Probabilistic Excitement of Sports Games

Jan Vecer Tomoyuki Ichiba Mladen Laudanovic

Department of Statistics, Columbia University,
http://www.stat.columbia.edu/∼vecer

Harvard University, September 29, 2007

Abstract

In this talk we introduce a quantitative measure of the excitement of sports games. This measure can be thought of as the variability of the expectancy of winning as a game progresses. We illustrate the concept of excitement at soccer games for which the theoretical win expectancy can be well approximated from a Poisson model of scoring. We show that in the Poisson model, higher scoring rates lead to increased expected excitement. Given a particular strength of a team, the most exciting games are expected with opponents who are slightly stronger. We apply this theory to the FIFA World Cup 2006 games, where the winning expectancy was independently estimated by betting markets. Thus it was possible to compute the expected and the realized excitement of each given game from the trading data.

Motivation

Higher Excitement comes with

- More changes in the score lead.
- Changes that are more dramatic.
- Decisive moments later in the game.

Fluctuation of Win Expectancy satisfies the above criterions. More changes to the Win Expectancy, the less "predictable" is the outcome of the game.

Definition

We propose the following measure for excitement:

$$\text{Excitement} = \text{Variability of the Winning Expectancy.} \tag{1}$$

Variability can be measured as the Total Variation (TV):

$$TV(f) = \lim_{\max |t_{i+1} - t_i| \to 0} \sum |f(t_{i+1}) - f(t_i)|, \tag{2}$$

where $0 = t_0 < t_1 < \cdots < t_n = T$ is a partition of the interval $[0, T]$. Total variation can be viewed as the vertical component of the arc-length of the graph of a given function f. The longer the path of win expectancy for a given team, the more swings there are in the game, and thus the game is more exciting.

Definition - Continued

Formally, we can define

Definition

$$\begin{aligned}\text{Excitement} = \text{TV}(&\text{Probability of Team 1 Wins}) \\ &+ \text{TV}(\text{Probability of Team 2 Wins}). \quad (3)\end{aligned}$$

Definition - Continued

This definition makes sense if a draw is not an option, such as in the elimination round games. If a draw is possible, we can use a modified version:

Definition

$$\begin{aligned}\text{Excitement} = \text{TV}(&\text{Probability that Team 1 Wins}) \\ &+ \text{TV}(\text{Probability of Draw}) \\ &+ \text{TV}(\text{Probability that Team 2 Wins}). \quad (4)\end{aligned}$$

Definition Explained

In general, the total variation of winning expectancy is mostly changed if there is a game deciding event close to the end of the game, if there are a number of events where the game lead is changed, or if the weaker team unexpectedly wins or draws the game. On the other hand, only small changes in the total variation of winning probabilities occur when the game is one sided, with an early lead from the favorite team.

Win Expectancy

Our approach to quantitatively measure the excitement via variability of the win expectancy described in this talk is novel and it could be applied to all sports played with two competing teams.

Win Expectancy can be computed/observed from

- The Model.
- Betting Market.

Measurement from Model

There are already some cases of sports specific attempts to measure excitement based on model. The problem is that for many sports, it is not entirely clear how to determine the win expectancy during the course of the game. The evolution of the game can be modeled with rather simplifying assumptions for only a small number of sports. The simplest and analytically tractable models assume no memory during the game which suggests the use of Markov models. Sports whose evolution could be approximated well by Markov models include baseball, tennis, soccer, or hockey.

Measuring from Betting Market.

Betting market (in play) provides excellent feedback on measuring probability of win/loss of a certain team. Betting markets are now organized as major financial exchanges (such as betfair.com), and thus are very efficient and reliable in estimating the correct Win Expectancy at any given point during the game.

US entities cannot participate in the market for legal reasons, but the betting exchange trades US sports events as well (traded by foreign entities).

Focus of the Talk

We will focus on soccer in this talk since the it is both tractable from the model, and there is enough data from the betting market.

