ORGANISING SCHEME OF NATIONAL CHAMPIONSHIPS OF FOOTBALL

Gjergji Pano

Email: *georgefefa@gmail.com*

(Accepted 5 October 2021)

https://doi.org/10.33598/V8I2202111

Abstract

There are many factors that influence the positive performance of football. In this aspect, the organizational scheme of this major activity cannot and should not be overlooked, which in itself does not require any expenditure. The methodology of this study consists of using the applied probability, theoretical and statistical, to discover the conditions that are **generally expected** in the performance of the championship and, based on this general lawfulness, to formulate even the most optimal rules for the way of development of this activity. This is precisely the main goal of this study: to find out an organizational scheme that is most effective for national football championships, in different countries, according to the specifics of each.

Keywords: football, championship, scheme, countries

Introduction

In every football league, in the Super League (and not only), a certain number of teams 'K', face one another in a district system, with a certain number of stages, 'b'. At the end of the activity, **the team with the best results is declared champion, a certain number of teams, 'Q', win the right to participate in European cups and another number of teams, 'R', are relegated from category.** The above three goals are the motivating elements that affect the competitive spirit among teams during a championship. The purpose of this study is to discover **what is expected** to be the impact of these stimulus elements during the course of a championship with the existing scheme:

- When each of the motivating elements ceases to act?

-When the three ones, cease to act altogether, simultaneously?

-What negative phenomena appear as a result of their inaction?

After that only, we can judge if we HAVE TO improve this scheme and HOW can we do it.

Subject and methods

Let's see what is going to happen during the $n = b(K-1)$ weeks of the activity, from the beginning up to the end, in the $N = \dfrac{bK(K-1)}{2}$ matches of it. If all the N games of the event ended in victory, the sum of points of all the teams together at the end of the event would be:

$3 \times N = \dfrac{3bK(K-1)}{2}$ points, while if all the N games were to draw, then the points of all the teams together would be: $2 \times N = \frac{2bK(K-1)}{2} = bK(K-1)$ points. It is very clear that neither the first, nor the second possibility, there is not much chance of happening. From a statistical study of a six-year period, for 5 European countries, (England, Germany, Spain, Italy, France), countries with the most developed football, it turns out to have an average ratio of D : W(L) = 1 : 3. **So, it is expected that ¾ of all the - N - matches to be wins-loses whilst ¼ of them draws** and consequently, the sum of all team points at the end of the event, with a VERY big advantage, is expected to be:

$$\sum Points = (\frac{3}{4}x3 + \frac{1}{4}x2)xN = \frac{11}{4}x\frac{bk(k-1)}{2} = \frac{11}{8}bk(k-1)$$

points.

It is quite logical to think that the place in the middle of the ranking table would be expected with a number of points in the end of the activity, that can be calculated, with a high degree of certainty as the ratio of the sum of all points accumulated, to the number of teams:

$$P_M = \frac{\sum Pik}{K} = \frac{11bK(K-1)}{8K} = \frac{11}{8}b(K-1) = \frac{11}{8}n$$

points. (1)

Analyzing the above relation, we see that the expectation P_M of points for the middle line V_M is the production of a coefficient $11/8$, with the duration of the activity.

This coefficient $11/8 \cong 4/3$ is none other than the PPG coefficient, (points per game), that you can find in all the statistics of football championships. I am marking it with the symbol X_M.

After that in the complete study, with mathematical methods, I found that $X_1 = 7/3$, while $X_K = 1/3$. These values, compared to the multi-year average, (of over 100 European championships), are convincingly the same in the values $X_1 = 7/3$ and $X_M = 4/3$, while for X_K the discrepancy is unacceptable. The theoretical $X_K = 1/3$ is significantly different from its statistical one, $X_K = 2/3$.

SUCH A PHENOMENON CANNOT BE **IGNORED** **if we want to come to the right conclusions** about what happens during the course of a championship.

