Condensed Matter Physics 2009, Vol. 12, No 4, pp. 739–752

Football fever: self-affirmation model for goal distributions

W. Janke[1], E. Bittner[1], A. Nußbaumer[1], M. Weigel[2,3]

[1] Institut für Theoretische Physik and Centre for Theoretical Sciences (NTZ) – Universität Leipzig, Augustusplatz 10/11, Postfach 100920, 04009 Leipzig, Germany

[2] Theoretische Physik, Universität des Saarlandes, 66041 Saarbrücken, Germany

[3] Institut für Physik, Johannes Gutenberg-Universität Mainz, Staudinger Weg 7, 55099 Mainz, Germany

Received July 22, 2009

The outcome of football games, as well as matches of most other popular team sports, depends on a combination of the skills of players and coaches and a number of external factors which, due to their complex nature, are presumably best viewed as random. Such parameters include the unpredictabilities of playing the ball, the player's condition of the day or environmental influences such as the weather and the behavior of the audience. Under such circumstances, it appears worthwhile to analyze football score data with the toolbox of mathematical statistics in order to separate deterministic from stochastic effects and see what impact the cooperative and social nature of the "agents" of the system has on the resulting stochastic observables. Considering the probability distributions of scored goals for the home and away teams, it turns out that especially the tails of the distributions are *not* well described by the Poissonian or binomial model resulting from the assumption of uncorrelated random events. On the contrary, some more specific probability densities such as those discussed in the context of extreme-value statistics or the so-called negative binomial distribution fit these data rather well. There seemed to be no good argument to date, however, why the simplest Poissonian model fails and, instead, the latter distributions should be observed. To fill this gap, we introduced a number of microscopic models for the scoring behavior, resulting in a Bernoulli random process with a simple component of *self-affirmation*. These models allow us to represent the observed probability distributions surprisingly well, and the phenomenological distributions used earlier can be understood as special cases within this framework. We analyzed historical football score data from many leagues in Europe as well as from international tournaments, including data from all past tournaments of the "FIFA World Cup" series, and found the proposed models to be applicable in all cases. To complete the picture, we conducted a field study with visitors of a science showcase to collect additional data from matches of tabletop football. As it turns out, also the latter data are represented well with our feedback models, underscoring their apparently rather universal applicability.

Key words: *sport statistics, extreme-value distributions, feedback models*

PACS: *89.20.-a, 02.50.-r*

1. Introduction

Throughout Europe football undoubtedly is a social and economic phenomenon on a large scale with thousands of players organized in several levels of leagues in most countries and an audience in the arenas as well as in front of the television sets counting at least tens of millions. Naturally, an event of such dimensions has triggered a significant number of research studies, for instance concerned with the improvement of game tactics or the question of the predictability of results. Clearly less effort has been devoted, it seems, to the understanding of football (and other ball sports) from the perspective of the stochastic behavior of co-operative "agents" (i. e., players) in abstract models. From a statistical mechanics point-of-view, however, this problem appears to be an ideal testing ground for the application of its simplified stochastic models to subjects outside of condensed-matter physics [1,2]. A number of such attempts to study ball sports with a statistical mechanics machinery have been reported earlier; see, for instance, the examples collected in [3].

Score distributions of football and other ball games have been occasionally considered by mathematical statisticians for more than fifty years [4–10]. The first studies, involving only very limited statistical data, seemed to indicate that the resulting score distributions were compatible with the simplest assumption of a completely random, Poissonian process with a fixed, albeit possibly team-dependent, scoring probability [4]. Somewhat later, due to the observed deviations (especially) of the tails of the empirical distributions from the Poisson shape, the latter was, on purely empirical grounds, abandoned in favor of a negative binomial distribution (NBD) [11]. The NBD occurs naturally for a mixture of Poissonian processes with a certain distribution of (independent) success probabilities [5]. Much more recently, the availability of significantly more extensive collections of scoring data led the authors of [10] to the conclusion that, at least for some leagues, histograms of football scores resembled rather the generalized distributions of extreme value statistics [12] than any other distribution that had been considered so far. In total, these studies yielded a rather heterogeneous picture, offering statistical descriptions of the empirically observed goal distributions purely on the grounds of best fit, that is, without suggesting any microscopic justification for the manifestation of the assumed analytical forms of probability densities. Furthermore, for a system of highly co-operative entities it might be presumed that such models without correlations cannot be an adequate description anyway.

The distribution of extremes, i. e., the probability density function of (k^{th}) maximal or minimal values of independent realizations of a random variable, is described by only a few universality classes, depending on the asymptotic behavior of the original distribution [12]. Apart from the direct importance of the problem of extremes in actuarial mathematics and engineering, generalized extreme value (GEV) distributions have been found to occur in such diverse systems as the statistical mechanics of regular and disordered systems [13–19], turbulence [20] or earthquake data [21]. However, in most cases global properties were considered instead of explicit extremes, and the occurrence of GEV distributions led to speculations about hidden extreme processes in these systems, which could not be identified in most cases, though. It was only realized recently that GEV distributions can also arise naturally as the statistics of sums of *correlated* random variables [22–24], which could explain their ubiquity in physical systems.

For the problem of scoring in football, correlations naturally occur through processes of (positive and negative) feedback of scoring on both teams, and we shall see how the introduction of simple rules for the adaptation of the success probabilities in a modified Bernoulli process upon scoring a goal leads to systematic deviations from Gaussian statistics. We find simple models with a single parameter of *self-affirmation* to best describe the available data, including cases with relatively poor fits of the NBD. The latter is shown to result from one of these models in a particular limit, offering an explanation for the relatively good fits observed earlier. For the models under consideration, exact recurrence relations and precise closed-form approximations of the probability density functions can be derived. Although the limiting distributions of the considered models in general do *not* follow the statistics of extremes, it is demonstrated how alternative models leading to GEV distributions could be constructed. The best fits are found for the models where each extra goal encourages a team even more than the previous one: a true sign of *football fever*.

