Competitive Balance in the Portuguese premier league of professional soccer

António Marques
amarques@alpha2.ubi.pt

Abstract

Competitive balance is what makes the sport becoming an attractive event. Contrary to other industries, a reduction in competition and a hypothetically dominant position could mean the end of the championship. If there is only one team, the criterion for league cannot be met. In this paper we consider the advantage of the team that plays at home and, using an ordered probit model, the strength for all teams is estimated. At the same time, is made an evaluation on the evolution of the competitive balance in the Portuguese premier league of professional soccer. The conclusions point that this balance has increased, but also the difference between the best team and the worst one.

Keywords: Competitive Balance, Soccer, Ordered Probit
JEL: L19, L83

I- INTRODUCTION

In the middle of the 20th century the interest in the study of professional sporting clubs started to increase. The founding articles were: Rottenberg (1956)[1], The baseball players labor market, Journal of Political Economy; Neale (1964), The peculiar economics of professional sports, Quarterly Journal of Economics and Sloane(1971), The economics of professional football: the football club as a utility maximiser, Scottish Journal of Political Economy.

A professional sporting club needs to keep a competitive level which guarantees its continuity in the future. In the other sectors companies might prosper with the elimination of the competitors, trying to stay alone in the market, and thus becomes a monopoly. The uncertainty of the outcome of a game attracts a larger number of spectators (the demand), and reduces the dominant position of a particular team in a championship. This happens because a single club cannot satisfy all the sports consumers. So, all the teams must cooperate between themselves, in order to constitute a competitive championship. Based on this joint endeavor, Dobson and Goddard (2001a:6) pointed out that a championship will have to be regarded as a company which produces on several plants, according to the decisions that are taken and implemented collectively in this context of championship. Thus, the clubs maximize profits, while the objective of a championship is to maximize the demand of this sport and the joint profits of the clubs (Palomina and Risotti, 2000).

The competitive balance in a championship, is characterized by the uncertainty of the result. It is absolutely necessary in order to guarantee that the adepts look for the sporting events. So, the more difficult is to foresee the result of a game, the higher will be number of spectators. If the teams are equally likely to win, then the championship will be more attractive. Considering the advantage of playing at home, this text estimates the quality of each team as a parameter. After this, through the dispersion of these estimated parameters, we analyze the evolution of the competitive balance in the Portuguese premier league of professional soccer. We also calculated the probability of the team that plays at home to win when it plays against other team with an equal

[1] *Cf* Dobson and Goddard (2001a)

strength. Finally, we present, for each season, the team with a higher strength and a lower strength. The econometrical Package used is the Limdep.

In section two there is a brief summary of texts produced on this theme which constitutes theoretical support for the use of the considered model. Section three is reserved for the presentation of the methodology, the specification of the model estimated and the method of estimation. On the fourth section, the empirical results are presented. The first point of the 4th section is the performance of teams by season. This is followed by an analyze of the competitive balance in the Portuguese premier league of the professional soccer. This paper finishes with the conclusions, which were obtained, resulting ideas and thus limitations and suggestions for future works.

II - COMPETITIVE BALANCE

The economic literature there are not many studies where there are models that give the results of the soccer games, although there is an increasing interest in the area. Dobson and Gooddard (2001a:170) presents the most important contribution, pointing out the difference in methodologies. The first works of Moroney (1956) and Reep, Polard and Banjamin (1971) use the binomial negative and Poisson distributions in the models of the number of goals by game. Maher (1982) considered that the points of the teams when they play at home and when they are a visiting team follows an independent Poisson distribution. The model, however, does not have to have forecasting capacity before points are won on the end of the match. Dixon and Coles (1997) retake the model of Maher (1982), in developing a forecast model, capable of generating, ex-ante, the probabilities of the possible results of the matches, through direct adjustment in order to allow the interdependence of the marginal Poisson probabilities. Rue and Salvessen (2000), using a similar formalization to the Dixon and Coles one have concentrated on the estimation of the behavior parameters of the defense and the attack of each team. García and Rodríguez (2002), using panel data model, estimate for the Spanish soccer league price elasticities. They also measure the contribution of each group by explaining factors for the demands for a game, concluding that ex-ante variables are related with to the quality of both teams are the ones with the highest explaining power.