Poisson Model

Poisson process is a memoryless counting process (i.e., taking integer values) in continuous time. Its evolution is given by

$$\mathbb{P}(X_{t+\Delta t} = X_t | X_t) = 1 - \lambda(\Delta t) + o(\Delta t), \quad (5)$$

$$\mathbb{P}(X_{t+\Delta t} = X_t + 1 | X_t) = \lambda(\Delta t) + o(\Delta t). \quad (6)$$

Score Evolution

We assume that the scores of the two teams evolve as independent Poisson processes. In particular, if $X_T : Y_T$ denotes the final score of the game, we have

$$\mathbb{P}(X_T = X_t + k) = \exp[-\lambda(T-t)]\frac{[\lambda(T-t)]^k}{k!},$$

and

$$\mathbb{P}(Y_T = Y_t + k) = \exp[-\mu(T-t)]\frac{[\mu(T-t)]^k}{k!},$$

where $X_t : Y_t$ is the current score at time t in the game. Variable T is the end time of the game, and it is assumed to be fixed.

Score Evolution

Parameters λ and μ are called the scoring intensities of the two teams. They are related to the expected score by the following relationship:

$$\mathbb{E}[X_T - X_t | X_t] = \lambda(T-t), \qquad \mathbb{E}[Y_T - Y_t | Y_t] = \mu(T-t).$$

Thus one should expect to see on average $\lambda(T-t)$ and $\mu(T-t)$ goals for the two teams in the remaining $T-t$ time of the game.

Win Expectancy

The Poisson model of scoring could be used for modeling the evolution of scoring in soccer or hockey games. The memoryless property implies that the time between the goals is exponentially distributed with parameters λ and μ respectively. In this model we can get explicit formulas for the win expectancy of each team, and the expectancy of a draw.

Win Expectancy - Continued

If the current time is $t \in [0, T]$, and the current score is $X_t : Y_t$, we have

$$\mathbb{P}(\text{Team 1 Will Win}) = \mathbb{P}(X_T > Y_T) =$$

$$= \sum_{k=0}^{\infty} \mathbb{P}(X_T = k - X_t, Y_T < k - X_t)$$

$$= \sum_{k=0}^{\infty} \left[e^{-\lambda_t} \frac{\lambda_t^k}{k!} \cdot \sum_{i=0}^{k+X_t-Y_t-1} e^{-\mu_t} \frac{\mu_t^i}{i!} \right] \quad (7)$$

Win Expectancy - Continued

$$\mathbb{P}(\text{Draw}) = \mathbb{P}(X_T = Y_T) =$$

$$= \sum_{k=0}^{\infty} \left[e^{-(\lambda_t+\mu_t)} \cdot \frac{\lambda_t^{(k+\max(X_t+Y_t)-X_t)}}{(k+\max(X_t+Y_t)-X_t)!} \cdot \frac{\mu_t^{(k+\max(X_t+Y_t)-Y_t)}}{(k+\max(X_t+Y_t)-Y_t)!} \right] \tag{8}$$

