



ELSEVIER

Physica A 316 (2002) 615–624

PHYSICA A

www.elsevier.com/locate/physa

Football goal distributions and extremal statistics

J. Greenhough*, P.C. Birch, S.C. Chapman, G. Rowlands

Department of Physics, University of Warwick, Coventry CV4 7AL, UK

Received 1 November 2001

Abstract

We analyse the distributions of the number of goals scored by home teams, away teams, and the total scored in the match, in domestic football games from 169 countries between 1999 and 2001. The probability density functions (PDFs) of goals scored are too heavy-tailed to be fitted over their entire ranges by Poisson or negative binomial distributions which would be expected for uncorrelated processes. Log-normal distributions cannot include zero scores and here we find that the PDFs are consistent with those arising from extremal statistics. In addition, we show that it is sufficient to model English top division and FA Cup matches in the seasons of 1970/71–2000/01 on Poisson or negative binomial distributions, as reported in analyses of earlier seasons, and that these are not consistent with extremal statistics.

© 2002 Elsevier Science B.V. All rights reserved.

PACS: 05.20.-y; 05.90.+m; 89.90.+n; 01.80.+b

Keywords: Football; Extremal statistics

1. Introduction

Few authors have considered football scores from a statistical point of view. Moroney [1] showed that the numbers of goals scored by individual teams, and the total goal scores, were well described by a “modification of the Poisson”; Reep et al. [2] later identify this as the negative binomial distribution, and found similar results for other ball games. Maher [3] then pointed out that a negative binomial distribution may arise from the aggregate of Poisson-distributed scores with a different mean for each team. The short-term predictability of results has subsequently been modelled using independent Poisson distributions with means dependent on teams’ past performances

* Corresponding author.

E-mail address: greenh@astro.warwick.ac.uk (J. Greenhough).

[4,5]; an improved model [6] includes the scoring-rate dependence on both time and the current score.

Other aspects of the game have been examined including the effects of certain conditions on the scores—see Ref. [7] for a review. Seeking a broader understanding, it has been suggested [8] that the distribution of goals per player may be linked with anomalous diffusion via the Zipf–Mandelbrot law. In this paper we show, in agreement with analyses of matches from the 1960s [1,2], that 13,000 English top division and 5000 FA Cup matches between the seasons of 1970/71 and 2000/01 [9] are closely-fitted by either Poisson or negative binomial distributions. However, we find that the number of goals scored by home and away teams, and the total goals, in over 135,000 domestic football games (leagues and cups, hereafter referred to as domestic matches) from 169 countries between 1999 and 2001 [10], cannot be fitted over their entire ranges by Poisson or negative binomial distributions. Instead, we find that the data can be modelled by extremal statistics (explained in Section 3).

The ubiquity of power-law relationships in both nature [11–13] and the field of econophysics [14–17] has spawned a significant amount of literature in recent years. Intriguingly, extremal statistics in a global measure are found in turbulent fluids and other highly-correlated systems [18–23]. Hence the significance and origin of extremal and power-law-tailed distributions are currently of considerable interest in statistical physics; the use of probability distributions in the modelling of complex systems is a topical approach to the inverse problem. From an operational perspective, knowledge of the statistics would be an important constraint on any model for the game.

2. The probability density function (PDF)

The first step in our analysis of each data set is the construction of its PDF. The PDF $P(x)$ of a variable X is defined such that the probability that X lies within a small interval dx centred on $X = x$, is equal to $P(x)dx$. $P(x)$ is normalised so that

$$\int_{x_{\min}}^{x_{\max}} P(x) dx = 1. \quad (1)$$

Here, x takes the integer values of goals scored ($x_{\min} = 0$ and x_{\max} is the maximum number of goals scored in the sample of matches) so Eq. (1) becomes $\sum_{x_{\min}}^{x_{\max}} P(x) = 1$. We further normalise each PDF to the sample mean μ and standard deviation σ to enable comparison with extremal distributions (see Section 3).