There were many studies in the nineties. They had tried to evaluate the impact of the introduction of specific factors for the explanation of the results of games, such as the expulsion of players in the game, the percentage of ball possession time or the number of spectators and the claque charisma.

One question that recently has motivated more works is the analysis of the impact of the rights and revenues of media broadcasting transmission in the competitive balance of championships. The transmission rights contracts have been the target of a lot of attention, where the clubs seek to maximize their revenues in or out of the stadium, because these contracts allow them to increase direct revenues and bring the game to the consumer. At the same time the staffs responsible for the championship try to make the form of these contracts negotiated in order to prevent diminishing the competitiveness in championships. This theme was developed by El Hodri and Quirk (1971), with their modeling of the seminal work of Rottemberg, and Fort and Quirk (1995), Vrooman (1995) and Eckard (1998)[2]*. Szymanski (1998:15) concluded that the share of stadium revenues tends to influence negatively the competitive balance in championship while the share of television revenues stimulates this competitive balance. The reason is that television revenues are negotiated collectively, and are generated independently the teams' performance.*

Szymanski (2000), refers that, in England, there exists little evidence that the income concentration is associated with the decline in competitive balance. Moreover there is little evidence that this increase in the inequality of the team's wealth reduces interest in soccer championship. It concludes, therefore, that competitive balance is sensitive only to a great change in income distribution. Dobson and Goodard (2001b) observe that the clubs of big cities have proved an increase in its base of demand relatively to small cities. Thus, it is possible to verify divergences in the sharing of revenues.

Köning (2000) modeled the results of games, but, rather than using the points of the teams individually, he used an estimation of an ordered probit model. It follows the research line of Neumann and Tamura (1996) that, using the strength of each team as a parameter to measure the competitive balance for the dispersion of estimated parameters.

Köning (2000) considered that one soccer league has perfect competitive balance during one year if the probability of that team to win a game when it plays at home is constant and, therefore, does not vary with the adversary team.

In the present work we follow this research line studying the competitive balance in the Portuguese premier league of professional soccer.

[2] Authors cited by Dobson and Goodard (2001b).

III - THE MODEL

Let us consider the existence of an advantage for a team when it plays at home. This can result from the best knowledge of the field, from the supporters' charisma or yet from the physical saving not have to dislocate the players from the other stadium. Let us consider Y_{ij}*, a variable that represents the match result between team* i*, when it plays at home, and team* j*, the visiting team. It can assume values the 0, 1 or 2, depending if the home team loses, draws or win, respectively. For the preferences of the clubs the victory strict dominates draw and this dominates strictly loses. It is possible establish a hierarchy in the results, that constitutes an incentive for the use of the ordered probit model. The variable* Y_{ij}^**, with* $Y_{ij}^* = a_i - a_j + e_{ij}$*, is a latent variable, that determines the game result. If* Y_{ij}^* *is positive, the quality of the team that plays at home is superior to the visitor team, with* e_{ij} *representing the random element, the luck, or unexpected element of the result of the game. Thus:*

$$Y_{ij} = \begin{cases} 0 & Y_{ij}^* \leq m_1 \\ 1 & m_1 < Y_{ij}^* \leq m_2 \\ 2 & Y_{ij}^* \leq m_2 \end{cases}$$

The model admits a latent a_i *for each team, which represents its quality. So, it represents the heterogeneity among the various teams. This parameter is independent on the adversary and where the game takes place, with the assumption that it is constant for each season. The talent differences* $a_i - a_j$ *influence the probability of the results* Y_{ij}*, through a response function F that will be considered the normal distribution function. The probabilities of the possible results of the game are (see annexes 1):*

$$Prob(Y_{ij} = 0) = \Phi(m_1 - a_i + a_j);$$
$$Prob(Y_{ij} = 1) = \Phi(m_2 - a_i + a_j) - \Phi(m_1 - a_i + a_j);$$
$$Prob(Y_{ij} = 2) = 1 - \Phi(m_2 - a_i + a_j)$$