2. Probability distributions

Considering a simplified statistical description of a football match, the most natural quantity to start with is certainly given by the overall score of the game. Remaining on this level of a first approximation, we restrict ourselves here to the analysis of the distributions of goals scored by the home and away teams in football league or cup matches. Disregarding any effects of player skill, and thus degrading football to a pure game of chance, one might start out assuming independent and constant probabilities of each team for scoring during an appropriate time interval of the match. Since the scoring probabilities will be small, the resulting probabilities of final scores will follow a Poissonian distribution,

$$P^{\mathrm{h}}_{\lambda_h}(n_{\mathrm{h}}) = \frac{{\lambda_{\mathrm{h}}}^{n_{\mathrm{h}}}}{n_{\mathrm{h}}!}\exp(-\lambda_{\mathrm{h}}), \qquad P^{\mathrm{a}}_{\lambda_a}(n_{\mathrm{a}}) = \frac{{\lambda_{\mathrm{a}}}^{n_{\mathrm{a}}}}{n_{\mathrm{a}}!}\exp(-\lambda_{\mathrm{a}}), \tag{1}$$

where n_{h} and n_{a} are the final scores of the home and away teams, respectively, and the parameters λ_{h} and λ_{a} are related to the average number of goals scored by a team, $\lambda = \langle n \rangle$. As an additional check of the fit to the data, one might then also consider the probability densities of the sum $s = n_{\rm h} + n_{\rm a}$ and difference $d = n_{\rm h} - n_{\rm a}$ of goals scored under the assumption of Poissonian distributions for $n_{\rm h}$ and $n_{\rm a}$,

$$\begin{aligned} P^{\Sigma}_{\lambda_{\rm h},\lambda_{\rm a}}(s) &= \sum_{n=0}^{s} P^{\rm h}_{\lambda_{\rm h}}(n) P^{\rm a}_{\lambda_{\rm a}}(s-n) = \frac{(\lambda_{\rm h}+\lambda_{\rm a})^s}{s!} \exp[-(\lambda_{\rm h}+\lambda_{\rm a})], \\ P^{\Delta}_{\lambda_{\rm h},\lambda_{\rm a}}(d) &= \sum_{n=0}^{\infty} P^{\rm h}_{\lambda_{\rm h}}(n+d) P^{\rm a}_{\lambda_{\rm a}}(n) = {\rm e}^{-(\lambda_{\rm h}+\lambda_{\rm a})} \left(\frac{\lambda_{\rm h}}{\lambda_{\rm a}}\right)^{d/2} I_d(2\sqrt{\lambda_{\rm h}\lambda_{\rm a}}), \end{aligned} \tag{2}$$

where I_d is the modified Bessel function (see [25], p. 374). Note that $P^{\Sigma}_{\lambda_{\rm h},\lambda_{\rm a}}(s)$ is itself a Poissonian distribution with parameter $\lambda = \lambda_{\rm h} + \lambda_{\rm a}$.

Clearly, even the most pessimistic football fan would tend to allow for some effects of player behavior on the outcome of a match and hence consider the assumption of constant and independent scoring probabilities for the teams as not appropriate for real-world football matches. Since we are interested in averages over the matches during one or several seasons of a football league or cup, one might expect a *distribution* of scoring probabilities λ depending on the different skills of the teams, the lineup for the match, tactics, weather conditions etc., leading to the notion of a *compound* Poisson distribution. It can be easily shown [26,27] that for the special case of the scoring probabilities λ following a gamma distribution,

$$f(\lambda) = \begin{cases} \dfrac{a^r}{\Gamma(r)} \lambda^{r-1} {\rm e}^{-a\lambda}, & \lambda > 0, \\ 0, & \lambda \leqslant 0, \end{cases} \tag{3}$$

the resulting compound Poisson distribution has the form of a NBD [27],

$$P_{r,p}(n) = \int_0^\infty {\rm d}\lambda\, P_\lambda(n) f(\lambda) = \frac{\Gamma(r+n)}{n!\,\Gamma(r)} p^n (1-p)^r, \tag{4}$$

where $p = 1/(1+a)$. This distribution has been found to rather well describe football score data [6,10]. The underlying assumption of the scoring probabilities following a gamma distribution seems to be rather *ad hoc*, however, and fitting different seasons of our data with the Poissonian model (1), the resulting distribution of the parameters λ does not resemble the gamma form (3). At the level of discussion up to this point, the parameter r introduced in equation (3) only appears as an empirical fit parameter. As will be shown below, however, it corresponds to the ratio of initial scoring probability and "self-affirmation factor" in the context of one of the microscopic models considered here. Analogous to equation (2), for the NBD (4) one can evaluate the probabilities for the sum s and difference d of goals scored by the home and away teams,

$$\begin{aligned} P^{\Sigma}_{r_{\rm h},p_{\rm h},r_{\rm a},p_{\rm a}}(s) &= (1-p_{\rm h})^{r_{\rm h}} (1-p_{\rm a})^{r_{\rm a}} {p_{\rm a}}^s \frac{\Gamma(r_{\rm a}+s)}{s!\,\Gamma(r_{\rm a})}\, {}_2F_1\left(-s, r_{\rm h}; 1-s-r_{\rm a}; \frac{p_{\rm h}}{p_{\rm a}}\right), \\ P^{\Delta}_{r_{\rm h},p_{\rm h},r_{\rm a},p_{\rm a}}(d) &= (1-p_{\rm h})^{r_{\rm h}} (1-p_{\rm a})^{r_{\rm a}} {p_{\rm h}}^d \frac{\Gamma(r_{\rm h}+d)}{d!\,\Gamma(r_{\rm h})}\, {}_2F_1\left(r_{\rm h}+d, r_{\rm a}; 1+d; p_{\rm h}\, p_{\rm a}\right), \end{aligned} \tag{5}$$

where ${}_2F_1$ is the hypergeometric function (see [25], p. 555). Restricting to $p_{\rm h} = p_{\rm a} = p$, the distribution of the total score simplifies to $P^{\Sigma}_{r_{\rm h},p,r_{\rm a},p}(s) = P_{r_{\rm h}+r_{\rm a},p}(s)$, i. e., one finds a composition law similar to the case of the Poissonian distribution.

Starting from the observation that the goal distributions of certain leagues do not seem to be well fitted by the NBD, Greenhough et al. [10] considered fits of the GEV distributions,

$$\begin{aligned} P_{\xi,\mu,\sigma}(n) &= \frac{1}{\sigma}\left(1+\xi\frac{n-\mu}{\sigma}\right)^{-1-1/\xi} \exp\left[-\left(1+\xi\frac{n-\mu}{\sigma}\right)^{-1/\xi}\right] \quad \text{for} \quad \xi \neq 0, \\ P_{\mu,\sigma}(n) &= \frac{1}{\sigma}\exp\left[-\exp\left(-\frac{n-\mu}{\sigma}\right) - \frac{n-\mu}{\sigma}\right] \quad \text{for} \quad \xi = 0, \end{aligned} \tag{6}$$

to the data, obtaining good fits in some cases. Depending on the sign of the parameter ξ, these distributions are called Weibull ($\xi < 0$), Gumbel ($\xi = 0$) and Fréchet ($\xi > 0$) distributions, respectively. The shape parameter ξ controls the asymptotic decay of $P_{\xi,\mu,\sigma}(n)$ for large n, such that increasing values of ξ correspond to stronger feedback effects in terms of the self-affirmation models discussed in the next section.