The Poisson distribution is defined by

$$P(x|\mu) = \frac{\mu^x}{x!} e^{-\mu} I_{(0,1,\dots)}(x), \quad (2)$$

where $I_{(0,1,\dots)}(x)$ ensures that $P(x) = 0$ for non-integer x . In the Poisson PDF $\mu = \sigma^2$; for data to be well-fitted by this distribution we require $\mu \approx \sigma^2$. It is explained in Refs. [1,2] that this condition does not hold for football goals because a constant probability per unit time of scoring a goal is not a valid assumption. Instead, a

compound Poisson or negative binomial distribution is used, defined by

$$P(x | r, p) = \binom{r+x-1}{x} p^r q^x I_{(0,1,\dots)}(x), \tag{3}$$

where x is the number of goals scored with probability q per goal before r “failed goals” (probability $p = 1 - q$) have occurred. The negative binomial PDF has $\mu = r(1 - p)/p$ and $\sigma^2 = r(1 - p)/p^2$; fitting to data thus requires $\mu/\sigma^2 = p \leq 1$ and $\mu p/(1 - p) = r$, where we round r to the nearest integer.

3. Extremal statistics

Our data analysis presented in Section 4 (Figs. 1–3) shows that the tails of the PDFs of goal scores in the domestic matches clearly deviate from both the Poisson and the negative binomial distributions. Here, we compare the PDFs of the data with those arising from extremal statistics. We choose extremal distributions fitted over the entire dataset in preference to a piecewise fit of arbitrary functions as (1) they have been observed in a wide variety of natural systems; (2) they may suggest a physical interpretation, as they arise in situations where only the largest events are observed;

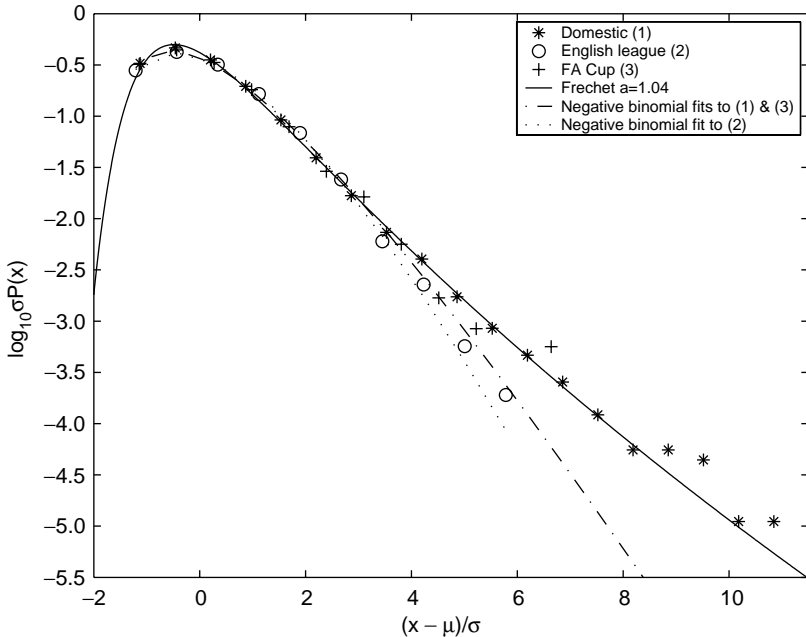


Fig. 1. PDFs of goals scored by home teams, normalised with respect to μ and σ , showing how domestic matches are more closely fitted by a Fréchet distribution than by a negative binomial. Coincident curves are plotted as a single line as indicated in the legend.