The estimation of the limit parameters m_1 *and* m_2 *and* a*'s is made by the maximization of the likelihood function, which is the product result of the individual contributions to all the games. The model, however, is not identifiable, because only the quality differences are considered in the likelihood function. In adding a* *constant* $a_i, i = 1,...n$ *, it will not modify the likelihood function. It is needed, therefore, to introduce an identification restriction*[3]*, which is* $\sum_{i=1}^{n} a_i = 0$. *The quality of the teams is calculated simultaneously, and therefore, the boarding adjusts automatically the quality of the adversary team and the advantage of playing at home. The parameters* a *can be interpreted as the distance in relation to one hypothetical average team with quality zero (Köning, 2000) and, therefore, a negative value of* a_i *implies that the talent of team i is inferior than the average team. The maximization of the likelihood function, in order to obtain the estimative of the parameters, follows:*

$$LogL = \sum_{t} \sum_{(i,j) \in x_t} \{ I_{(Y_{ij}=2)} \log[1 - \Phi(m_2 - a_i + a_j)] + I_{(Y_{ij}=1)} \log[\Phi(m_2 - a_i + a_j) - \Phi(m_1 - a_i + a_j)] + I_{(Y_{ij}=0)} \log \Phi(m_1 - a_i + a_j) \},$$

where t *represents the season and* x_t *the teams set that plays in the first league in the season* t *. The used information considers all the soccer games in the championship of the first league in Portugal, between the seasons 1960_61 and 2000_01*[4]*. Until 1970_71 fourteen teams had participated in the first league. After this season this number jumped to sixteen. In the seasons 1987_88 and 88_89 twenty teams had participated. At the season following, eighteen teams and twenty, again, in 90_91. After 1991_92, inclusive, the number stabilized at eighteen teams. Combined two by two and taking into consideration the fact that each team plays two times per season in different fields, the total number of considered games is 10029. In all seasons, fifty-four teams had participated.*

In an exploratory data analysis, we can observe a trend for the reduction of the average performance of teams. At the same time we observe an improvement in performance of teams when they play away, basically after the beginning of the nineteen-nineties (cf table 1). These data follows the long term international trend, that is, reduction in the advantage to play at home, and the last ten seasons had

Table 1 – Wins, draws and loses

Seasons	*Home wins (%)*	*Away team wins (%)*	*Draws (%)*
!960_61 a 1969_70	0,547	0,258	0,195
1970_71 a 1079_80	0,541	0,218	0,241
1980_81 a 1989_90	0,539	0,184	0,277
1990_91 a 2000_01	0,517	0,220	0,263

[3] The invariance property of the maximum likelihood estimator make sure that the estimated team's quality is equivalent, wichever it may be the used restriction (Cox e Hinkley, 1974).
[4] From the time 1999_00, inclusive, tThe first division has been called first professional soccer league.

accentuated the rhythm of this trend. The Spanish, German, French and English championships had verified, however, a more accentuated decreasing (cf Dobson and Goddard, 2001a:152).

IV - EMPIRICAL RESULTS

We present, in the first instance, the results from the teams' performance by season. In the second instance, we conclude on competitive balance.

IV-I - Performance per Season

Table 2 synthesizes the reached results of the estimative made to each season. In the third and fifth columns it is presented the victory probability of one team with the quality a *, when it plays with a hypothetical team with zero quality. This probability is gotten using* $1-\Phi(m_2-a)$ [5]. *The variation in these probabilities does not represent changes in home advantage but the variation in quality (Köning, 2000), given we compare victory probabilities.*

