. and WOLFOWITZ, J. (1948) Optimum character of the sequential probability ratio test. *Ann. Math. Statist.,* **19**, 326–339.

WATSON, D. (1970) On scoring in games. *The Mathematical Gazette,* **54**, 110–113.

ZELEN, M. (1969) Play-the-winner rule and the controlled clinical trial. *J. Am. Statist. Ass.,* **64**, 131–146.

POLLARD, G. H. (1983b, unpublished paper)

AN ANALYSIS OF THE INTERNATIONAL SYSTEM OF SCORING IN SQUASH

G. H. POLLARD

SCHOOL OF INFORMATION SCIENCES

CANBERRA COLLEGE OF ADVANCED EDUCATION

SUMMARY

Explicit expressions are found for the distribution of the score and the number of rallies in a game of squash played under international rules.

1. INTRODUCTION

The first analysis of international squash by Ap Simon (1951) evaluated the effect of winning the toss. Ap Simon (1957) attempted to find general expressions for the probability that a player wins a game. Watson (1970) used difference equations to evaluate numerically the probability that a player wins a game. Renick (1978) and Clarke and Norman (1978) analysed the relative merits of calling for a "long" or "short" game. Clarke and Norman (1979) used recurrence relationships to evaluate the expected value and variance of the number of rallies in a game of squash. Schutz and Kinsey (1978) made a comparison of international and North American scoring systems using simulation methods. Pollard (1980) used random walk procedures to evaluate the distribution of the score in a game and to evaluate the relative "importances" of each of the states within a game. This paper extends the results of Clarke and Norman (1979) and Pollard (1980).

In this paper it is assumed that player A serves first in the game and has a constant probability $p(q)$ of winning (losing) a rally. It is assumed that rallies are independent and that a "short" game (first to 9) is called for if the score 8-8 is reached. Explicit results for the distribution of the scores and the number of rallies in a game are obtained.

A. The Case in which Player A wins the Game 9-i (i = 0,1,2,....8).

As an example, consider the case in which player A wins the game 9-4. These 4 points that player A lost may have been lost in the following ways:

(a) 4 at once, denoted by (4);

(b) 3 at once, and the others separately denoted by (3,1) or (1,3);

(c) (2,2)

(d) (2,1,1), (1,2,1), (1,1,2);

(e) (1,1,1,1).

Note that, in (a) for example, player A may have won one or more rallies or handouts (but no points) during this period in which he lost 4 consecutive points. The above 5 cases are now considered separately.

Case (a)

Figure 1 gives an example of such a situation. In this example, the 4 points lost (L) consecutively occur when the score is 3-0. In this figure the handout losses (L) and wins (W) giving rise to a change in score at the next rally are underlined.

[Insert Figure 1 somewhere here.]

The downward arrows indicate the 9 positions in which the sequence L L L L L W could have occurred. Note that the first loss in this sequence, L, gives service to player B and the last win in this sequence W, returns service to player A. The 13 dots in Figure 1 indicate the positions in which handout cycles can occur. Handout cycles are rallies whose outcomes do not lead to a change in score. Such handout cycles are LW when preceding a W and WL when preceding an L as the upward arrows in Figure 1 demonstrate. Note that a dot in this figure cannot occur before an underlined L or W. Figure 1 thus shows the realization of a game in which there was a handout cycle when the score was 1-0, and double handout cycles at 3-2 and 6-4. In this realization, there were 13 points scored and 25 rallies played.

Making use of the occupancy problem in Feller (1957. p.36) to distribute all possible unscoring handout cycles, the probability that player A wins a game 9-4, losing all 4 points consecutively, is

$$\binom{9}{1}p^{10}q^{5}\sum_{n=0}^{\infty}\binom{12+n}{n}(pq)^{n}$$

Figure 2 shows a realization of the second type within case (b) in ... player A lost a point at 3-0 and 3 more consecutive points at 5-1 ... unscoring handout cycle is

[Insert Figure 2 somewhere here.]