Therefore, the PPG coefficients, in the following reasoning, are finally accepted

$X_1 = 7/3$, $X_M = 4/3$ and $X_K = 2/3$, and enable us to calculate **the expected points** of the first, middle and last place in the standings, whether at the end of the activity, or their expected points **UP TO ANY WEEK** of the championship.

In addition, it becomes possible to calculate the **EXPECTED PPG** values for each team in the ranking table, from the first place to the middle one, as well as from the last to the middle place, too! I am not aggravating this summary of my study with the relevant formulas. I'm just showing the values of - X_i - for $K = 18$, in the table below:

Place	1	2	3	4	5	6	7
X_i	2.33	2.23	2.09	1.98	1.86	1.74	1.63
Place	8	9	M	10	11	12	13
X_i	1.51	1.39	1.33	1.29	1.22	1.14	1.06
Place	14	15	16	17	19		
X_i	0.99	0.91	0.83	0.75	0.67		

After that it is possible to calculate the expected points of any team, either at the end of the activity, or up to **whatever week** of it. F. ex.:

For K=20, n=38, the first team, at the end of

the activity will get $P_1 = n \cdot X_1 = 2.33 \cdot 38 \approx 89$ points, while up to the twentieth week it will have got $P_1 = X_1 \times 20 = 2.33 \times 20 \approx 47$ points.

The sixth place: $P_6 = X_6 \cdot 38 = 1.81 \cdot 38 \approx 69$ points, while up to the fifteenth week it is expected to get $P_6 = X_6 \cdot 15 = 1.81 \cdot 15 \approx 27$ points, and so on.

Subject

To make the following reasoning clearer, as a start I'm presenting the progress of a championship, schematically with a rectangle, with sides, numerically, equal to: vertical rib, - the number of participating teams, K, and horizontal rib, - the duration of the activity in weeks, n=b(K-1).

The surface of this rectangle, S = n K = b K (K-1), is numerically equal to twice the total number of championship matches, i.e., 2N (weeks x teams)

Week after week, after the beginning, teams start to differentiate by their results and there comes a moment when, after "e_1" weeks of the activity, **the last place team with the number of points earned up to this week, even if it wins all the remaining matches, it CANNOT REACH the expected points of the first place**. After this moment, the incentive to win the title, **STOPS WORKING**.

We can find out this number from the relation: $\frac{2}{3}e_{1,K} + 3(n - e_{1,K}) < \frac{7}{3}b(K-1)$. After calculations we find:

$$e_{1,K} > \frac{2b(K-1)}{7} \text{ weeks}$$

. In general, **we can calculate** AFTER HOW MANY WEEKS EACH TEAM IN THE PLACE "i" of the ranking table, loses the opportunity towards the title, in the relation:

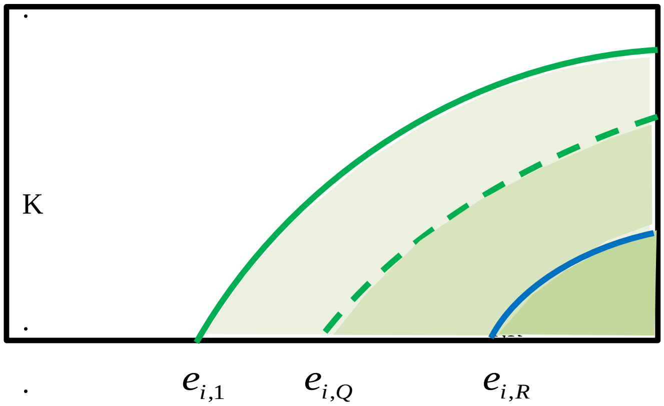
$X_i e_{1,i} + 3(n - e_{1,i}) < \frac{7}{3}b(K-1)$ and,

$$e_{1,i} > \frac{2b(K-1)}{7} \text{ weeks x team.} \quad (2)$$

The sum of the values $e_{1,i}$ gives us the number **weeks x teams** with open prospects towards the title. The difference $2N - \sum e_{1,i}$, shows how many **weeks x teams** can no longer win it. The ratio $\frac{2N - \sum e_{1,i}}{2N} = l_1$, (which is signed l_1), fixes what percentage of activity matches remains without any possibility to the title goal.