Table 2 – Team with bigger and smaller quality, by season

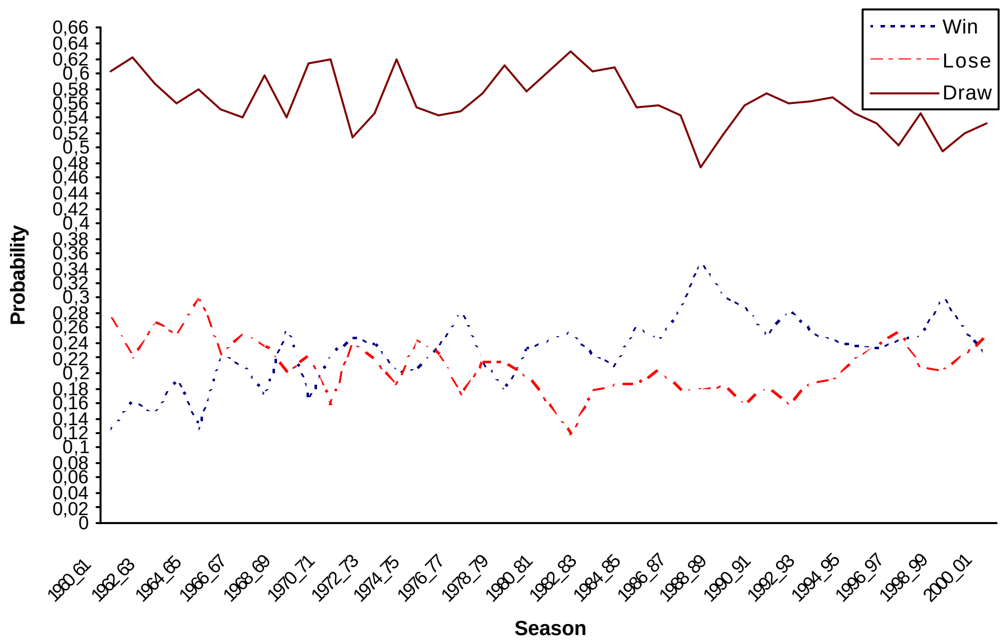
Season	Worst team	$1-\Phi(m_2-a)$	Best team	$1-\Phi(m_2-a)$	Standard-error
1960_61	FCB	0,085542	FCP	0,422649	0,3316
1961_62	SCC	0,087420	SLB	0,417298	0,2890
1962_63	CDF	0,067149	SLB	0,502088	0,3682
1963_64	SFC	0,117379	SLB	0,454170	0,2766
1964_65	SCUT	0,098276	AAC	0,391392	0,2719
1965_66	SCBM	0,120526	VFC	0,350373	0,2435
1966_67	VFC	0,160910	SCP	0,327818	0,2094
1967_68	LSC	0,156510	SCP	0,384719	0,1886
1968_69	CUF	0,118267	VFC	0,423538	0,3122
1969_70	AAC	0,099367	SCP	0,390353	0,3150
1970_71	EVEC	0,110998	FCP	0,390186	0,2772
1971_72	FCT	0,135175	SLB	0,469430	0,2356
1972_73	SCP	0,130826	SLB	0,434491	0,2568
1973_74	CUF	0,163602	VFC	0,482967	0,2424
1974_75	SCF	0,129669	FCP	0,421214	0,3039
1975_76	AAC	0,131029	SLB	0,481168	0,2707
1976_77	ACP	0,107965	SLB	0,394478	0,2554
1977_78	SCE	0,128937	SCP	0,399542	0,2486
1978_79	CAF	0,112515	SLB	0,376688	0,2491

[5]For the boundary-value m_2 we considered the estimated m_2 average in the 41 seasons (0,7491). By this way we caught the esteem boundary-values in each season.

1979_80	SCE	0,128706	BFC	0,378664	0,2534
1980_81	EVEC	0,072546	FCP	0,410886	0,3199
1981_82	VSC	0,118601	SCP	0,471063	0,2922
1982_83	RAFC	0,151048	SLB	0,341257	0,1739
1983_84	SCE	0,082510	SLB	0,455205	0,2765
1984_85	VSC	0,097507	FCP	0,435811	0,2756
1985_86	VFC	0,140848	SLB	0,354114	0,2258
1986_87	PSC	0,119777	FCP	0,370923	0,2024
1987_88	SCC	0,145216	PSC	0,306582	0,1634
1988_89	SCBM	0,148933	SLB	0,382523	0,2035
1989_90	PSC	0,128713	FCP	0,411423	0,2084
1990_91	SCB	0,134796	SLB	0,500734	0,3019
1991_92	SCBM	0,137005	VSC	0,420249	0,1985
1992_93	GVFC	0,118494	FCP	0,394146	0,2200
1993_94	SCB	0,118010	SLB	0,397856	0,2572
1994_95	UDL	0,130244	FCP	0,466824	0,2622
1995_96	SCS	0,140226	SLB	0,387403	0,1843
1996_97	VSC	0,143857	FCP	0,479095	0,2540
1997_98	CFB	0,070736	VSC	0,338438	0,2728
1998_99	FCA	0,133321	SLB	0,363551	0,1768
1999_00	SCB	0,153648	SCP	0,333028	0,1580
2000_01	FCA	0,120401	FCP	0,369669	0,2226

As it would be expected and paying attention to the championship history of the premier league in Portugal, in forty-one seasons only eight teams had reached the best set. Of these, only three (SLB, FCP and SCP) had been, in thirty-three seasons, the best soccer

Graph 1 – Victory probability of the home team when plays with a team of the same quality



team, that is, with the highest quality. We still reveal the evidence that in only 46% of the cases the best team won effectively the championship.