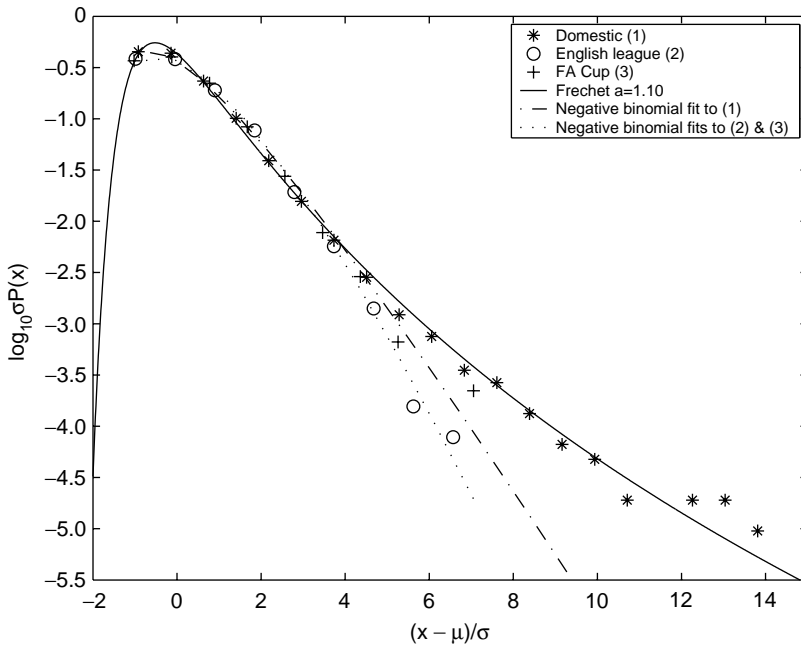


Fig. 2. PDFs of goals scored by away teams, normalised with respect to μ and σ , showing how domestic matches are more closely fitted by a Fréchet distribution than by a negative binomial. Coincident curves are plotted as a single line as indicated in the legend.

(3) following normalisation of the data, only one parameter (a) remains to be estimated, and (4) unlike log-normal PDFs, they can be applied to data including zero values.

The two limiting distributions of interest are “Gumbel’s asymptote” and Fréchet [24–26]. In outline, the limiting distributions result from selecting the maximum value x_{\max} from each of a large number of large samples whose individual members are drawn from a distribution $P(x)$. When $P(x)$ decreases more rapidly than any power-law (as $x \rightarrow \infty$), “Gumbel’s asymptote” has the form

$$P_G(x_{\max}) = K(e^{u-e^u})^a \quad \text{with } u = b(x - s), \tag{4}$$

where in the limit of an infinite number of measurements $a \equiv 1$; the constants K , b , and s are fixed by normalisation as in Section 2 (see Refs. [27,28]). Selecting the *second* largest values from the same large samples produces a PDF of the same functional form as Eq. (4) but with $a \equiv 2$.

Fréchet distributions $P_F(x_{\max})$ arise in the same manner when the underlying PDF $P(x)$ is power-law; the power-law tail of this underlying distribution is preserved in the Fréchet, thus lending itself to the fitting of heavy-tailed data. Mathematically, $P_F(x_{\max})$ can be defined by Eq. (4) but with $u = \alpha + \beta \ln(1 + x/G)$, where K , α , and G are again fixed by normalisation, and $\beta = (1 - a)^{-1}$ [28]. These curves exist for $a > 1$.

A simple heavy-tailed distribution often encountered in nature is the log-normal. Log-normal distributions with the same means and variances as the datasets provide