Win Expectancy - Continued

$$\mathbb{P}(\text{Team 2 Will Win}) = \mathbb{P}(Y_T > X_T) =$$

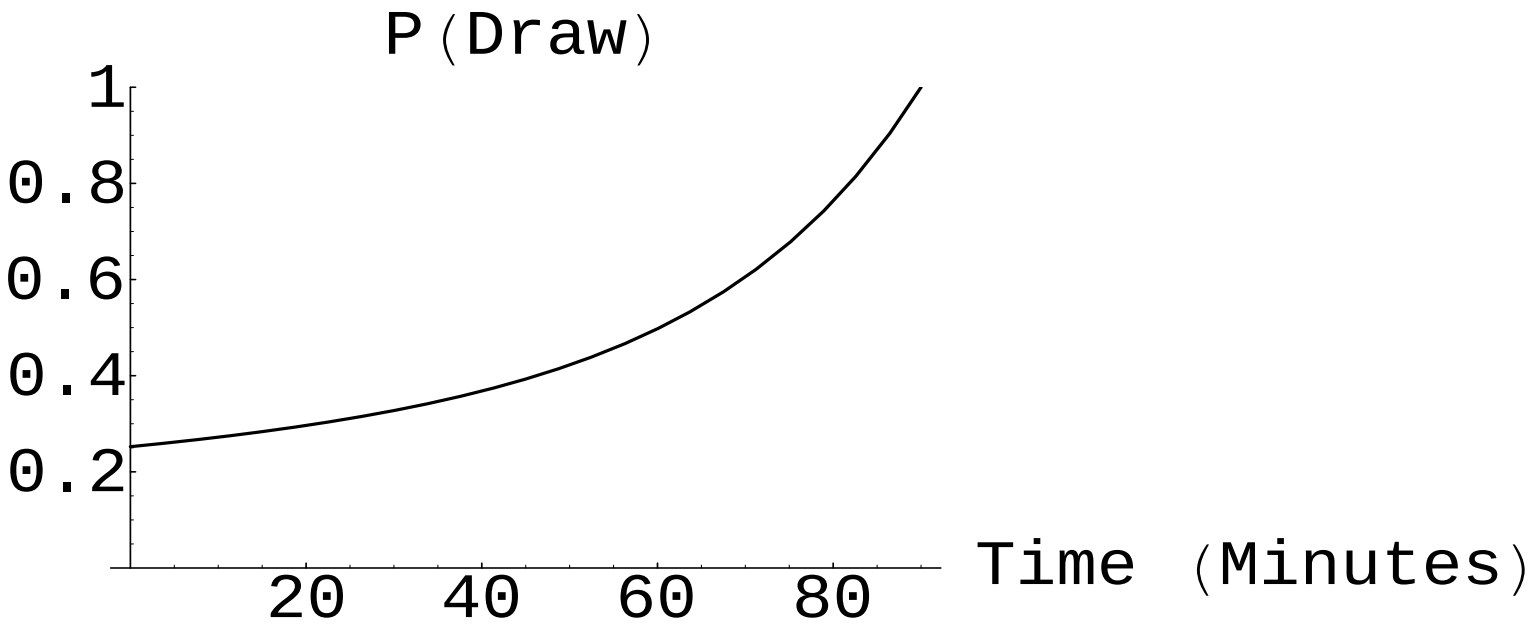
$$= \sum_{k=0}^{\infty} \mathbb{P}(Y_T = k - Y_t, X_T < k - Y_t)$$

$$= \sum_{k=0}^{\infty} \left[e^{-\mu_t} \frac{\mu_t^k}{k!} \cdot \sum_{i=0}^{k-X_t+Y_t-1} e^{-\lambda_t} \frac{\lambda_t^i}{i!} \right] \quad (9)$$