3. Scoring models

From the discussion of the previous section it should be apparent that the previously employed probability distributions for modelling football scores were chosen rather *ad hoc* mainly from the criterion of best fit to the observed data, but without offering any explanation for their suitability to the purpose. The only exception to this observation is the use of the Poisson distribution which, however, has the drawback of *not* describing the empirical distributions well. The major shortcoming of this latter description appears to be the assumption of *independent* scoring events, ignoring the fact that scoring certainly has a profound feedback on the motivation and thus the likelyhood of subsequent scoring of both teams (via direct motivation/demotivation of the players, but also, e. g., by a strengthening of defensive play in case of a lead), i. e., there is a fundamental component of (positive or negative) feedback in the system. We include such correlation effects by introducing feedback into the binomial model (being a discrete version of the Poissonian model (1) above): consider a football match divided into N time steps (we restrict ourselves here to the natural choice $N = 90$, but good fits are found for any choice of N within reasonable limits) with both teams having the possibility to either score or not score in each time step. Feedback is introduced into the system by having the scoring probabilities p depend on the number n of goals scored so far, $p = p(n)$. Several possibilities arise. For our model "A", upon each goal the scoring probability is modified as

$$p(n) = p(n-1) + \kappa, \tag{7}$$

with some fixed constant κ. Alternatively, one might consider a multiplicative modification rule,

$$p(n) = \kappa p(n-1), \tag{8}$$

which we refer to as model "B". The resulting modified binomial distributions $P_N(n)$ for the total number of goals scored by one team can be computed exactly from a Pascal type recurrence relation [28],

$$P_N(n) = [1-p(n)]P_{N-1}(n) + p(n-1)P_{N-1}(n-1), \tag{9}$$

where, e. g., $p(n) = p_0 + \kappa n$ for model "A" and $p(n) = p_0\kappa^n$ for model "B". For the case of the additive model "A", it can be shown that the continuum limit of $P_N(n)$, i. e., $N \to \infty$ with $p_0 N$ and κN kept fixed, is given by the NBD (4) with $r = p_0/\kappa$ and $p = 1 - \mathrm{e}^{-\kappa N}$ [28]. Thus the good fit of a NBD to the data can be understood from the "microscopic" effect of self-affirmation of the teams or players, without making reference to the somewhat poorly motivated composition of the pure Poissonian model with a gamma distribution. To elaborate on these simple models, one might relax the condition of independence of the scoring of the home and away teams by coupling the adaptation rules upon scoring, for instance as

$$\begin{array}{rclll} p_\mathrm{h}(n) & = & p_\mathrm{h}(n-1)\kappa_\mathrm{h}, & p_\mathrm{a}(n) = p_\mathrm{a}(n-1)/\kappa_\mathrm{a}, & \text{for a goal of the home team,} \\ p_\mathrm{h}(n) & = & p_\mathrm{h}(n-1)/\kappa_\mathrm{h}, & p_\mathrm{a}(n) = p_\mathrm{a}(n-1)\kappa_\mathrm{a}, & \text{for a goal of the away team,} \end{array} \tag{10}$$

which we refer to as model "C". If both teams have $\kappa > 1$, this results in an incentive for the scoring team and a demotivation for the opponent. But a value $\kappa < 1$ is conceivable as well.

These microscopic models are not only related to the Poissonian ansatz and the NBD used earlier, but also distributions of the GEV type can result from a modified microscopical model with feedback. To see this, consider again a series of trials for a number N of time steps. Assume that the probability to score U_1 goals in time step 1 is distributed according to $P_1(U_1) = P(U_1)$ (e. g., with a Poisson distribution P), the probability to score U_2 goals in time step 2 is $P_2(U_2) = P(U_1+U_2)/Z_2$

etc., such that $P_i(U_i) = P(\sum_{j=1}^{i-1} U_j + U_i)/Z_i$. For any continuous distribution P, this means that due to the normalization factors Z_i the distribution of U_i will possess enhanced tails compared to the distribution of U_{i-1} (unless $U_{i-1} = 0$) etc., resulting in a positive feedback effect similar to that of models “A”, “B” and “C”. We refer to this prescription as model “D”. From the results of Bertin and Clusel [23,24] it then follows that the limiting distribution of the total score $n = \sum_{i=1}^{N} U_i$ is a GEV distribution, where the specific form of distribution [in particular the value of the parameter ξ in (6)] depends on the falloff of the original distribution P in its tails.

Our “coarse-grained” scoring models with a single parameter of self-affirmation are clearly a gross over-simplification of the complex psycho-social phenomena on a football pitch and thus a plethora of opportunities for improvement of the description and further studies opens up. For instance, considering the averages over whole leagues or cups, we do not take into account the differences in skill between the teams. Likewise, if time-resolved scoring data were made available, a closer investigation of the intra-team and inter-team motivation and demotivation effects would provide an intriguing future enterprise to undertake. Such data would allow us to investigate the behavior of the (average) scoring probability as a function of playing time, and hence a direct test of our basic assumption of score-dependent scoring probabilities incorporated into the models “A”–“D” discussed above. In particular, the functional form of the thus extracted scoring probability $p(n)$ could be compared to the linear or exponential forms implied by equations (7) and (8) for models “A” and “B”. Some data of this type have been analyzed in [29], showing a clear increase of scoring frequency as the match progresses, thus supporting the presence of feedback as discussed here.

4. Bundesliga and Oberliga

We now turn to the discussion of football matches played in leagues, using the example of football played in Germany. Our main data set consists of the matches played in the “Bundesliga” (men’s premier league FRG, 1963/64–2004/05, ≈ 12 800 matches), the “Oberliga” (men’s premier league GDR, 1949/50–1990/91, ≈ 7700 matches), and the “Frauen-Bundesliga” (women’s premier league FRG, 1997/98–2004/05, ≈ 1050 matches) [30–33]. Beyond the question of which probability distribution or microscopic model might describe these data, we here wanted to see how the score distributions depend on cultural and political circumstances and are possibly different between men’s and women’s leagues. We first determined histograms estimating the probability density functions (PDFs) $P^{\mathrm{h}}(n_{\mathrm{h}})$ and $P^{\mathrm{a}}(n_{\mathrm{a}})$ of the final scores of the home and away teams, respectively [34]. Similarly, we determined histograms for the PDFs $P^{\Sigma}(s)$ and $P^{\Delta}(d)$ of the sums and differences of final scores. To arrive at error estimates on the histogram bins, we utilized the bootstrap resampling scheme [35].