Graphically, this phenomenon is presented in figure 1, (the full green line).

n = b(K-1) Fig. 1.



On the other hand, the first team after f_R weeks is expected to get $\frac{7}{3} f_R$ points and when these ones are more than the points of place i = K - (R-1), the leading team ceases to be threatened by relegation. In general, we can calculate the number $f_{K-(R,i)}$ of the weeks in the relation: $X_i f_{R,i} > X_{K-(R-1)}.n$, from where, $f_{R,i} > \frac{2b(K-2R+1)}{3X_i}$ (the red line in figure 2).

Finally, the top teams near the end of the activity, ensure their participation in the European editions, even if they lose the remaining matches, when their points become bigger than the expected points of Q place. From the relation,

$X_i f_{Q,i} > X_Q.n = (X_1 - \frac{2Q}{K-1}).b(K-1)$, we find out: $f_{Q,i} > \frac{b(7K-6Q-7)}{3X_i}$, (the red dashed line).

By the same logic, we can calculate the sum of **week x teams,** in which teams lose the chances to be ranked in the first Q places, to secure the participation in European editions, in the relation:

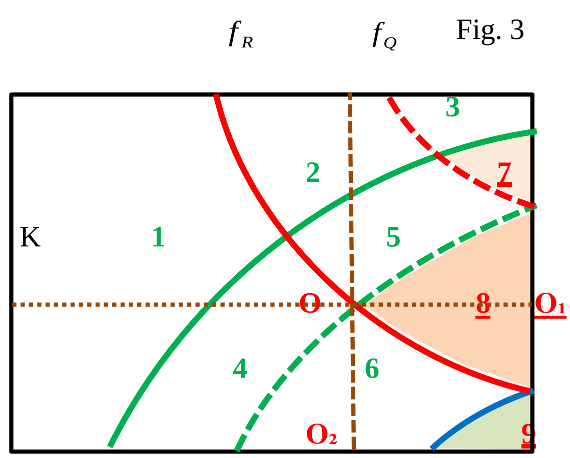
$X_i e_{Q,i} + 3(n - e_{Q,i}) < X_Q \cdot n$, from where: $e_{Q,i} > \frac{2b(K-1+3Q)}{3(3-X_i)}$. (the dashed green line in figure 1).

Likewise, we can find out that after $e_{R,i}$ weeks, the teams of R last places, cannot avoid relegation, in the relation:

$X_i e_{R,i} + 3(n - e_{R,i}) < X_R b(K-1)$ from where,

$e_{R,i} > \frac{b(7K-4R-3)}{3(3-X_i)}$, **(the blue line)** .

Fig. 3



There are 9 zones in the figure above. In the first six of them, from 1 to 6, **there is at least** one of the incentive elements **that works.**

The other three ZONES **7, 8** and **9** , were the object of this study. In each of them, the three incentive elements, **CEASES TO ACT.**

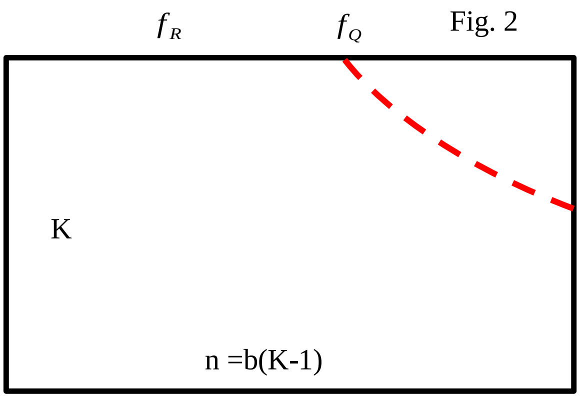


Fig. 2

ANALYSE

Let's overlap the dividing lines of stimulus elements in one figure, to judge their joint action, both at different levels of the teams, and in their extent in time, fig. 3