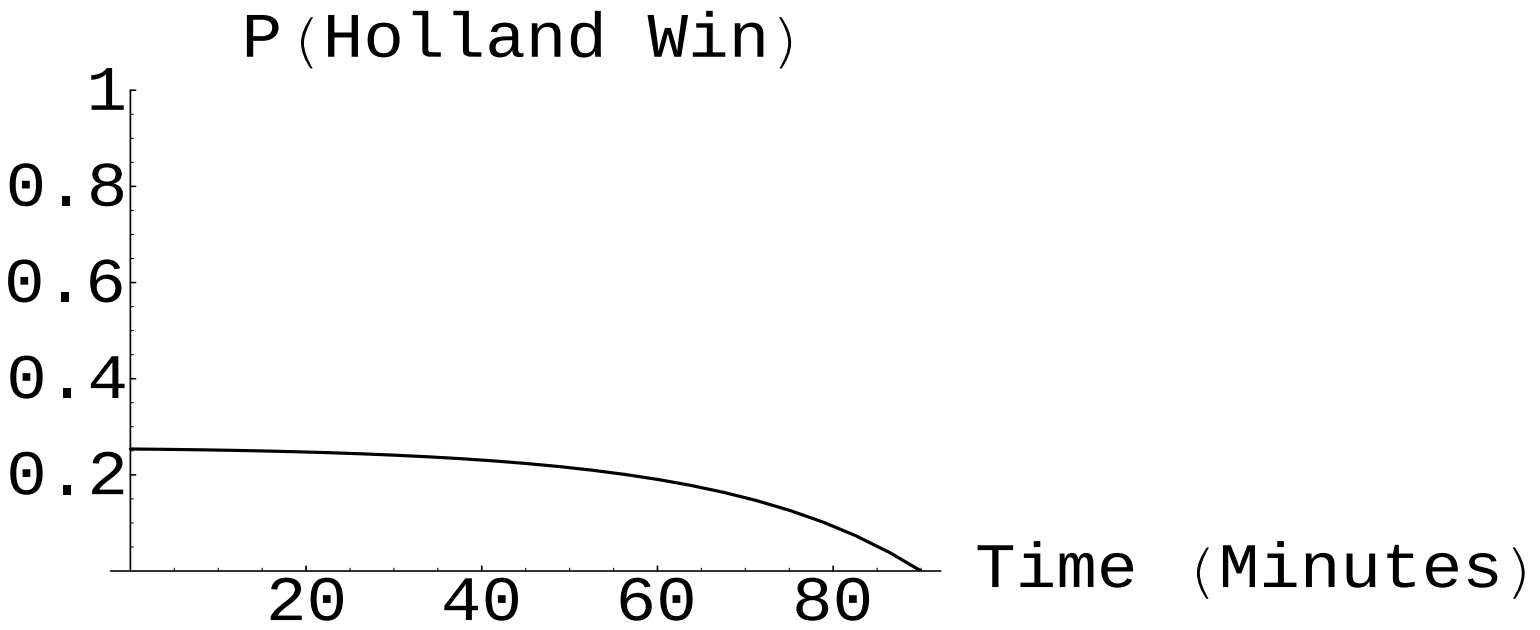
Holland - Argentina

Evolution of probabilities (implied by a Poisson model of scoring) of a draw, win of Holland and win of Argentina during the Holland - Argentina game. The game ended with a 0:0 draw. The scoring intensity for Holland was estimated at 1.05, and for Argentina at 1.57.

