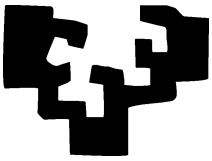
eman ta zabal zazu



Universidad del País Vasco

Euskal Herriko Unibertsitatea

Activity yyy

1 Synopsis.

What we are set about to do. This activity will illustrate the use of statistical testing to decide whether football tactics has evolved over time.

What you need to know. In order to benefit from this activity you need:

1. A working knowledge of R.
2. A rather thorough understanding of hypothesis testing, and in particular of tests for independence and homogeneity in contingency tables.
3. Preferably a good understanding on how to perform a Fisher's exact test.

2 Context.

It has been claimed that football has lost part of its appeal because of more conservative tactics that lead to fewer goals than used to be the rule. Table 1 totals the number of matches and goals in the World Cup[1].

It looks indeed as if prior to 1960 the number of goals per match was somewhat higher. Your task is to check with statistical criterion whether the data supports said hypothesis.

3 Questions.

There are several ways you could go about to test the hypothesis mentioned.

1. Matches are a relatively rare event. Only in a few of the 90 minutes duration of a football match we see a goal[2]. One might think that each minute there is a (small) risk of one of the two teams scoring a

[1] The data has been obtained from http://www.guardian.co.uk/football/datablog/2010/jun/17/opta-world-cup-2010-data. Download from the original source. You also have it in ready-to-read format in MOODLE (file `WorldCup.dge`, readable into R with a `dget` instruction.)

[2] In the 1984 UEFA European Football Championship, Spain won over Malta by 12-1, with a whopping average of 0.20 goals per minute in the second half; a very unusual event, and the second biggest win in Spain's history.

Year	Goals	Matches	Goals/match
1930	70	18	3.89
1934	70	17	4.12
1938	84	18	4.67
1950	88	22	4.00
1954	140	26	5.38
1958	121	35	3.46
1962	88	32	2.75
1966	89	32	2.78
1970	95	32	2.97
1974	97	38	2.55
1978	102	38	2.68
1982	146	52	2.81
1986	132	52	2.54
1990	115	52	2.21
1994	141	52	2.71
1998	171	64	2.67
2002	161	64	2.52
2006	147	64	2.30
2010	25	16	1.56

Table 1: World Cup Statistics. Source: see text.

goal. What would this suggest as a plausible approximation for the distribution of the total number of goals in a match?

2. If we sum the goals in a large number of matches, what does the central limit theorem say about the distribution of the total number of goals? About the average number of goals per match?

3. Choose a statistical model. Phrase the question of interest as a hypothesis test on the parameter(s) of said model.

4. What are your results? Do you find support for the belief that more conservative modern football tactics leads to fewer goals per match?

5. Re-read the previous question. And now, do you find evidence that in modern editions of the World Cup there have been fewer goals per match than in the pre-1960 period?

6. What is the difference between the previous two questions?

7. You have performed an statistical test on the parameters of a model; the initial (vague) statement of the problem ventured an *explanation* why there might be fewer goals in modern football (more conservative tactics). What does your statistical result support? (HINT: Are there alternative plausible explanations, different than conservative tactics, that would agree with your statistical result?)