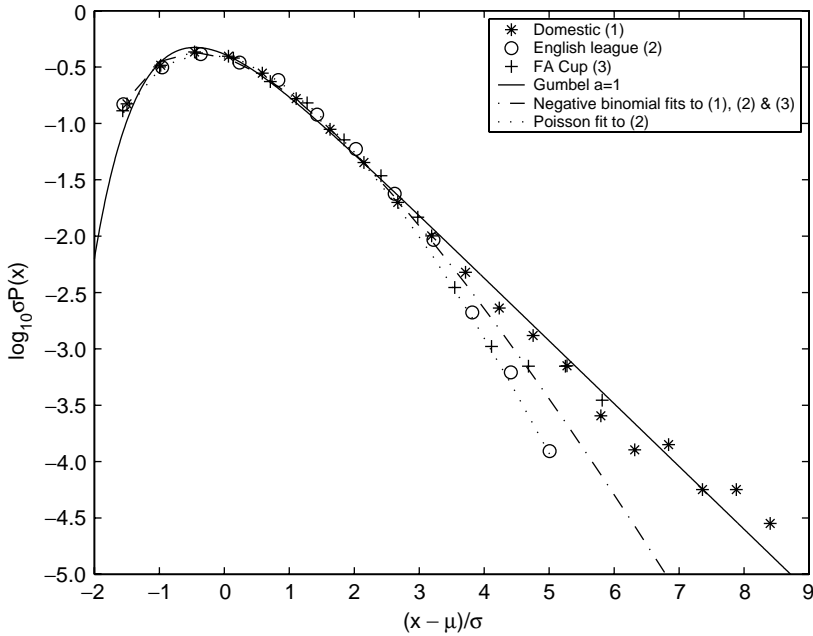


Fig. 3. PDFs of total goals scored, normalised with respect to μ and σ , showing how domestic matches are more closely fitted by a Gumbel distribution than by a negative binomial. Coincident curves are plotted as a single line as indicated in the legend. Domestic $\mu=2.9$, $\sigma=1.9$; league $\mu=2.6$, $\sigma=1.7$; Cup $\mu=2.8$, $\sigma=1.8$.

very poor models in all cases if scores of zero are included. The domestic home and away scores are quite well modelled by log-normal PDFs *providing zero scores are neglected*. Alternatively, one goal can be added to all scores but, since the log-normal is not invariant under translation, the results are no more meaningful. Scores of zero occur frequently and should not be removed; we seek a single heavy-tailed PDF appropriate for modelling integer data from zero upwards.

4. Results

As discussed in Section 1, the Poisson distribution has been demonstrated to be inferior to the negative binomial when modelling football scores; only where this is *not* the case do we include a Poisson fit in Figs. 1–3. In Fig. 1, we show the PDFs of home team scores with their respective negative binomial PDFs (fitted to μ and σ) along with the best-fit extremal distribution (see Section 3) for the domestic matches. While the league scores follow a negative binomial PDF, it is clear that the domestic scores are better described by a Fréchet distribution beyond about $\mu+3\sigma$ (a home score of about 6 goals). Although the Cup scores are suggestive of some departure from a negative binomial PDF, we cannot quantify the functional form of this tail. Counting errors caused by binning a finite dataset are omitted from Figs. 1–3 for the purpose of

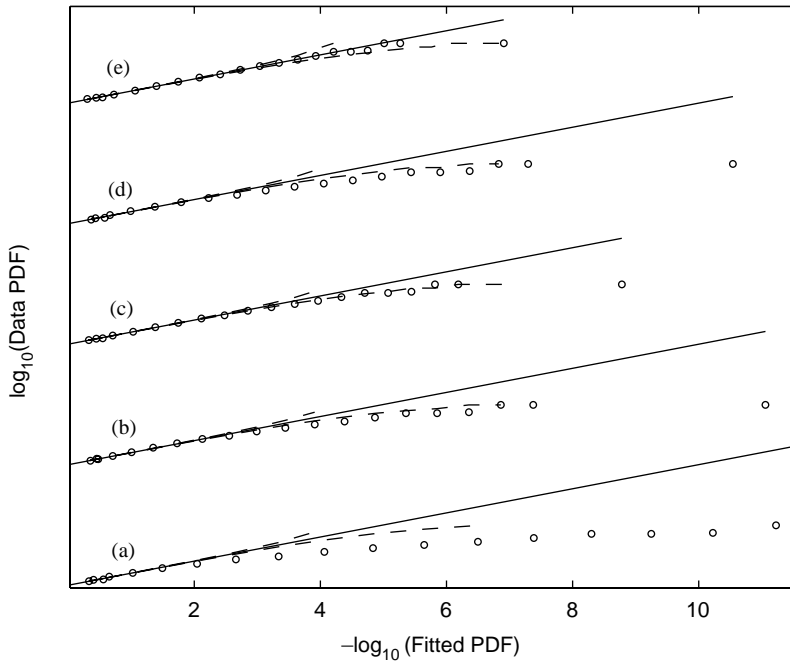


Fig. 4. Normalised PDF of domestic home scores plotted against a range of fitted PDFs. Straight lines indicate where the points would lie were the fits perfect, and are separated by an arbitrary vertical displacement; dashed lines indicate 95% binomial counting errors. Note how the Fréchet $a = 1.04$ PDF (e) provides a superior fit to the gumbel $a = 1, 2$ (c,d), negative binomial (b), and Poisson (a) distributions; compare Fig. 1.

clarity; typical sizes of counting errors are indicated by fluctuations around a smooth trend and become apparent in the final few bins.