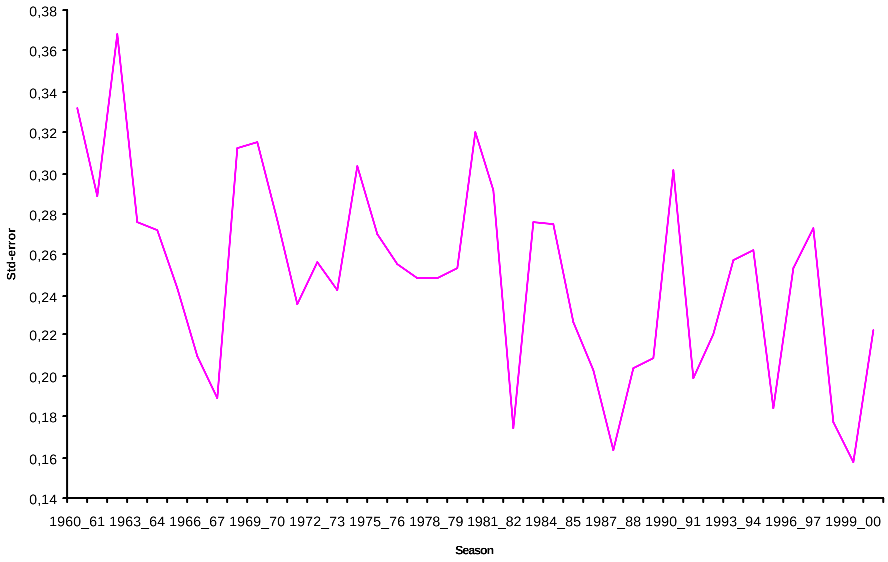
Admitting a context where the team that plays at home and the visitor team have the same quality, it is possible to calculate the victory probability of the team that plays at home. When we consider the talent of the teams being equal, then we separate the effect of the advantage of playing at home. The advantage of playing at home is presented in graph 1. It is calculated by $a_i = a_j$, *then*[6] $Prob(Y_{ij} = 2) = Prob(e_{ij} > m_2) = 1 - \Phi(m_2)$. *The losses probability is* $Prob(Y_{ij} = 0) = Prob(e_{ij} \leq m_1) = \Phi(m_1)$. *The victory probability of playing at home in the Portuguese championship has different behavior of the results estimated by Köning (2000) for the Dutch championship. During the sixties this probability was less than the probability of loosing. The reason could be the hegemony of few teams that, independently on the field where they played, won the matches. We observe that, during the seventies, when teams of equal quality played one against the other, the result was not influenced by the field where they made it. In the eighties, the victory probability of playing at home increased, but in the middle of the nineties we observe again the approximation of victory and losses probabilities. For this contributed the new methods of trainings, gradually accessible to all teams, because the players were strengthened, allowing them to have better performances when they playing away (Dobson and Goddard, 2001a:150). Besides that the psychological factor of playing at home lost importance. At the same time, the improvement of the ways of communication and transports diminishes the fatigue of the players in their displacement.*

IV-II Competitive Balance

[6] We reject the hypothesis $m_1 = m_2 = 0$ at a 5% significance level.