$$l_t = \frac{6Q(Q-1) + (5K-6Q-4R+5)(K-Q-R) + 5R(R-1)}{18K(K-1)}$$

The value of parameter K for any given country has an accepted value depending on a considerable number of specific factors of each of these countries, while the R parameter varies to values 2 or 3. In such conditions, for the values of K= 10, 12, 14, 16, 18, 20 and with the values of R= 2, 3 or even 4, we can find which are the optimal values for Q, that bring the minimum value $l_{t,\min}$. Substituting in relation above, the values of K and R, we always obtain a quadratic equation of the form $12Q^2 - B \cdot Q + C$, for its numerator. This relation graphically is represented by a parabola. In all cases the discriminant of this equation

The smaller the surface of the sum of ZONES **7, 8** and **9** , the smaller the omission of the sport battle during the championship.

We can easily notice that the surfaces S_7 and S_9 ,increase with **the increase of Q and R,** and, this is accompanied by **the decrease** of the surface S_8, and **vice versa**. After calculations, the total release of the sports battle would be counted:

$$l_t = \frac{S_7 + S_8 + S_9}{S_{total}}, \text{ where:}$$

$$S_7 = \frac{bQ(Q-1)}{3} = \frac{6bQ(Q-1)}{18},$$

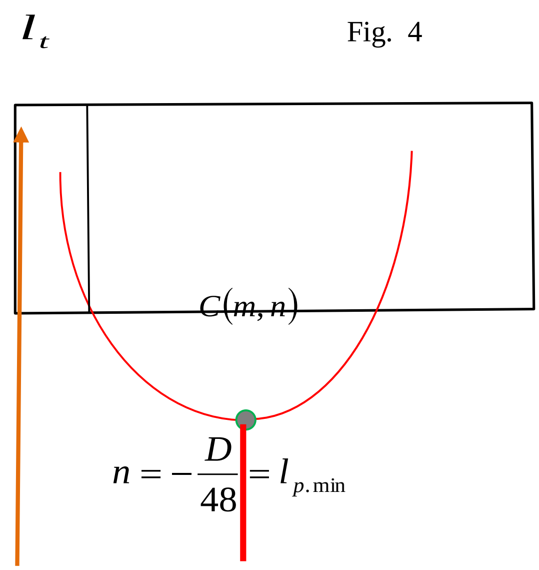
$$S_8 = \frac{b(5K-6Q-4R+5)(K-Q-R)}{18} \quad \text{and,}$$

$$S_9 = \frac{2bR(R-1)}{7} \approx \frac{5bR(R-1)}{18}.$$

The total release of the sports battle, would be calculated in the relation below

turns out to be negative, which means that the coordinates of the vertex of the parabola, are the values $Q_{opt.}$ and $l_{t,\min}$, (the figure below).

Fig. 4



After calculations, for the most common combinations of parameters $K..and..R$, we find out the values for $Q_{opt.}$ as well as the respective values of $I_{t,\min}$, in percentage, (the table below) .

K/R	20/3	18/3	16/3	14/2	12/2	10/2	10/3
$Q_{opt.}$	8	7	6	6	5	4	4
$I_{t,\min}$	11.2	11.0	10.8	11.7	11.1	10.7	9.8

After that, the main thing that stands out, is the fact that the values of $I_{t,\min}$, remain within narrow limits, from 9.8 to 11.7 % of all the matches of the activity.

$$m = -\frac{b}{24} = Q_{opt.} \qquad Q$$

It means that…… EVEN WITH THE BEST COMBINATION OF PARAMETERS K, Q AND R, **THE LACK OF INCENTIVE FOR SPORTING BATTLE CANNOT BE REDUCED MORE THAN THE ABOVE LIMITS.**

IF WE WANT **TO REDUCE** THEM FURTHER ON, WE MUST LOOK FOR A **DIFFERENT SCHEME** OF ORGANIZATION FOR THE ACTIVITY.