We first considered fits of the PDFs of the phenomenological descriptions considered previously, namely the Poissonian form (1), the NBD (4) and the distributions (6) of extreme value statistics. The parameters of the fits of these types to the data are summarized in table 1 comparing the East German “Oberliga” to the West German “Bundesliga” (1963/64–1990/91, ≈ 8400 matches) during the time of the German division, and in table 2 comparing the data for all games of the German men’s premier league “Bundesliga” to the German women’s premier league “Frauen-Bundesliga”. Not to our surprise, and in accordance with previous findings [5,10], the simple Poissonian ansatz (1) is not found to be an adequate description for any of the data sets. Deviations occur here mainly in the tails with large numbers of goals which in general are found to be fatter than what can be accommodated by a Poissonian model, whereas the distribution peaks are reasonably well represented. On the contrary, the NBD form (4) models all of the considered data well as is illustrated with fits of the corresponding form to our data in figure 1 comparing “Oberliga” and “Bundesliga” and in figure 2 presenting “Bundesliga” and “Frauen-Bundesliga”. Comparing the leagues, we find that the parameters r of the NBD fits for the “Bundesliga” are about twice as large as for the “Oberliga”, whereas the parameters p are smaller for the “Bundesliga”, cf. the data in table 1. Recalling that the form (4) is in fact the continuum limit of the feedback model “A” discussed above, these differences translate into larger values of κ and smaller values of p_0

Table 1. Fits of the phenomenological distributions (1), (4) and (6) to the data for the East German "Oberliga" between 1949/50 and 1990/91 and for the West German "Bundesliga" for the seasons of 1963/64–1990/91.

		Oberliga		Bundesliga	
		Home	Away	Home	Away
Poisson	λ	1.85 ± 0.02	1.05 ± 0.01	2.01 ± 0.02	1.17 ± 0.01
	χ^2/d.o.f.	12.5	12.8	6.53	7.31
NBD	p	0.17 ± 0.01	0.14 ± 0.01	0.11 ± 0.01	0.10 ± 0.01
	r	9.06 ± 0.88	6.90 ± 0.84	15.9 ± 2.10	11.3 ± 1.84
	p_0	0.0191	0.0112	0.0213	0.0126
	κ	0.0021	0.0016	0.0013	0.0011
	χ^2/d.o.f.	0.99	4.09	0.68	2.29
GEV	ξ	-0.05 ± 0.01	0.02 ± 0.01	-0.09 ± 0.01	-0.01 ± 0.01
	μ	1.12 ± 0.02	0.49 ± 0.02	1.28 ± 0.02	0.58 ± 0.02
	σ	1.30 ± 0.02	0.90 ± 0.02	1.36 ± 0.02	0.96 ± 0.02
	χ^2/d.o.f.	1.93	5.04	1.83	4.74
Gumbel	μ	1.12 ± 0.02	0.48 ± 0.02	1.28 ± 0.02	0.59 ± 0.01
	σ	1.25 ± 0.01	0.92 ± 0.01	1.25 ± 0.01	0.95 ± 0.01
	χ^2/d.o.f.	4.13	4.65	12.9	4.06

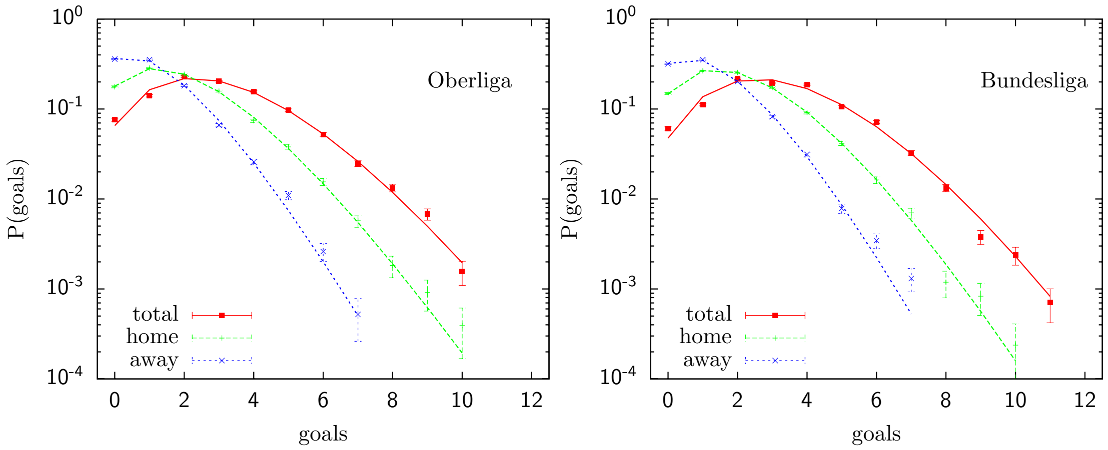

Figure 1. Probability density of goals scored by home and away teams, and of the total number of goals scored in a match of the GDR "Oberliga" (left) and the FRG "Bundesliga" (right), restricted to the seasons of 1963/64–1990/91. The lines for "home" and "away" show fits of the NBD (4) to the data; the line for "total" denotes the resulting distribution (5) for the sum.

for the "Oberliga" results. That is to say, scoring a goal in a match of the East German premier league was a more encouraging event than scoring a goal in a match of the West German league. Alternatively, this observation might be interpreted as a stronger tendency of the perhaps more professionalized teams of the West German league to switch to a strongly defensive mode of play in case of a lead. Consequently, the tails of the distributions are slightly fatter for the "Oberliga" than for the "Bundesliga". Comparing the results for the "Frauen-Bundesliga" to those for the "Bundesliga", even more pronounced tails are found for the former, resulting in very significantly larger values of the self-affirmation parameter κ for the matches of the women's league, see the fit parameters collected in table 2 and the fits of the NBD type presented in figure 2.

Considering the fits of the GEV distributions (6) to the data for all three leagues, we find that extreme value statistics are in general a reasonably good description of the data. The shape parameter ξ is always found to be small in modulus and negative in the majority of the cases,

Table 2. Fits of the phenomenological distributions (1), (4) and (6) to the data for the German men's premier league "Bundesliga" between 1963/64 and 2004/05 and for the German women's premier league "Frauen-Bundesliga" for the seasons of 1997/98–2004/05.

		Bundesliga		Frauen-Bundesliga	
		Home	Away	Home	Away
Poisson	λ	1.91 ± 0.01	1.16 ± 0.01	1.78 ± 0.04	1.36 ± 0.04
	χ^2/d.o.f.	9.21	9.13	14.6	14.4
NBD	p	0.11 ± 0.01	0.09 ± 0.01	0.45 ± 0.03	0.46 ± 0.03
	r	16.24 ± 1.82	12.08 ± 1.69	2.38 ± 0.24	1.97 ± 0.22
	p_0	0.0202	0.0125	0.0160	0.0133
	κ	0.0012	0.0010	0.0067	0.0068
	χ^2/d.o.f.	1.08	2.22	2.32	1.37
GEV	ξ	-0.10 ± 0.01	-0.02 ± 0.01	0.04 ± 0.04	0.25 ± 0.07
	μ	1.17 ± 0.02	0.57 ± 0.01	0.83 ± 0.08	0.77 ± 0.07
	σ	1.33 ± 0.01	0.96 ± 0.01	1.49 ± 0.06	1.18 ± 0.05
	χ^2/d.o.f.	3.43	7.95	3.40	1.55
Gumbel	μ	1.18 ± 0.01	0.58 ± 0.01	0.81 ± 0.08	0.58 ± 0.07
	σ	1.21 ± 0.01	0.94 ± 0.01	1.53 ± 0.05	1.31 ± 0.05
	χ^2/d.o.f.	24.5	7.26	3.17	4.09