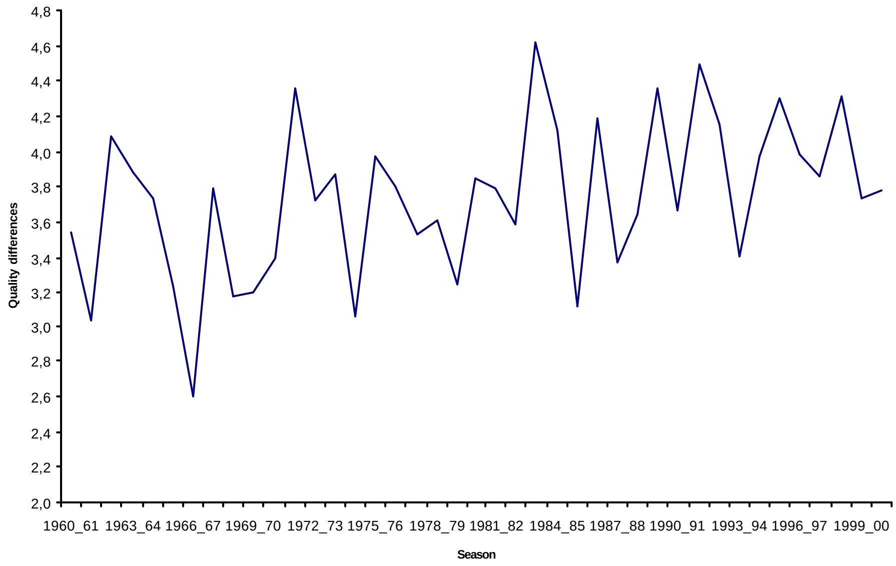
For the assessment of the competitive balance in the championships we can use several measures. In a similar way, the factors studied that will be able to influence this competitive balance are numerous. The concentration ratios, instability indexes, Gini's coefficient, Lorenz's curve, Auto-Regressive processes are examples of possible measures. In this work we consider only the standard-errors of the a *estimates in each season as well as the discrepancies of quality of the teams in each season. We imposed*

Graph 2 – Standard-Error of the estimated parameters



the standard normalization from the parameters estimated and after the difference was observed between a's. *The bigger the* a, *the bigger the heterogeneity of the teams in each season. Graph 2 shows the trend to decrease standard-error of estimated* a's *across considered seasons. This is a signal of the improvement concerning competitive balance. However, this trend is not homogeneous during all the concerning period. The increase of the competitive balance was superior from those of other seasons. In the sixties, the increasing trend of the balance is easily shown. The decade of the 70's discloses lesser levels of competitive balance which are approximately constant. From the 80's on it is observed the normality of the decreasing trend of the standard-error, and, therefore, the enhancement of competitiveness.*

Graph 3 – Quality differences by season



It was also verified the evolution of the differences of the team's strength. So, after the standardization of the estimated parameters, we constructed Graph 3, where we can observe that, in the same period, the lag between the highest and lowest quality teams in Portuguese championship of professional soccer has increased.

V - CONCLUSIONS

The understanding of the competitive balance in sport is an emerging area that has been the object of an increasing number of articles. These researches attend several explaining factors as well as they use several methodologies. In this work, we use an ordered probit model that has been used by Neumann and Tamura (1996) and it was firstly developed by Köning (2000). It is possible to conclude that in the Portuguese premier league of professional soccer, between the 1960 and 2001 seasons, the most talented teams were often the same. For example, only three teams were, in 80,5% of the seasons the best teams. The competitive balance has increased, even if the trend is not homogeneous during the whole period. At the same time, the difference between the teams with greater strength and those with less talent has also increased. This can be the reflex of the different resources that each team has access to.

It was also possible to conclude that, in contrast with other leagues of professional soccer such as Dutch's league, in Portugal when two teams with equal talent play, the result is not significantly influenced by the field where there play, that is, the home advantage in not very significant in the period of this analysis. That is equivalent of *saying that the victories must result from quality and not from the field where the matches are played which is the result of greater competitiveness in the Portuguese championship.*

The simplicity, which also is evident in this model, implies some limitations. The modeling of the soccer games results should consider the inclusion of other factors such as discipline, the efficiency of the attack and defense, the stability in clubs, the incentives from supporters and the moment of the season when the game is played. In future works we wish to consider the comparison of the competition levels, not only inside a championship, but also between the teams of different championships, in order to allow us to draw a conclusion on regarding the evolution of the competitive level in different countries. This is the work for the next study.

