Parabola Volume 39, Issue 1 (2003)

Winning at Soccer

Peter Donovan[1]

The previous article by Milan Pahor deals with the scoring system in tennis and shows that it tends to magnify differences in skill. The reason is that the number of scoring occasions in a serious tennis match is quite large: there are 5 sets each of about 10 games with each game containing about 6 scoring shots. In a top-level soccer game there will be about 6 or 7 scoring shots at the outside and a rather different phenomenon occurs: there is a quite appreciable chance that the weaker team wins. Indeed, consider the following table:

	1.0	**1.2**	**1.4**	**1.6**	**1.8**	**2.0**	**2.2**	**2.4**	**2.6**	**2.8**	**3.0**
1.0	35	31	27	24	21	18	16	14	12	11	9
1.2	41	36	32	29	25	23	20	18	16	14	12
1.4	46	42	37	34	30	27	24	21	19	17	15
1.6	51	47	42	38	35	31	28	25	23	20	18
1.8	56	51	47	43	39	35	32	29	26	24	21
2.0	61	56	51	47	43	40	36	33	30	27	25
2.2	65	60	56	51	48	44	40	37	34	31	28
2.4	68	64	60	55	51	48	44	41	37	34	31
2.6	72	67	63	59	55	51	48	44	41	38	35
2.8	75	71	67	63	59	55	51	48	45	41	38
3.0	77	74	70	66	62	58	55	51	48	45	42

Here, under the mathematical simplification explained below, is given the percentage probability that team A will beat team B. The average score that may be expected for team A over a long series of matches is given in the bold face column on the left and the average score that may be expected for team B over the same long series of matches is given in the bold face row on the top. Thus if A may be expected to score 2.4 on average while B may be expected to score 1.8 then A has a 51% chance of winning. Likewise if A may be expected to score 1.8 on average while B may be expected to score 2.4 then A has a 29% chance of winning. The remaining 20% is the probability of a draw.

The appreciable chance that the weaker team wins in soccer can be reduced by requiring the teams to play 10 matches and deciding the winner on the total score. In this case the corresponding table is:

[1]Peter Donovan is a Senior Lecturer at UNSW

	10	12	14	16	18	20	22	24	26	28	30
10	46	30	18	10	5	3	1	1	0	0	0
12	63	46	31	20	12	6	3	2	1	0	0
14	76	62	46	32	21	13	8	4	2	1	1
16	86	75	61	46	33	23	14	9	5	3	2
18	92	84	73	60	47	34	24	16	10	6	3
20	96	91	83	72	60	47	35	25	17	11	7
22	98	95	90	81	71	59	47	36	26	18	12
24	99	97	94	88	80	70	59	47	36	27	19
26	100	99	97	93	87	79	69	58	47	37	27
28	100	99	98	96	92	86	78	69	58	47	37
30	100	100	99	98	95	91	85	77	68	58	47

Thus if A may be expected to score 24 over ten games and B may be expected to score 18 over ten games then in any ten game tournament between the two teams A has an 80% chance of winning.

The calculations that produced the above tables depend upon certain assumptions that are only approximately valid and some probability theory that is too technical to explain here. Yet something can be said in the available space:

The major mathematical approximation used in constructing the tables is that if team A scores a goals on average whenever it plays team B then the probability that it scores exactly m goals against B is[2]

$$p_m = e^{-a}\frac{a^m}{m!}. \tag{1}$$

The exponential constant $e = 2.718\ldots$ has the property that for a real a the exponential e^a can be expressed as the sum of the infinite series:

$$1 + \frac{a}{1!} + \frac{a^2}{2!} + \frac{a^3}{3!} + \ldots.$$

Once an appropriate decimal accuracy has been specified one may find an integer p so that the sum of the first p terms is a sufficiently accurate approximation to e^a. Thus there is the key formula:

$$1 = e^{-a} + e^{-a}\frac{a}{1!} + e^{-a}\frac{a^2}{2!} + e^{-a}\frac{a^3}{3!} + \ldots. \tag{2}$$

You may know that the Formula (2) now ensures that $\sum p_m = 1$.

As it is most unusual for two top-level soccer teams to play long series of matches with each other (1) can scarcely be verified directly. We would have to discuss the

[2]Editors Note: This formula for the probabilities is known as the Poisson Distribution. It is widely used in mathematics to estimate the probability of observing m discrete events if the average number of events is a. It is obtained from the Binomial Distribution by taking two limits; the number of possible events approaches infinity, the probability of a specific event approaches zero.

extent to which the general theory of random processes of Poisson type is applicable here. However this is much beyond the scope of this article. The analogous simplification is made for team B: the probability that it scores exactly n goals when it meets team A is

$$q_n = e^{-b} \frac{b^n}{n!}. \tag{3}$$

The second mathematical simplification in the tables is that the scores of the two teams are independent. As it is quite impossible for both teams to score at once this simplification is to some extent unsound. Yet as scoring occurs in at most a small proportion of the total time of the match, this simplification may be reasonable anyway. In view of independence assumption, the probability that A scores exactly m goals and B scores exactly n goals is $p_m q_n$. Thus the probability that A wins is

$$p_1 q_0 + p_2 q_0 + p_2 q_1 + p_3 q_0 + p_3 q_1 + p_3 q_2 + p_4 q_0 + \ldots.$$

This infinite series can be used without much difficulty to calculate the information in the tables.