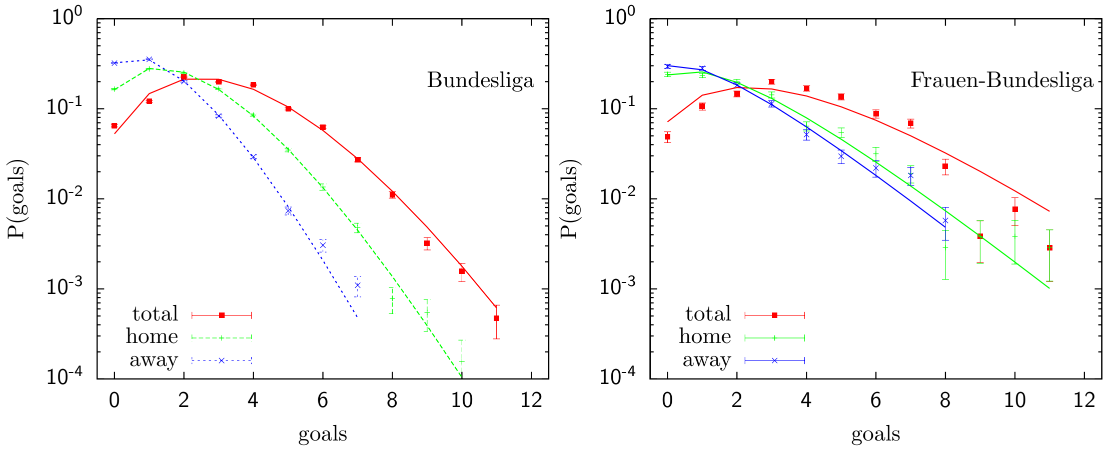

Figure 2. Probability density of goals scored in the German premier league "Bundesliga" for all seasons (left) and in the women's "Frauen-Bundesliga" (right).

indicating a distribution of the Weibull type (which is in agreement with the findings of [10]). On the other hand, fixing $\xi = 0$ yields overall clearly larger values of χ^2 per degree-of-freedom (d.o.f.), indicating that the data are hardly compatible with a distribution of the Gumbel type. Comparing "Oberliga" and "Bundesliga", we consistently find larger values of the parameter ξ for the former, indicative of the comparatively fatter tails of these data discussed above, see the data in table 1. The location parameter μ, on the other hand, is larger for the West German league which features a larger average number of goals per match (which can be read off also more directly from the λ parameter of the Poissonian fits), while the scale parameter σ is similar for both leagues. Compared to the results for the NBD, we do not find any cases where the GEV distributions would provide the best fit to the data, so clearly the leagues considered here are not of the type of the general "domestic" league data for which Greenhough et al. [10] found better matches with the GEV than for the NBD statistics. Similar conclusions hold true for the comparisons of "Bundesliga" and "Frauen-Bundesliga", with the latter taking on the role of the "Oberliga".

In total, the best fits so far are clearly achieved by the NBD ansatz. Since this distribution is

Table 3. Fit results for models "A" and "B". Fits were performed to the score distributions of the home and away teams only and the resulting model estimates for the sums and differences of goals compared to the data.

		Bundesliga 04/05	Bundesliga 90/91	Oberliga	Women
Model "A"	$p_{0,\mathrm{h}}$	0.0199 ± 0.0002	0.0210 ± 0.0002	0.0188 ± 0.0002	0.0159 ± 0.0005
	$\kappa_{0,\mathrm{h}}$	0.0015 ± 0.0001	0.0016 ± 0.0002	0.0024 ± 0.0002	0.0070 ± 0.0005
	$p_{0,\mathrm{a}}$	0.0125 ± 0.0002	0.0125 ± 0.0001	0.0112 ± 0.0001	0.0132 ± 0.0004
	$\kappa_{0,\mathrm{a}}$	0.0012 ± 0.0001	0.0013 ± 0.0002	0.0018 ± 0.0002	0.0071 ± 0.0007
Home	χ^2_h/d.o.f.	1.01	0.68	1.07	2.28
Away	χ^2_a/d.o.f.	2.31	2.37	4.23	1.44
Total	χ^2_Σ/d.o.f.	16.6	11.5	5.33	12.4
Difference	χ^2_Δ/d.o.f.	18.6	14.0	5.63	2.86
Model "B"	$p_{0,\mathrm{h}}$	0.0200 ± 0.0002	0.0211 ± 0.0002	0.0189 ± 0.0002	0.0166 ± 0.0005
	$\kappa_{0,\mathrm{h}}$	1.0679 ± 0.0060	1.0695 ± 0.0072	1.1115 ± 0.0083	1.3146 ± 0.0303
	$p_{0,\mathrm{a}}$	0.0125 ± 0.0001	0.0125 ± 0.0002	0.0112 ± 0.0001	0.0138 ± 0.0004
	$\kappa_{0,\mathrm{a}}$	1.0932 ± 0.0106	1.1015 ± 0.0124	1.1526 ± 0.0149	1.4115 ± 0.0543
Home	χ^2_h/d.o.f.	1.25	0.71	0.75	3.24
Away	χ^2_a/d.o.f.	1.96	20.2	3.35	0.95
Total	χ^2_Σ/d.o.f.	16.9	11.8	5.40	13.5
Difference	χ^2_Δ/d.o.f.	18.4	13.8	5.26	2.82

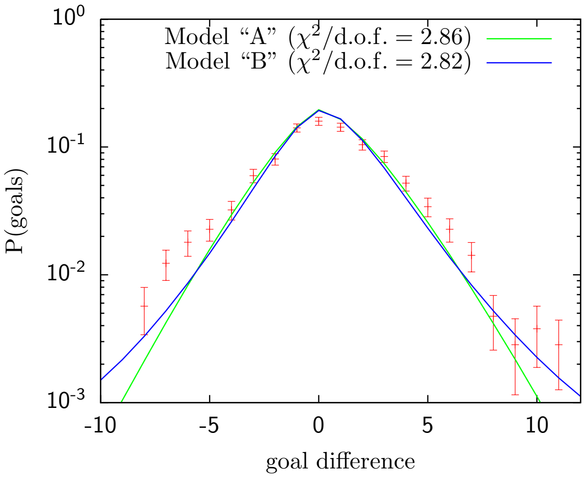

Figure 3. Goal differences in the German women's premier league together with fits of models "A" and "B".