shown at 2-0, and unscoring handout triple cycles are shown at 5-2 ... 7-4. Thus in this realization, there were 13 points scored and 31 ... played. Similarly to above, the probability that player A wins a ... 9-4, losing 3 points in a row and the other separately, is

$$2\binom{9}{2}p^{11}q^{6}\sum_{n=0}^{\infty}\binom{12+n}{n}(pq)^{n}$$

Case (c), (d) and (e)

The corresponding expressions for cases (c), (d) and (e) are:

Case (c) $\binom{9}{2}p^{11}q^{6}\sum_{n=0}^{\infty}\binom{12+n}{n}(pq)^{n}$

Case (d) $3\binom{9}{3}p^{12}q^{7}\sum_{n=0}^{\infty}\binom{12+n}{n}(pq)^{n}$

Case (e) $\binom{9}{4}p^{13}q^{8}\sum_{n=0}^{\infty}\binom{12+n}{n}(pq)^{n}$

The sum of these 5 expressions gives the probability that player A ... the game 9-4. Thus, combining cases (b) and (c) above, the probabil... that player A wins the game 9-4 is

$$\left\{\binom{9}{1}p^{10}q^{5}+3\binom{9}{2}p^{11}q^{6}+3\binom{9}{3}p^{12}q^{7}+\binom{9}{4}p^{13}q^{8}\right\}\sum_{n=0}^{\infty}\binom{12+n}{n}(pq)^{n}$$

Also, as an example, the joint probability that player A wins 9-4 ... the duration, D, of the game is 21 rallies (ie. there is 21-13=8 han... in total) is

$$\left\{\binom{9}{1}t_{3}+3\binom{9}{2}t_{2}+3\binom{9}{3}t_{1}+\binom{9}{4}t_{0}\right\}p^{13}q^{8}$$

Where $t_i = \binom{12+n}{\ldots}$ ... for i = 0, 1, 2 and 3

Corresponding expressions can be written down for the joint probability when $D = 15,17,19,21,.....$. Note that, in this case, even durations are not possible. Thus, the conditional distribution, the mean and the variance of the duration of a game given player A won (say) 9-4 can be evaluated (see Table 1 for the case in which $p = 0.5$).

[Insert Table 1 somewhere here.]

The cases in which player A wins $9-i$ ($i=1,2,.....,8$) can be correspondingly analysed. For $i=1,2,.....,8$, the probability that player A wins $9-i$ is

$$\left\{\sum_{j=0}^{i-1}\binom{i-1}{j}\binom{9}{j+1}p^{10+j}q^{i+1+j}\right\}\sum_{n=0}^{\infty}\binom{8+i+n}{n}(pq)^n$$

If this expression is denoted by $\sum_{k=0}^{\infty} W_{i,k}\, p^{10+k}\, q^{i+1+k}$,

then $W_{i,k}$ denotes the joint probability that player A wins $9-i$ and the duration of the game is $(11+i+2k)$ rallies.

The case in which player A wins 9-0 can be correspondingly examined. Combining these 9-0 results with those above, the distribution of the duration of a game conditional on player A winning that game can be determined. Tables 2 and 3 give further characteristics for the case in which $p = 0.5$, of a game which player A wins. The first column of Table 2 agrees with Pollard (1980,p22) who used random walk procedures to generate the probabilities.

[Insert Tables 2 and 3 somewhere here.]

B. The Case in which Player A loses the Game I-9 (I = 0,1,2,.....8).

As an example, consider the case in which player A loses the game 2-9. The 2 points that player A won may have been won in the following ways:

(a) Player A wins the first 2 rallies and no more points.

(b) Player A wins the first rally, and 1 point later on.

(c) Player A loses the first rally, and wins 2 points later on.

[Include Figure 3 somewhere here]

Figure 3 gives a representation, omitting any unscoring cyclical handouts (for simplicity), of examples of these 3 cases.

Note that case (c) in turn consists of 2 types: the 2 points won later on are either (i) won separately

or (ii) won together.