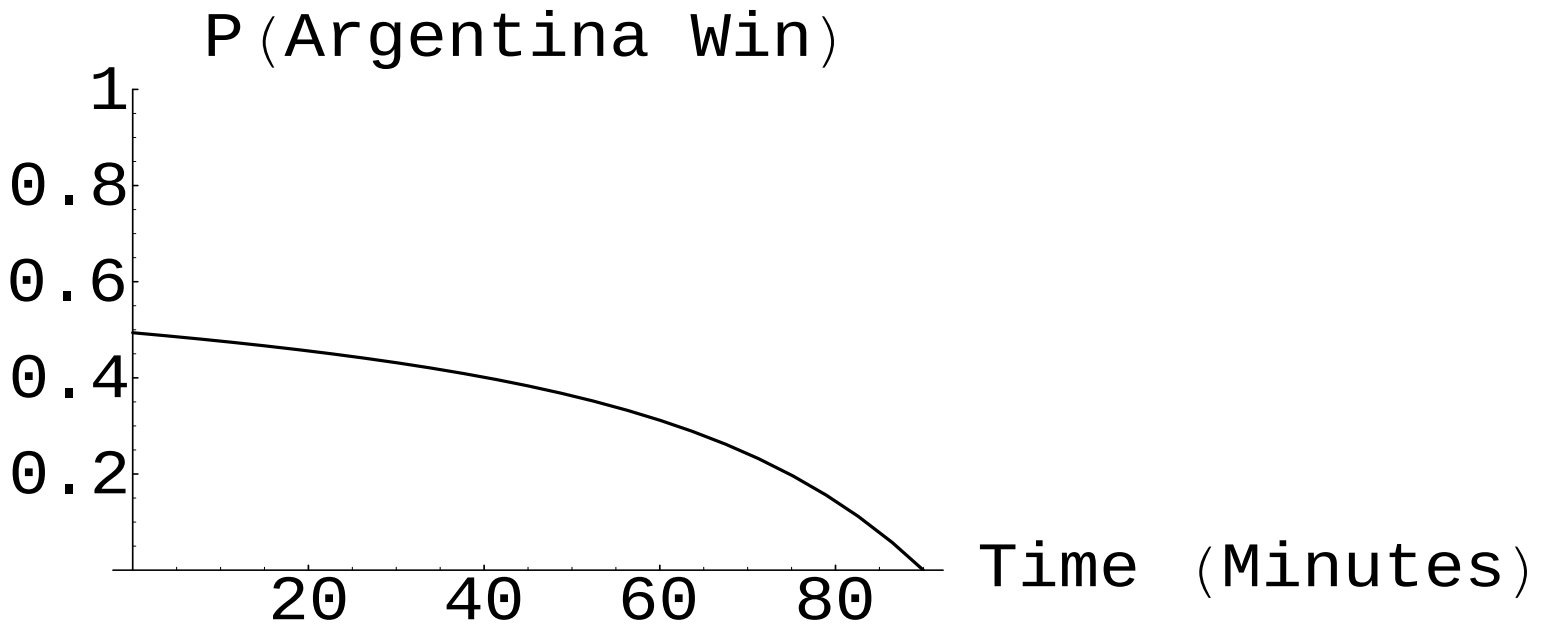
Draw



Holland



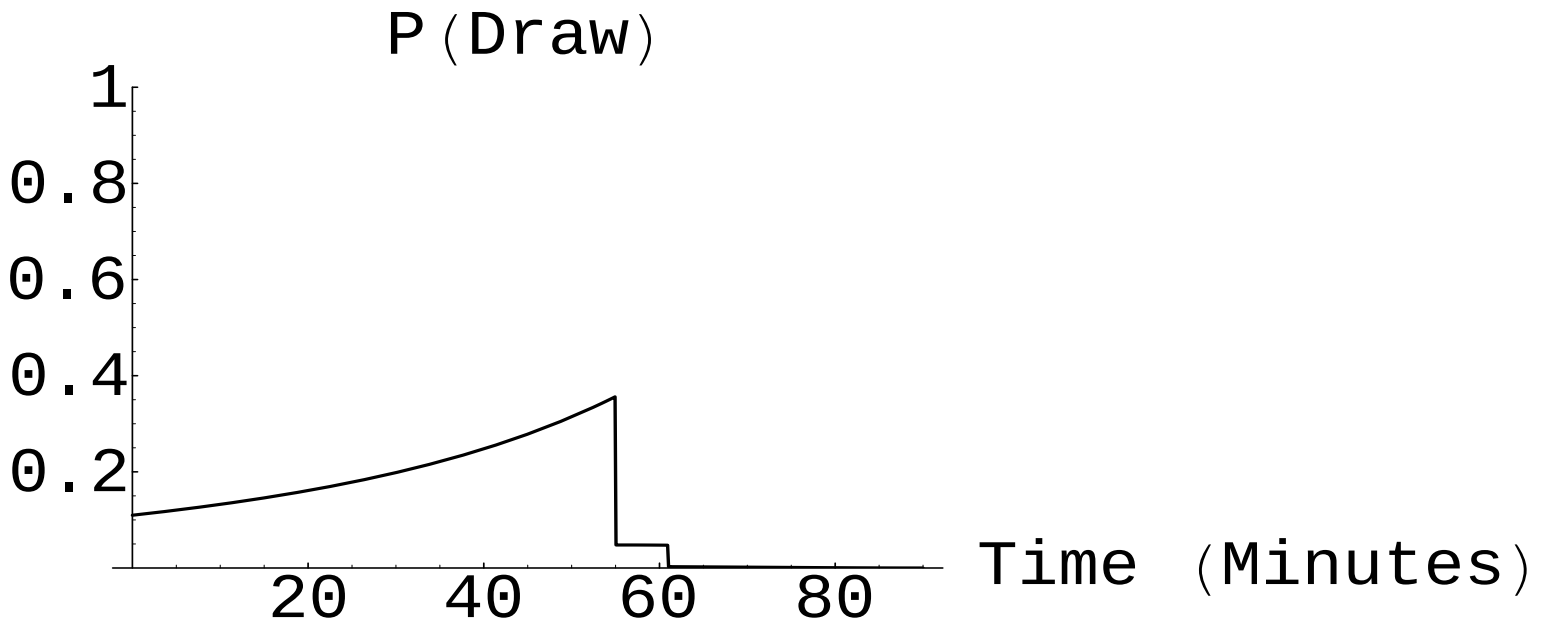
Argentina



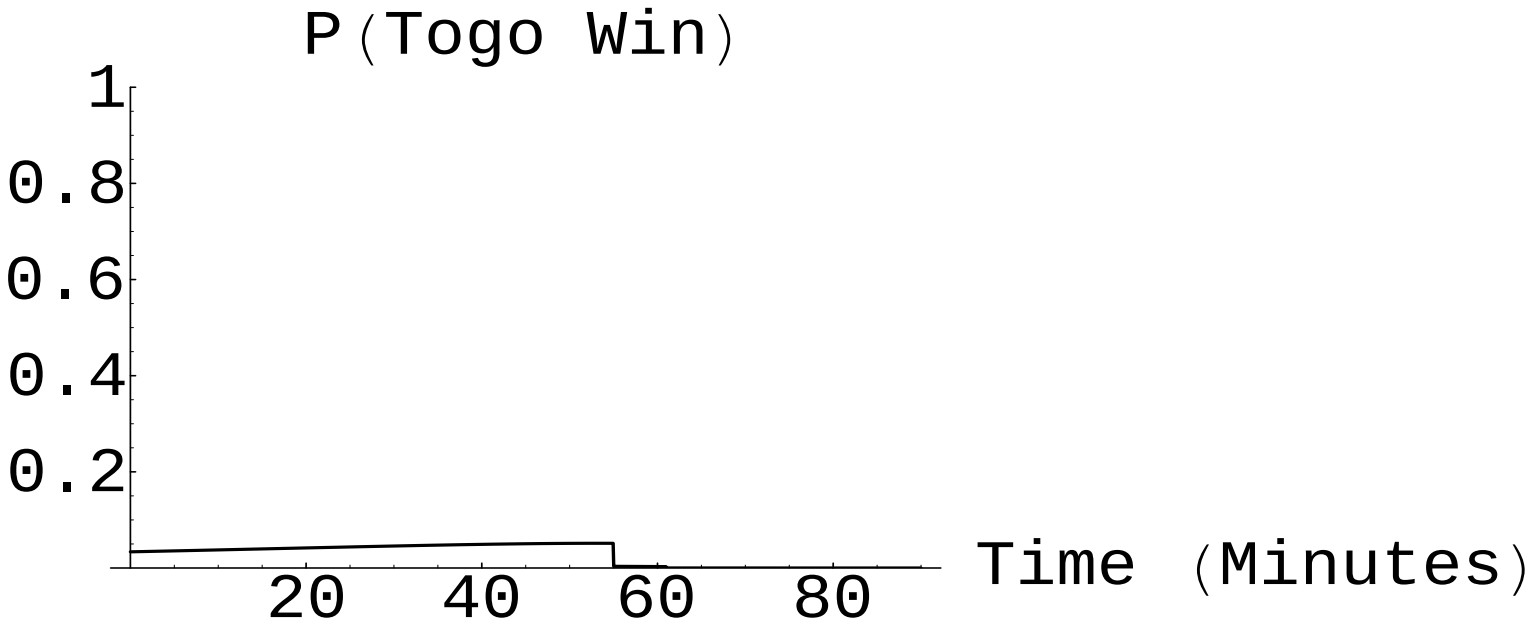
Togo - France

Evolution of probabilities (implied by a Poisson model of scoring) of a draw, win of Togo and win of France during the Togo - France game. France scored in the 55th and the 61st minute of the match. The betting market estimated the scoring intensities to be 0.37 for Togo and 2.65 for France. The expected excitement was 1.28, the lowest for all the games played in that championship.