We plot the away team scores in Fig. 2. Again, the domestic scores are consistent with a Fréchet distribution above $\mu + 4\sigma$ (an away score of about 6 goals) whereas negative binomial PDFs suffice for the league and Cup scores if the last few points are discounted as explained above. The total goal scores with fitted negative binomial PDFs are plotted in Fig. 3. Here we find that the domestic scores are consistent with a Gumbel distribution (see Section 3) above $\mu + 3\sigma$ (9 goals), and that the league scores are more suggestive of a Poisson than a negative binomial PDF.

We now provide more detailed analysis of the goodness of fit of the extremal PDFs to the domestic scores. Figs. 4–6 show in linear form the closeness of fit of various distributions to the domestic data. To quantify whether the data are consistent with the fitted PDFs, one must estimate the likely counting errors in the numbers of points in the bins (introduced by the finite size of the dataset). We are interested in the distribution of the number of points in a bin, given both the total number of data points and the probability that any point will lie in that particular bin (given directly by the fitted PDF). A binomial distribution of counting errors is a reasonable estimate, and one can thus estimate the upper and lower limits of the number of points one

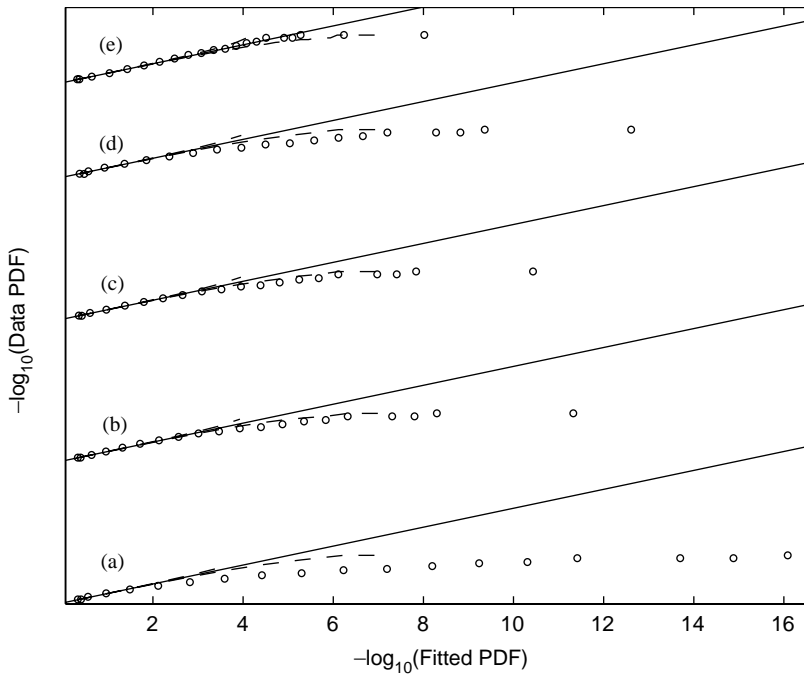


Fig. 5. Normalised PDF of domestic away scores plotted against a range of fitted PDFs. Straight lines indicate where the points would lie were the fits perfect, and are separated by an arbitrary vertical displacement; dashed lines indicate 95% binomial counting errors. Note how the Fréchet $a = 1.10$ PDF (e) provides a superior fit to the gumbel $a = 1, 2$ (c,d), negative binomial (b), and Poisson (a) distributions; compare Fig. 2.

would expect to find in any bin on 95% of occasions; these are plotted with dashed lines. The ending of a dashed line indicates that the lower/upper limit of the expected number of occurrences of the corresponding score is zero; where both limits stop, no higher scores are expected given the size of the dataset. From these plots we again conclude that a Fréchet $a = 1.04$ PDF is the best fit to domestic home scores, a Fréchet $a = 1.10$ PDF is the best fit to domestic away scores, and a Gumbel $a = 1$ PDF is the best fit to domestic total scores.

The empirical PDF varies between the negative binomial and extremal PDF in each case; for low scores both the negative binomial and extremal distributions provide satisfactory fits. However, as shown in the previous figures, there is a strong departure from the negative binomial on to heavier-tailed distributions for the higher scores. Our aim here is to identify a single distribution that fits the whole dataset rather than an arbitrary piecewise fit. The latter could always be achieved given a sufficient number of independent distribution functions to fit to different ranges of data, but would ultimately be less informative of the underlying processes.