PROPOSAL.

All of the above, help us to perceive a scheme to enable further minimization of $I_{t,\min}$. Judging the figure 3, we come to the conclusion that the dynamics of results during the development of the activity brings two groups WITH COMPLETELY DIFFERENT OPPORTUNITIES AND TARGETS, THAT DIFFERENTIATE THEMSELVES, either horizontally, (the line OO_1), or vertically, (the line OO_2, appearing approximately in the middle of the championship). It is this differentiation that shows us how to conceive a more efficient scheme for national football championships.

I am presenting below a scheme with division into two groups, (horizontally), and in two half-parts, (vertically). The first half of it is quite the same as up to now.

The teams ranked in the first group are **finally secured** from relegation. Further on, for the second-half, we have to decide **what will be the starting points of the teams in the second half.** The best and most practical way to continue in the second half, is to establish such a rule to enable a more positive action of the promoting elements for the sports battle in both halves, at the same time. It is not a good choice, either **SAVING** all the points accumulated in the first half, or **EXTINGUISHING** all the points obtained in it. A differentiated starting point in the second half, would be more than logical. Thus, in the ranking table from place $i = 1$, up to $i = K_1$ in the first group, the teams will start in the second half of the championship with

$a_i = K_1 - i$ points, while in the second group the starting points will be $a_i^{'} = K_2 - i$ points.

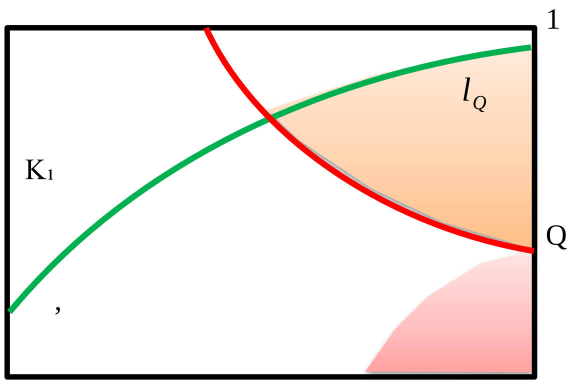
THE SECOND HALF PART

FIRST GROUP

The parameters of this group are: $K_1 = \frac{K}{2} + 2,$ Q_{opt}, $b_1 = 1, n_1 = K_1 - 1$ and $a_i = K_1 - i$.

In this group appear two areas where the motivating elements of the sports battle ceases to act.

Fig. 5



On the other hand, after f_Q weeks, the top teams ensure their stay within the European zone, even with the loss of matches in the remaining weeks. We can write the relation:

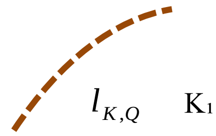
$K_1 - i + X_i \times f_i > K_1 - Q + n_1 \cdot X_Q,$

and after the transformations, we find out:

$f_{i,Q} = \frac{n_1 \cdot X_Q - Q + i}{3 - X_i}$, relation that makes us possible the drawing of red line in fig. 5

SECOND GROUP.

The parameters of the second group for the second half-part, are the following:



.

. $n_1 = K_1 - 1$

The relation

$(K_1 - i) + X_i \cdot e_i + 3(n_1 - e_i) < K_1 - 1 + X_1 \cdot n_1$

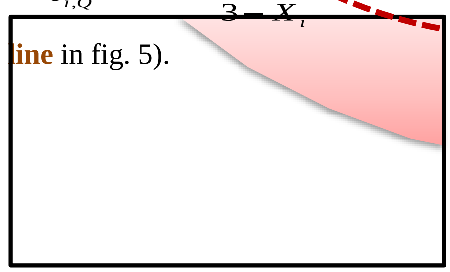
gives us the opportunity to calculate the weeks after which it is not possible to win the title, **(the green line** in the fig.5). After transformations we find: $e_{i,1} = \frac{n_1(3 - X_1) + 1 - i}{3 - X_i}$.