Now including all possible unscoring cyclical handouts, the probability associated with each of the above cases is:

Case (a) $\quad p^2q^{10}\sum_{n=0}^{\infty}\binom{8+n}{n}(pq)^n$

Case (b) $\quad \binom{9}{1}p^3q^{11}\sum_{n=0}^{\infty}\binom{9+n}{n}(pq)^n$

Case (c) (i) $\quad \binom{9}{2}p^4q^{12}\sum_{n=0}^{\infty}\binom{10+n}{n}(pq)^n$

(ii) $\quad \binom{9}{1}p^3q^{11}\sum_{n=0}^{\infty}\binom{10+n}{n}(pq)^n$

Thus, placing case (a) last in the following expression, the probability that player A loses 2-9 is

$$\left\{\left\{\binom{9}{2}p^4q^{12}+\binom{9}{1}p^3q^{11}\right\}S_{11}+\binom{9}{1}p^3q^{11}S_{10}\right\}+p^2q^{10}S_9$$

where $S_m = \sum_{n=0}^{\infty}\binom{m+n-1}{n}(pq)^n$.

Also, as an example, the joint probability that player A loses 2-9 and the duration, D, of the game is 16 rallies (i.e. there are 5 handouts in total) is

$$\left\{\left[\binom{9}{2}t_{11,0} + \binom{9}{1}t_{11,1}\right] + \binom{9}{1}t_{10,1} + t_{9,2}\right\} p^{4}q^{12}$$

where $t_{m,n} = \binom{m+n-1}{n}$

Corresponding expressions can be written down for the joint probability when $D = 12,14,16,18,\ldots\ldots$. Note that odd durations are not possible in this case. Thus, similar to Table 1, the conditional distribution, the mean and the variance of the duration of a game given player A lost the game 2-9 can be evaluated.

The cases in which player A loses i-9 ($i = 1,2,\ldots..8$) can be correspondingly analysed. For $i = 1,2,\ldots..8$, the probability that player A loses i-9 is

$$\sum_{k=1}^{i}\sum_{j=0}^{k-1}\binom{k-1}{j}\binom{9}{j+1}q^{11+j}p^{i+1+j}\,S_{9+k} + q^{10}p^{i}\,S_{9}$$

If this expression is denoted by $\sum_{k=0}^{\infty} L_{i,k}\, q^{10+k}p^{i+k}$, then $L_{i,k}$ denotes the joint probability that player A loses i-9 and the duration of the game is $(10 + i + 2k)$ rallies.

The case in which player A loses 0-9 can be correspondingly examined. Combining these 0-9 results with those above, the distribution of the duration of a game conditional on player A losing that game can be determined (in a similar way to Table 3). By combining the winning and losing distributions of this section, the overall distribution of the duration of a game of squash can be determined (see Table 4). Table 5 gives further characteristics of the duration of a game. These results are in agreement with Pollard (1980, p.15) and Clarke and Norman (1979, Table 1)

[Insert Tables 4 and 5 somewhere here]

Squash is typically played as a best-of-five games match. The methods of Pollard (1983) for analysing a best-of-three sets match of tennis can be applied to the results of this section to give distributional, mean and variance results for a best-of-five games match of squash.

Corresponding results to those obtained in this section for the international method of scoring can be obtained for the North American method of scoring by truncating the negative binomial distribution. Also, results for the "long" game case can be obtained using the methods of this section.

REFERENCES

Ap Simon, H. (1951). *The Luck of the toss in squash rackets.* Mathematical Gazette 35: 193-194.

Ap Simon, H. (1957). *Squash Chances.* Mathematical Gazette 41: 136-7.

Clarke, S. R., Norman, J. M., (1978). *Playing up to ten: what are the chances?* The Squash Player. 7.7: 50-51, July 1978.

Clarke, S. R., Norman, J. M. (1979). *Comparison of North American and International Squash scoring systems - analytical results.* Research Quarterly, 50, 723-728.

Feller, W. (1957). *An Introduction to Probability Theory and its Applications.* 2nd Edition, Vol. 1, Wiley.

Pollard, G. H. (1980). *A Statistical investigation of squash and tennis.* M. Sc. thesis, Australian National University, Canberra.