In this context, it is important to note that the distribution of the aggregate of many thin-tailed datasets (i.e., the pooled data) is heavy-tailed if the variances of the

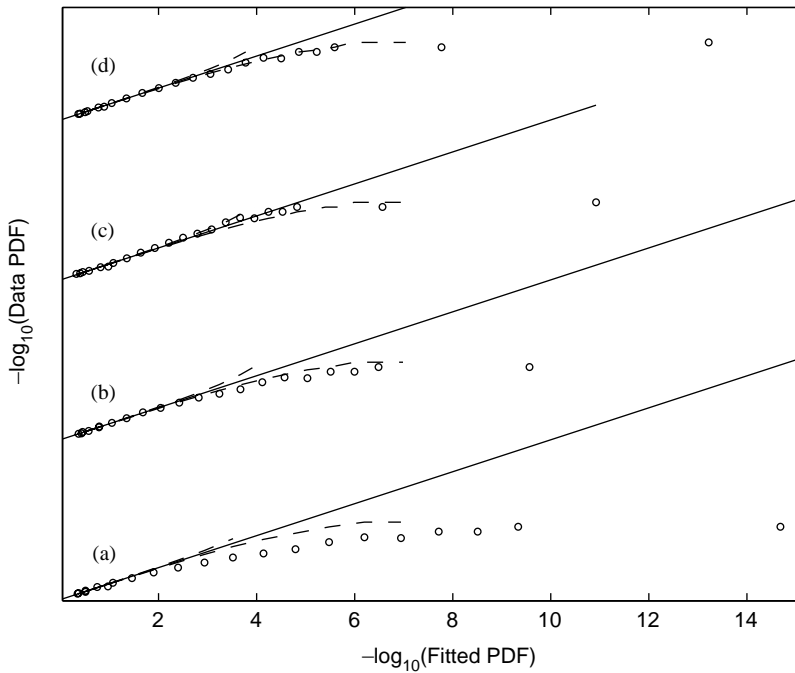


Fig. 6. Normalised PDF of domestic total scores against a range of fitted PDFs. Straight lines indicate where the points would lie were the fits perfect, and are separated by an arbitrary vertical displacement; dashed lines indicate 95% binomial counting errors. Note how the Gumbel $a=1$ PDF (c) provides a superior fit to the gumbel $a=2$ (d), negative binomial (b), and Poisson (a) distributions; compare Fig. 3.

component datasets differ [29,30]. Hence the heavy-tails seen in worldwide football results could arise simply from the aggregation of scores from many teams. Individual teams' scores may follow different Poisson distributions, which when pooled produce countries' scores following negative binomials, and then the aggregation of countries' scores is heavy-tailed. Testing this hypothesis would require significantly more data than used here, and would run over an interval of time that may imply significant changes in the game process. The alternative—and more interesting—possibility is that the heavy tails are the result of some inherent process that increases the likelihood of high scores over their Poisson-based expectations.

We also find that both the English data and the worldwide domestic results show a mean goal difference (home score minus away score) of 0.51, an aggregate home advantage (see Ref. [31]), and a bias towards uneven scores as the total score rises is evident in the larger domestic dataset; these trends are well-known in the world of football.

It is important to note that the observation of a departure from negative binomial distributions is not the result of a larger dataset for domestic matches. Whilst more rare events are observed in a larger sample and the distribution extends to higher values with lower probabilities, it is nevertheless possible to distinguish between the different

distributions, as we have shown, without considering these extreme values. We have looked briefly at other individual countries and find similar trends to those shown for English matches.

5. Conclusions

We have shown that the simplest models—the thin-tailed Poisson and negative binomial distributions based on the assumption of uncorrelated processes—do not fit domestic (worldwide) football matches between 1999 and 2001 beyond the low scores. Heavier-tailed distributions are required if these datasets are to be fitted with single PDFs. Log-normal distributions do not include zero scores whereas extremal distributions can model the entire range of scores. Extremal distributions have been observed in a variety of complex systems and our results may then inform the modelling of football games.