Likewise, we can find out e_Q weeks, after which the last teams, cannot reach any more the points of place Q , in the relation:

$(K_1 - i) + X_i \cdot e_{i,Q} + 3(n_1 - e_{i,Q}) < K_1 - Q + X_Q \cdot n_1$

from where: $e_{i,Q} = \frac{n_1(3 - X_Q) + Q - i}{3 - X_i}$

(**the brown line** in fig. 5).



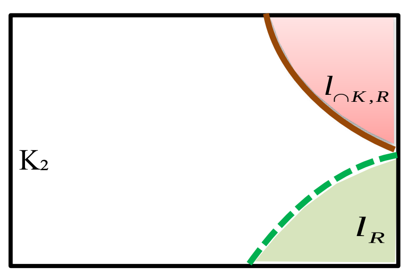


Fig. 6.

$K_2 = K - K_1, \quad n_2 = K_2 - 1, \quad b_2 = 1,$ $a_i^{'} = K_2 - i$ and R.

The top teams of this group aim to escape the relegation from the category. We can write:

$K_2 - i + x_i \cdot f_i^{'} = K_2 - R + X_R \cdot n_2$, and

$f_i > \frac{X_R \cdot n_2 + i - R}{X_i}$.

This relation gives us the data to draw the brown line in the fig. 6.

Finally, the teams of the last places after e_R weeks, even winning the remaining matches, does not manage to come out on top with the points of the fourth place from the end. Writing the relation

$$(K_2 - i) + X_i \cdot e_{i,R} + 3(n_2 - e_{i,R}) < K_2 - R + X_R \cdot n_2,$$

and making then transformations and replacements, we find at last:

$$e_{i,R} > \frac{n_2(3 - X_R) + R - i}{3 - X_i}$$

, that enables us to find now the values of $e_{i,R}$ weeks and to draw, as well, the green line of fig. 6.

$n_2 = K_2 - 1$

Conclusion

At last, let's compare the existing scheme of organizing the national football championship with the suggested one.

First of all, we must affirm that although the above three stimulus elements cease to act, the teams into the area 7, are interested in

Accepting zone 7 no longer as an uninterested one, the comparison of the two schemes should be done by considering the areas 8 and 9, only. So,

The existing scheme:

$l_{\cap K,Q,R} + l_R = 9.8 + 1.2 = 11.0\%$., or

11.0% x 306= **<u>34</u> uninterested matches.**

The proposed one:

$l_{\cap K,Q} + l_{\cap K,R} + l_R = 4.7\%$, or

4,7% x 226=**<u>11</u> matches, unmotivated**

using the remaining matches as a preparation for the European editions, in which they will very soon take part.

Second, the teams of European zone were not "donated" this position, on the contrary, they deserved it with their results. From this point of view, such an area, is "the least evil" in football activities.

Therefore, the comparison of the existing way with the proposed one, should focus on zones **8** and **9**.

Currently, the existing organizational scheme brings:

The zone **7,** $l_{\cap Q} = 2.8\%$

The zone 8, $l_{\cap K,Q,R} = 9.8\%$

The zone 9, $l_R = 1.2\%$

The parameters of the proposed scheme above, are:

The zone 7, $l_{\cap Q} = 5.7\%$

The zone 8'+8", $l_{\cap K,Q} + l_{\cap K,R} = 3.3\%$

The zone 9, $l_R = 1.4\%$.

References

Mina Naqo (1972). Theory of probabilities Part I and II, Tirana, Albania

Mina Naqo (1972). Mathematical statistics Tirane, Albania

Ylli Pango (1985). Statistical methods in psychological studies, Tirana, Albania

.