obtained only as the continuum limit of the microscopic model "A", it is interesting to see how fits of the *exact* distribution (for $N = 90$) resulting from the recurrence (9) for model "A", but also fits of the multiplicatively modified binomial distribution of model "B" compare to the results found above. We performed fits to the exact distributions of both models by employing the simplex method [36] to minimize the total χ^2 of the data for the home and away scores. Alternatively, we also considered fitting additionally to the sums and differences in a simultaneous fit and found very similar results with an only slight improvement of the fit quality for the sums and differences at the expense of somewhat worse fits for the home and away scores. We summarize the fit results in table 3. We also performed fits to the more elaborate model "C", but found the results rather similar to those of the simpler model "B" and hence do not present the results here. Comparing the results of model "A" to the fits of the limiting NBD, we find almost identical fit qualities for the final scores of both teams. However, the sums and differences of scores are considerably better described by model "A", indicating that here the deviations from the continuum limit are still relevant. In figure 3, we present the differences of goals in the German women's premier league together with the fits of models "A" and "B". The multiplicative model "B", where each goal motivates a team even more than the previous one, within the statistical errors yields fits of the same quality as model “A”, such that a distinct advantage cannot be attributed to either of them, cf. the data in table 3.

5. FIFA World Cup

Somewhat different conditions than for football in premier leagues apply to the case of international football tournaments. In particular, we considered the score data of the “FIFA World Cup” series from 1930 to 2006, focusing on the results from the qualification stage ($\approx$ 4800 matches) [37] (the final knockout stage follows different rules: matches are played on neutral grounds – apart from the team of the host country – and games cannot end in a draw). The results of fits of the phenomenological distributions (1), (4) and (6) as well as the models “A” and “B” are collected in table 4. Compared to the domestic league data discussed above, the results of the World Cup show distinctly heavier tails, cf. the presentation of the data in figure 4. Considering the fit results, this leads to good fits for the heavy-tailed distributions, and, in particular, in this case the GEV distributions provide a better fit than the NBD, similar to what was found by Greenhough et al. [10] for some of their data. This difference to the German league data discussed above can be attributed to the possibly very large differences in skill between the opposing teams occurring since all countries are allowed to participate in the qualification round. A glance back to table 2

Table 4. Fit results for the qualification phase of the “FIFA World Cup” series from 1930 to 2006.

		Home	Away
Poisson	λ	1.52 ± 0.02	0.90 ± 0.01
	χ^2/d.o.f.	21.7	28.5
NBD	p	0.36 ± 0.02	0.37 ± 0.02
	r	3.08 ± 0.20	1.84 ± 0.12
	p_0	0.0152	0.0095
	κ	0.0050	0.0051
	χ^2/d.o.f.	2.88	1.91
GEV	ξ	0.10 ± 0.02	0.17 ± 0.02
	μ	0.85 ± 0.03	0.37 ± 0.03
	σ	1.21 ± 0.03	0.87 ± 0.02
	χ^2/d.o.f.	1.13	2.44
Gumbel	μ	0.79 ± 0.03	0.27 ± 0.03
	σ	1.30 ± 0.02	0.95 ± 0.02
	χ^2/d.o.f.	3.68	13.7
Model “A”	p_0	0.0151 ± 0.0002	0.0094 ± 0.0002
	κ	0.0052 ± 0.0003	0.0053 ± 0.0003
	χ^2/d.o.f.	3.12	2.08
Model “B”	p_0	0.0154 ± 0.0002	0.0096 ± 0.0002
	κ	1.2725 ± 0.0130	1.4490 ± 0.0281
	χ^2/d.o.f.	1.00	0.85

reveals a remarkable similarity with the parameters of the “Frauen-Bundesliga” (e. g., in both cases the NBD parameters p are comparatively large while r is small, and the GEV parameters ξ are positive), where a similar explanation appears quite plausible since the very good players are concentrated in two or three teams only. Turning to the fits of the models “A” and “B”, we again find model “A” to fit rather similar to its continuum approximation, the NBD. On the other hand, model “B” describes the data extremely well, for the away team even better than the GEV distributions (6).

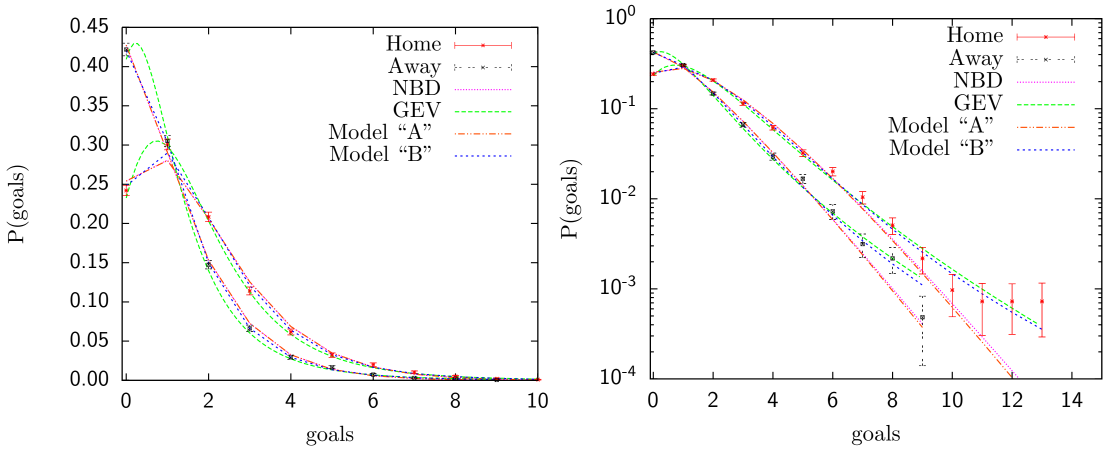

Figure 4. Probability density of goals scored by the home and away teams in the qualification stage of the "FIFA World Cup" series on a linear (left) and logarithmic (right) scale.