In addition, using English top division and FA Cup matches in the seasons of 1970/71–2000/01, we confirm the Poisson or negative binomial nature of English scores as reported in analyses of earlier football seasons.

Acknowledgements

We are grateful to Nick Watkins and Richard Dendy for helpful discussions. J.G. and P.C.B. acknowledge Research Studentships from the UK Particle Physics and Astronomy Research Council. Data provided by members of the Recreational Sport Soccer Statistics Foundation and SportingData.com.

References

- [1] M.J. Moroney, *Facts from Figures*, Penguin Books, London, 1956.
- [2] C. Reep, R. Pollard, B. Benjamin, *J. Roy. Statist. Soc. Ser. A* 134 (1971) 623.
- [3] M.J. Maher, *Statist. Neerlandica* 36 (1982) 109.
- [4] M.J. Dixon, S.G. Coles, *Appl. Statist.* 46 (1997) 265.
- [5] D. Dyte, S.R. Clarke, *J. Opl. Res. Soc.* 51 (2000) 993.
- [6] M.J. Dixon, M.E. Robinson, *The Statistician* 47 (3) (1998) 523.
- [7] J.M. Norman, in: J. Bennett (Ed.), *Statistics in Sport*, Arnold, London, 1998.
- [8] L.C. Malacarne, R.S. Mendes, *Physica A* 286 (2000) 391.
- [9] <http://www.SportingData.com>.
- [10] <http://www.rssf.com/histdom.html>.
- [11] P. Bak, *How Nature Works: The Science of Self-Organised Criticality*, Oxford University Press, Oxford, 1997.
- [12] C. Tsallis, S.V.F. Levy, A.M.C. Souza, R. Maynard, *Phys. Rev. Lett.* 75 (20) (1995) 3589.
- [13] C. Tsallis, *Phys. World* 10 (1997) 42.
- [14] R.N. Mantegna, H.E. Stanley, *An Introduction to Econophysics: Correlations and Complexity in Finance*, Cambridge University Press, Cambridge, 1999.
- [15] H.E. Stanley, L.A.N. Amaral, X. Gabaix, P. Gopikrishnan, V. Plerou, *Physica A* 299 (1–2) (2001) 1.
- [16] H.E. Stanley, L.A.N. Amaral, X. Gabaix, P. Gopikrishnan, V. Plerou, *Physica A* 302 (1–4) (2001) 126.
- [17] S. Solomon, P. Richmond, *Physica A* 299 (1–2) (2001) 188.

- [18] S.T. Bramwell, P.C.W. Holdsworth, J.-F. Pinton, *Nature* 396 (1998) 552.
- [19] S.T. Bramwell, K. Christensen, J.-Y. Fortin, et al., *Phys. Rev. Lett.* 84 (17) (2000) 3744.
- [20] V. Aji, N. Goldenfeld, *Phys. Rev. Lett.* 86 (6) (2001) 1007.
- [21] S.T. Bramwell, J.-Y. Fortin, P.C.W. Holdsworth, et al., *Phys. Rev. E* 63 (2001) 41 106.
- [22] B. Zheng, S. Trimper, *Phys. Rev. Lett.* 87 (18) (2001) 188 901.
- [23] S.T. Bramwell, K. Christensen, J.-Y. Fortin, et al., *Phys. Rev. Lett.* 87 (18) (2001) 188 902.
- [24] R.A. Fisher, L.H.C. Tippett, *Proc. Cambridge Philos. Soc.* XXIV (1928) 180.
- [25] E.J. Gumbel, *Statistics of Extremes*, Columbia University Press, New York, 1958.
- [26] D. Sornette, *Critical Phenomena in Natural Sciences*, Springer, Berlin, 2000.
- [27] K. Bury, *Statistical Distributions in Engineering*, Cambridge University Press, Cambridge, 1999.
- [28] S.C. Chapman, G. Rowlands, N.W. Watkins, *Nonlinear Processes in Geophysics*, 2002, in press.
- [29] C.W. Granger, D. Orr, *J. Am. Statist. Assoc.* 67 (1972) 275.
- [30] S.J. Kon, *J. Finance* 39 (1984) 147.
- [31] S.R. Clarke, J.M. Norman, *The Statistician* 44 (4) (1995) 509.