Pollard, G. H. (1983). *An Analysis of Classical and Tie-Breaker Tennis.* Austral. J. Statist., 25(3), 1983 to appear.

Renick, J. (1978). *Tie Point Strategy in Badminton and International Squash,* Research Quarterly, 48, 492-498.

Schutz, R. W., Kinsey, W. J. (1978). *Comparison of North American and International Squash Scoring Systems - A Computer Simulation.* Research Quarterly, 48, 248-251.

Watson, D. (1970). *On Scoring in Games,* Mathematical Gazette 54:110.

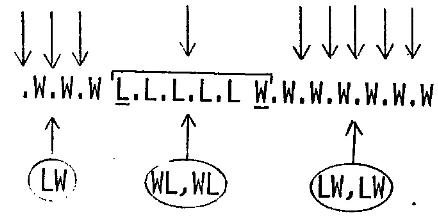


Figure 1

A diagrammatical representation of a game of duration 25 which player *A* won 9-4 losing the 4 points in a row.

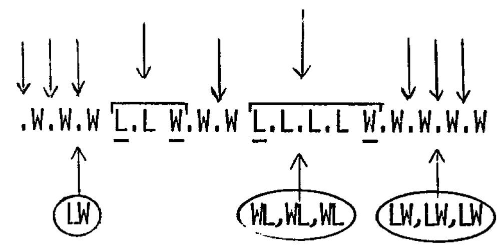


Figure 2

A diagrammatical representation of a game of duration 31 which player *A* won 9-4, losing 3 points in a row and the other separately.

W W L.L.L.L.L.L.L.L.L.L

) W L.L.L.L.L W.W L.L.L.L.L.L

(i) L.L.L.L W.W L.L.L.L W.W L.L.L.L

ii) L.L.L.L.L.L W.W.W L.L.L.L.L

Figure 3

Diagrammatical representations of a game (ignoring possible unscoring cyclical handouts) which player A lost 2-9

Table 1

Conditional Distribution of the Duration of a Game given Player A wins the Game 9-4 (for the case in which $p = 0.5$).

Duration	Probability
10-14	0.00
15-19	0.10
20-24	0.28
25-29	0.42
30-34	0.13
35-39	0.06
≥40	0.01
Total	1.00
mean	26.1
variance	25.4

Table 2

The Probability that Player A wins $9-i$, the Mean and the Variance of the Duration given Player A wins $9-i$ (for $i = 0,1,2,\ldots..8$ and $p = 0.5$).

i	Probability	Mean	Variance
0	0.026	15.0	16.0
1	0.039	18.7	17.8
2	0.052	21.3	20.6
3	0.062	23.8	23.0
4	0.069	26.1	25.4
5	0.073	28.3	27.6
6	0.073	30.4	29.8
7	0.072	32.5	31.9
8	0.069	34.5	33.9
Overall	0.535	27.1	56.2

Table 3

Distribution of the Duration of a Game given Player A wins the Game (for the case in which $p = 0.5$).

Duration	Probability
5 - 9	0.00
10 - 14	0.04
15 - 19	0.12
20 - 24	0.21
25 - 29	0.26
30 - 34	0.20
35 - 39	0.11
≥ 40	0.05
Total	1.00
Mean	27.1
Variance	56.2

Table 4

Distribution of the Duration of a Game (for the case in which $p = 0.5$).

Duration	Probability
10 - 14	0.03
15 - 19	0.11
20 - 24	0.21
25 - 29	0.26
30 - 34	0.21
35 - 39	0.11
$\geq$ 40	0.06
Total	1.00
Mean	27.5
Variance	54.5

Table 5

The Probability that Player *A* wins; the Mean and the Variance of the Duration of a Game (assuming Player *A* is the first player to serve in the game).

p value	Probability that Player *A* wins the Game	Expected Duration	Variance of Duration
0.5	0.535	27.5	54.5
0.52	0.627	27.2	55.4
0.54	0.713	26.7	56.3
0.56	0.788	25.9	56.7
0.58	0.850	25.0	56.2
0.60	0.899	23.9	54.5
$p = 0.70$	0.994	18.3	33.7