6. Tabletop football

Finally, we also conducted our own empirical experiments relying on the football fever of the visitors of the "Science Summer 2008" ("Wissenschaftssommer 2008") exhibition, the central opening event of the German "Year of Mathematics" held on Leipzig's Augustus Square in July 2008. To this end we rented two football tables on which visitors could play matches in teams of up to two players, see figure 5. All results were recorded and analyzed on-site in order to involve the visitors as closely as possible. By the end of the week, a total of about 2500 visitors had participated in the table football matches, in total contributing about 1000 results. With a fixed playing time of three minutes, a typical match resulted in about 5–10 goals, quite significantly more than the average number of goals (≈ 3) scored in matches of the professional football leagues considered above. Still, the overall trend of the goal distribution turned out to be surprisingly similar to the features seen for professional leagues, indicating a certain degree of universality of our interpretation – the football fever was indeed already visually apparent during the whole exhibition week (which, admittedly, was particularly suited since it ended with the final match of the UEFA Euro 2008 in Vienna between Germany and Spain (0:1)). The results can be inspected in the left panel of figure 6, where the empirical goal distribution (for one-to-one matches) is compared to our various fit models. As before, the simple Poissonian ansatz (with a $\chi^2/\text{d.o.f.} = 8.9$) does not work at all, but both feedback models "A" and "B" give satisfactory fits to the data (with $\chi^2/\text{d.o.f.} \approx 2$). For the self-affirmation parameter κ we find here $\kappa = 0.0074 \pm 0.0006$ for the additive model "A" and $\kappa = 1.11 \pm 0.01$ for the multiplicative model "B". In order to compare directly with the goal distribution of the home teams of the German Bundesliga, in the right panel of figure 6 we show the table football data together with the results from the German Bundesliga, where the former was renormalized to yield the same average number of goals per match. While it does not come unexpected that the two curves do not really fall on top of each other, the overall trend is surprisingly similar, given the completely different set-ups leading to the two data sets.

7. Conclusions

By analyzing German domestic and international football score data we have shown that the goal distributions can be modeled with a certain class of modified binomial models supplemented by a built-in effect of self-affirmation of the teams upon scoring a goal. The simple Poissonian ansatz assuming independent scoring probabilities is clearly ruled out. The NBD suggested earlier [5], which fits many of our data sets quite satisfactorily, can in fact be understood as the limiting distribution of our additive model "A". It should be stressed that the *exact* distribution of model

Figure 5. Football fever infected visitors of the “Science Summer 2008” exhibition in Leipzig fighting in a tabletop football match. From left to right: Former Foreign Minister and Vice-Chancellor of the Federal Republic of Germany Dr. Klaus Kinkel, Vice-Rector for Research of the University of Leipzig Prof. Dr. Martin Schlegel, Lord Mayor of the City of Leipzig Burkhard Jung, and Parliamentary State Secretary to the German Federal Minister of Education and Research Thomas Rachel.

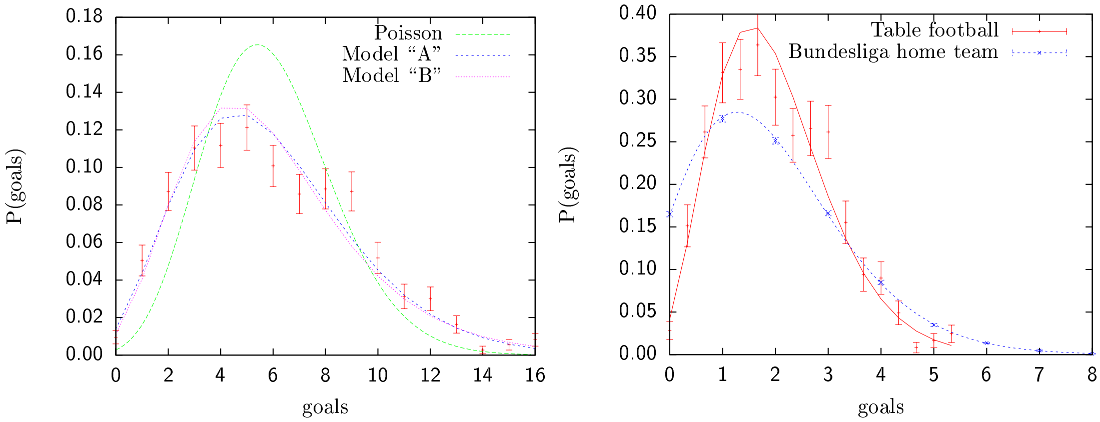

Figure 6. Probability density of goals scored in the tabletop football matches during the “Science Summer 2008” exhibition on the Augustus Square of the City of Leipzig (left plot). The right plot shows the same data renormalized to the average number of goals per game scored by the home teams of the German Bundesliga, whose goal distribution is shown for comparison.

“A” provides in general rather better fits to the data than the limiting NBD. This is particularly pronounced for the sums and differences of goals scored. However, the quality of the fits is limited in cases with heavier tails such as the qualification round of the “FIFA World Cup” series. Here the multiplicative model “B”, in which each goal motivates the team even more than the previous one, provides an outstanding fit to these data as well as the data from the German domestic leagues. Thus, the contradicting evidence for better fits of some football score data with NBD and other data with GEV distributions is reconciled with the use of a plausible microscopic model covering both cases by successfully interpolating between the two extremes. Also the tabletop football score data of our field study with visitors of the “Science Summer 2008” are well represented with our feedback models, underscoring their apparently rather universal applicability.

Comparing the score data between the separate German premier leagues during the cold war times, we find heavier tails for the East German league. In terms of our microscopic models, this corresponds to a stronger component of self-affirmation as compared to the West German league. Similarly, the German women's premier league "Frauen-Bundesliga" shows a much stronger feedback effect than the men's premier league, with at first glance surprisingly many parallels to the "FIFA World Cup" series. We also analyzed the results from further leagues, such as the Austrian, Belgian, British, Bulgarian, Czechoslovak, Dutch, French, Hungarian, Italian, Portuguese, Romanian, Russian, Scottish and Spanish premier leagues, and arrived at similar conclusions. In general, we find less professionalized leagues to feature stronger components of positive feedback upon scoring a goal, perhaps indicating a still stronger infection with the football fever there ...

8. Acknowledgements

The authors are grateful to O. Penrose and S. Zachary for fruitful discussions. This work was partially supported by the Deutsche Forschungsgemeinschaft (DFG) under grant No. JA483/22–1, the EU RTN–Network 'ENRAGE': *Random Geometry and Random Matrices: From Quantum Gravity to Econophysics* under grant No. MRTN–CT–2004–005616, and the Graduate College CDFA–02–07 of the Deutsch-Französische Hochschule (DFH–UFA). M.W. acknowledges support by the DFG through the Emmy Noether Programme under contract No. WE4425/1–1.