Annexes

Annex 1 - ***List of teams***

ABBREVIATION	TEAM	ABBREVIATION	TEAM
AAC	Académica de Coimbra	*FCV*	Vizela
ACP	Atlético	*GCA*	Alcobaça
ADF	Fafe	*GDC*	Chaves
ADS	Sanjoanense	*GDEP*	Estoril
AFC	Amora	*GDR*	Riopele
BFC	Boavista	*GVFC*	Gil Vicente
CAD	Elvas	*LFC*	Leça
CAF	Académico de Viseu	*LGC*	Lusitano de Évora
CDA	Aves	*LSC*	Leixões
CDF	Feirense	*PSC*	Portimonense
CDM	Montijo	*RAFC*	Rio Ave
CDN	Nacional	*RDA*	Águeda
CDSC	Santa Clara	*SCB*	Braga
CFB	Belenenses	*SCBM*	Beira-Mar
CFEA	Estrela da Amadora	*SCCA*	Campomaiorense
CFU	União da Madeira	*SCC*	Covilhã
CFUC	União de Coimbra	*SCE*	Espinho
COL	Oriental	*SCF*	Farense
CDM	Montijo	*SCO*	Olhanense
CSM	Marítimo	*SCP*	Sporting
CUF	CUF	*SCS*	Salgueiros
EVEC	Varzim	*SCUT*	Torreense
FCA	Alverca	*SFC*	Seixal
FCB	Barreirense	*SLB*	Benfica
FCF	Famalicão	*UDL*	Leiria
FCP	Porto	*UFCIT*	União Tomar
FCPF	Paços de Ferreira	*VFC*	Vitória de Setúbal
FCPI	Penafiel	*VSC*	Guimarães
FCT	Tirsense		

Annex 2 - ***Calculation of the results probabilities:***

LOSES :

$$Prob(Y_{ij}=0)=Prob(Y_{ij}^{*}\leq m_1)=Prob(a_i-a_j+e_{ij}\leq m_1)=Prob(e_{ij}\leq m_1-a_i+a_j)$$
$$=\Phi(m_1-a_i+a_j)$$

DRAW:

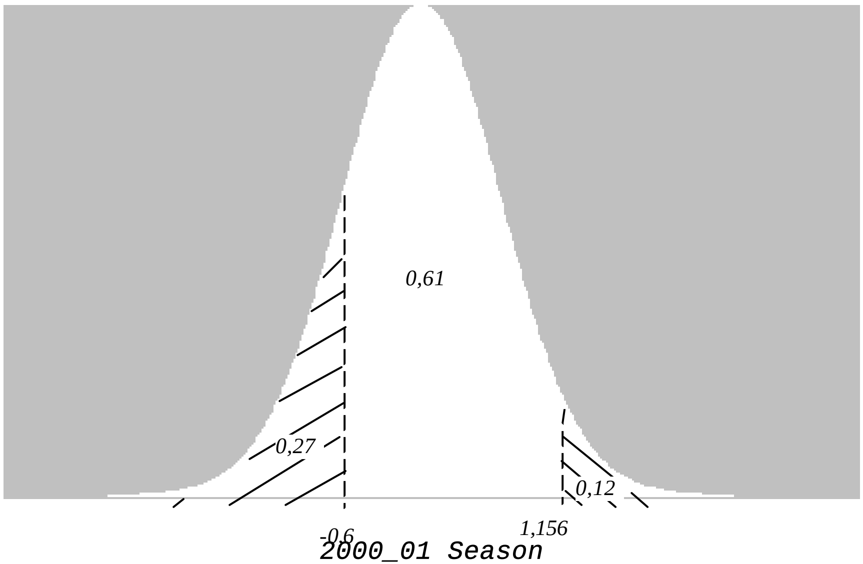
$$Prob(Y_{ij}=1)=Prob(m_1<Y_{ij}^*\le m_2)=Prob(m_1<a_i-a_j+e_{ij}\le m_2)$$
$$=Prob(a_i-a_j+e_{ij}\le m_2)-Prob(a_i-a_j+e_{ij}<m_1)$$
$$=Prob(e_{ij}\le m_2-a_i+a_j)-Prob(e_{ij}<m_1-a_i+a_j)$$
$$=\Phi(m_2-a_i+a_j)-\Phi(m_1-a_i+a_j)$$

WIN :