References

1. Stauffer D., Physica A, 2004, **336**, 1.
2. Econophysics and Sociophysics: Trends and Perspectives, Chakrabarti B.K., Chakraborti A., Chatterjee A., eds. Wiley-VCH, Berlin, 2006.
3. Clarke S.R., Norman J.M., The Statistician, 1995, **45**, 509; Malacarne L.C., Mendes R.S., Physica A, 2000, **286**, 391; Dobson S., Goddard J., Eur. J. Oper. Res., 2003, **148**, 247; Goddard J., Asimakopoulos I., J. Forecast., 2004, **23**, 51; Onody R.N., de Castro P.A., Phys. Rev. E, 2004, **70**, 037103; Linthorne N.P., Evertt D.J., Sports Biomechanics, 2006, **5**, 5; Sprem J. Simulation der Fußball-Bundesliga. Master's thesis, Universität zu Köln, 2006; Heuer A., Rubner O., Eur. Phys. J. B, 2009, **67**, 445.
4. Moroney M.J. Facts from Figures, 3rd edition. Penguin, London, 1956.
5. Reep C., Pollard R., Benjamin B., J. Roy. Stat. Soc. A, 1971, **134**, 623.
6. Pollard R., J. Am. Stat. Assoc., 1973, **68**, 351.
7. Clarke S.R., Norman J.M., The Statistician, 1995, **44**, 509.
8. Dyte D., Clarke S.R., J. Op. Res. Soc., 2000, **51**, 993.
9. Malacarne L.C., Mendes R., Physica A, 2000, **286**, 391.
10. Greenhough J., Birch P.C., Chapman S.C., Rowlands G., Physica A, 2002, **316**, 615.
11. Arbous A.G., Kerrich J.E., Biometrics, 1951, **7**, 340.
12. Kotz S., Nadarajah S. Extreme Value Distributions: Theory and Applications. World Scientific, Singapore, 2000.
13. Bramwell S.T., Holdsworth P.C.W., Pinton J.-F., Nature, 1998, **396**, 552.
14. Bramwell S.T., Christensen K., Fortin J.-Y., Holdsworth P.C.W., Jensen H.J., Lise S., López J.M., Nicodemi M., Pinton J.-F., Sellitto M., Phys. Rev. Lett., 2000, **84**, 3744.
15. Bouchaud J.-P., Mézard M., J. Phys. A, 1997, **30**, 7997.
16. Berg B.A., Billoire A., Janke W., Phys. Rev. E, 2002, **65**, 045102(R).
17. Dayal P., Trebst S., Wessel S., Würtz D., Troyer M., Sabhapandit S., Coppersmith S.N., Phys. Rev. Lett., 2004, **92**, 097201.
18. Bittner E., Janke W., Europhys. Lett., 2006, **74**, 195.
19. Weigel M., Phys. Rev. E, 2007, **76**, 066706.
20. Noullez A., Pinton J.-F., Eur. Phys. J. B, 2002, **28**, 231.
21. Varotsos P.A., Sarlis N.V., Tanaka H.K., Skordas E.S., Phys. Rev. E, 2005, **72**, 041103.
22. Dahlstedt K., Jensen H.J., J. Phys. A, 2001, **34**, 11193.
23. Bertin E., Phys. Rev. Lett., 2005, **95**, 170601.
24. Bertin E., Clusel M., J. Phys. A, 2006, **39**, 7607.
25. Abramowitz M., Stegun I.A. Handbook of Mathematical Functions. Dover Publications, New York, 1970.
26. Fisz M. Wahrscheinlichkeitsrechnung und Mathematische Statistik. VEB Deutscher Verlag der Wissenschaften, Berlin, 1989.
27. Feller W. An Introduction to Probability Theory and its Applications, vol. 1, 3rd edition. Wiley, New York, 1968.
28. Bittner E., Nußbaumer A., Janke W., Weigel M., Europhys. Lett., 2007, **78**, 58002; Bittner E., Nußbaumer A., Janke W., Weigel M., Eur. Phys. J. B, 2009, **67**, 459.
29. Dixon M.J., Robinson M.E., The Statistician, 1998, **47**, 523.
30. http://www.fussballdaten.de.
31. http://www.fussballportal.de.
32. http://www.nordostfussball.de.
33. http://www.sportergebnise.de.
34. To ensure reliable error estimates, in the fits presented below we ignored histogram bins consisting of single or isolated entries, i. e., outliers.
35. Efron B., SIAM Review, 1979, **21**, 460.
36. Press W.H., Teukolsky S.A., Vetterling W.T., Flannery B.P. Numerical Recipes in C – The Art of Scientific Computing, 2nd edition. CUP, Cambridge, 1992.
37. http://www.rdasilva.demon.co.uk/football.html.

Футбольна лихоманка: модель розподілу голів із самопідтвердженням

В. Янке[1], Е. Біттнер[1], А. Нусбаумер[1], М. Вайгель[2,3]

[1] Інститут теоретичної фізики та Центр теоретичних студій університету Лейпцига, Аугустпляц 10/11, 04009, Ляйпциг, Німеччина

[2] Інститут теоретичної фізики, Університет землі Саар, 66041 Саарбрюкен, Німеччина

[3] Інститут фізики, Університет ім. Йоганна Гутенберга в Майнці, 55099, Майнц, Німеччина

Отримано 22 липня 2009 р.

Результати футбольних матчів, як і більшості інших ігрових видів спорту, залежать від таких чинників, як майстерність гравців, досвід тренерів, а також багатьох зовнішніх факторів, які, внаслідок своєї складної природи, мабуть, краще вважати випадковими. Серед згаданих факторів – непередбачуваність руху м'яча, ігрова форма футболістів у день матчу, а також умови середовища, такі як погода та поведінка вболівальників. За таких обставин варто проаналізувати футбольні рахунки крізь призму математичної статистики з метою розділення детерміністичних ефектів від випадкових і з'ясування, який внесок привносять колективна та соціальна природа складових системи на результуючі спостережні величини. Розглядаючи розподіли ймовірностей забитих голів для команд господарів та гостей, виявилось, що "хвости" цих розподілів не можуть бути описані на основі моделі Пуассона чи біноміальної моделі, які випливають із припущення про нескорельовану природу випадкових подій. З іншого боку, деякі інші характерні розподіли густини ймовірності, які обговорюються в контексті статистики екстремальних значень або ж так званих від'ємних біноміальних розподілів, достатньо добре відтворюють ці результати. Так виглядає, що на сьогодні немає аргументованих пояснень, чому найпростіша модель Пуассона не є застосовною в згаданих випадках, а спостерігаються замість цього вищеописані розподіли. Щоб заповнити цей пробіл, ми ввели низку мікроскопічних моделей для опису результатів матчів, які описують випадкові процеси Бернуллі з простою компонентою самопідтвердження. Ці моделі дозволили нам несподівано добре описати спостережні розподіли ймовірностей, а феноменологічні розподіли, які використовувалися перед тим, трактуються в рамках нашого підходу як часткові випадки. Ми проаналізували результати футбольних матчів багатьох європейських чемпіонатів, а також міжнародних турнірів, включно з даними усіх останніх турнірів Чемпіонату світу ФІФА, і показали, що запропоновані моделі є застосовними в усіх вищезгаданих випадках. Для повноти картини, ми провели польові дослідження з відвідувачами наукових виставок з метою збору додаткових даних про результати матчів з настільного футболу. Як виявилось, ці останні дані також достатньо добре описуються в рамках наших моделей зі зворотнім зв'язком, підкреслюючи їх очевидну універсальність.

Ключові слова: *спортивна статистика, екстремальні розподіли, моделі зі зворотнім зв'язком*

PACS: *89.20.-a, 02.50.-r*