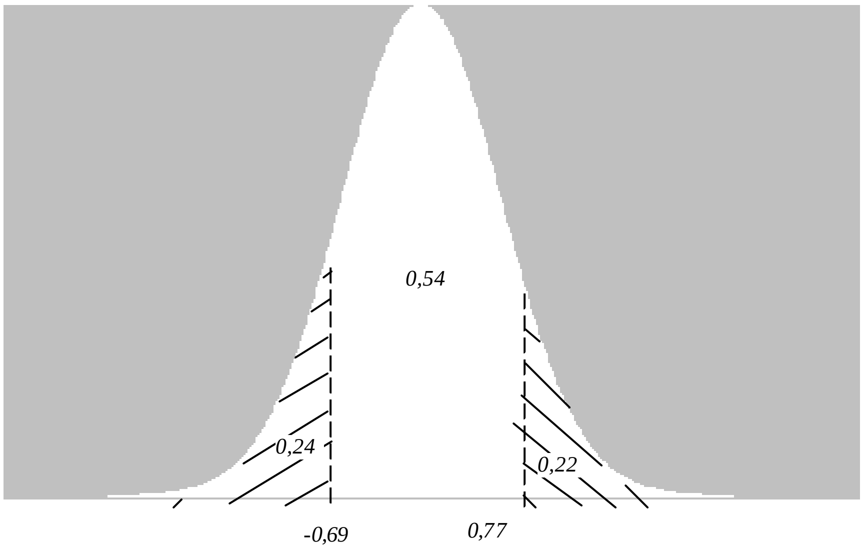
$$Prob(Y_{ij}=2)=Prob(Y_{ij}^*>m_2)=Prob(a_i-a_j+e_{ij}>m_2)=Prob(e_{ij}>m_2-a_i+a_j)$$
$$=1-\Phi(m_2-a_i+a_j)$$

Annex3 - ***Normal Standardized***

1960_61 Season



2000_01 Season



REFERENCES

COX, D. R. and HINKLEY, D. V. (1974), Theoretical Statistics, London: Chapman and Hall.

DOBSON, S. and GODDARD, J (2001a) The Economics of Football, Cambridge University Press.

DOBSON, S. and GODDARD, J (2001b) Revenue divergence and competitive balance in a professional sports league, Department of Economics at Trinity College, Dublin Economic Workshop http://econserv2.bess.tcd.ie/DEW/qub1dobson.doc

GARCÍA, J. and RODRÍGUEZ, P. (2002), The determinants of football match attendance revisited: empirical evidence from the Spanish football league, Journal of sports economics, vol3, 1, pp. 18-38.

GREENE, W. H. (2000), Econometric Analysis, fourth edition, Prentice-Hall Inc., Upper Sadle River, New Jersey.

HAAN, M., KONING, R.H. and WITTELOOSTUIJN, A.V. (2002), Market Forces in European Soccer, SOM Research Report 02F18.

KNORR-HELD, L. (1999), Dynamic rating of sports teams, Discussion Paper 98, Institut für Statistik, München, http://www.stat.uni-muenchen.de/sfb386/papers/dp1997.html

KONING, R.H. (1999), Competitive Balance in Dutch Soccer, SOM Research Report 99B04, Faculty of Economics, University of Groningen.

HAAN, M.A., KONING, R.H. and WITTELOOSTUIJN A. van (2002), Market Forces in European Soccer, SOM Research Report 02F18, http://www.rhkoning.com/publications/2002/market-forces-som.pdf

KONING, R.H. (2000), Balance in competition in Dutch Soccer, *The Statistician*, 49(3), pp.419-431.

NEUMANN, G. R. and TAMURA, R. F. (1996), Managing competition: the case of the national football league, Manuscript, Clemson University, http://hubcap.clemson.edu/~sixmile/sports/tamura.htm

PALOMINO, F. and RIGOTTI, L. (2000), The sport league's dilemma: competitive balance versus incentives to win, WPS, E00-292, University of California at Berkeley.

PALOMINO, F. and SAKOVICS, J. (2001), Inter-league competition for talent *vs* competitive balance, WP, http://cwis.kub.nl/~few5/center/staff/palomino/comp0601.pdf

ROSEN, S. (1981), The economics of Superstars, *American Economic Review*, 71, pp 845-858.

SZYMANSKI, S. (1998), Hermetic leagues, open leagues and the impact of revenue sharing on competitive balance, Imperial College, Management School, http://www.ms.ic.ac.uk/stefan/CB99.pdf

SZYMANSKI, S. (2000), Income inequality, competitive balance and the attractiveness of team sports: some evidence and a natural experiment from English soccer, forthcoming in the *Economic Journal* Feature on the economics of team sports.

SZYMANSKI, S. (2001), Competitive Balance and Income Redistribution in Teams Sports, Royal Economic Society, University of Warwick, http://www.warwick.ac.uk/res2002/papers/Szymanski.pdf