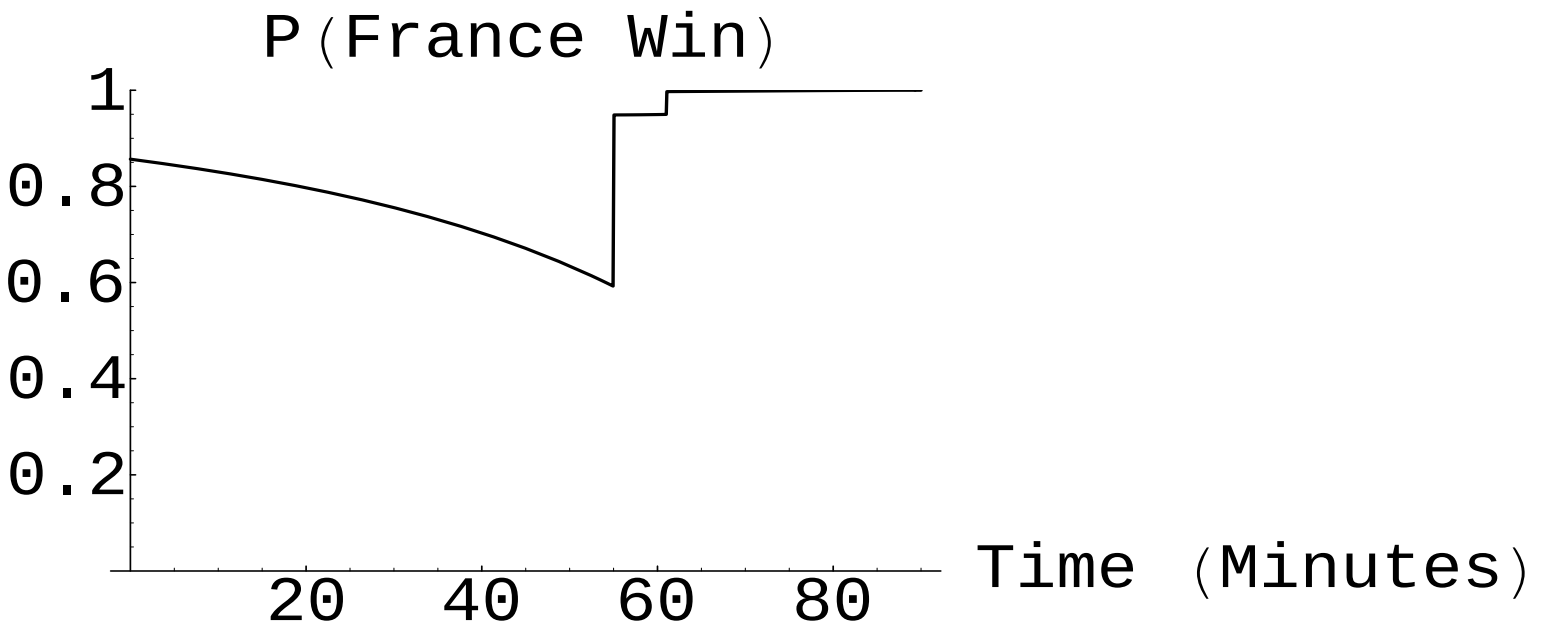
Draw



Togo



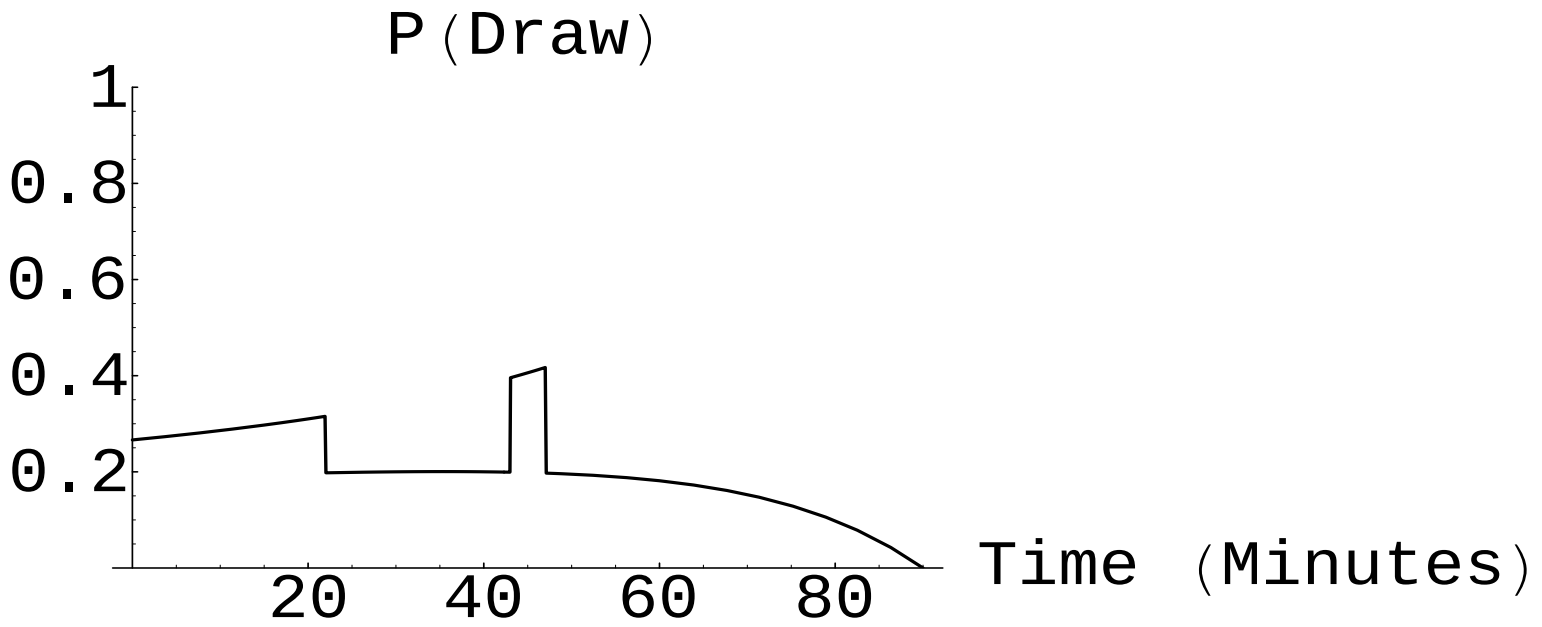
France



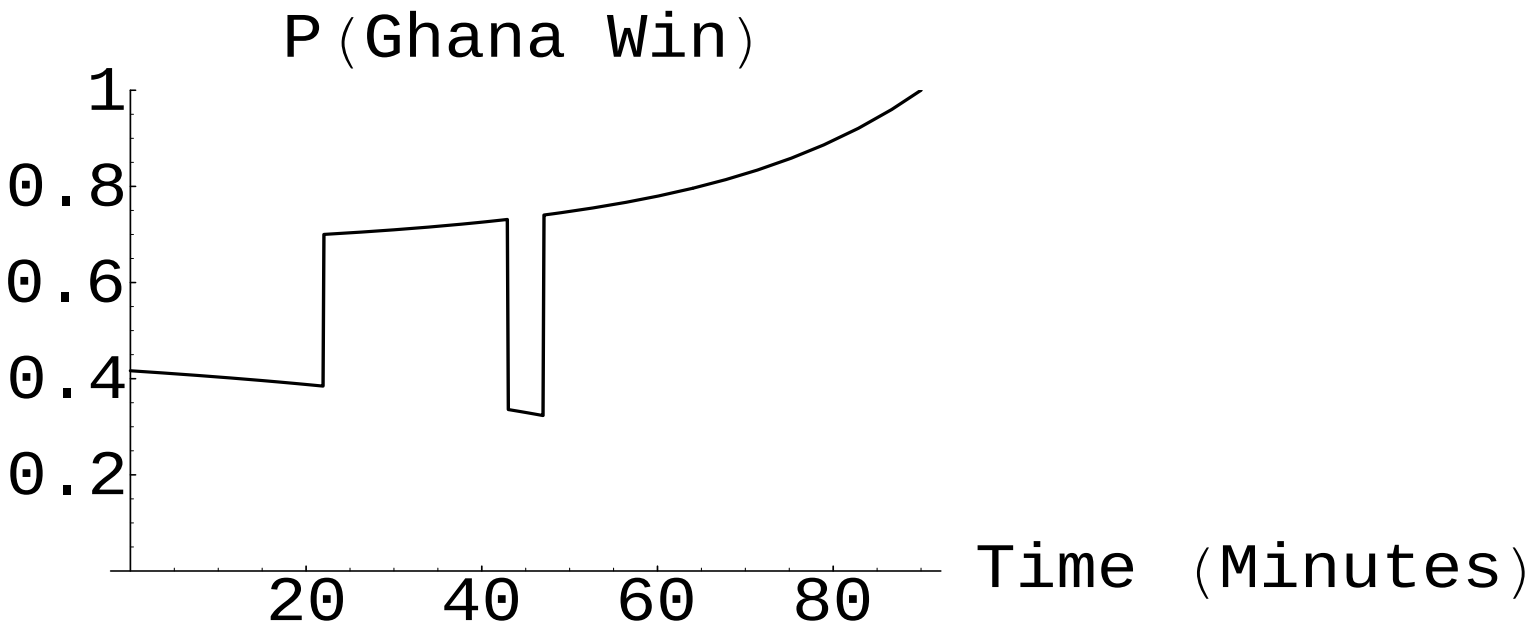
Ghana - USA

Evolution of probabilities (implied by a Poisson model of scoring) of a draw, win of Ghana and win of United States during the Ghana - United States game. Ghana scored in the 22nd minute, followed by the goal of United States in the 43rd minute and the second goal of Ghana in the 47th minute. The scoring intensity for Ghana was estimated at 1.37, and for the United States at 1.16. The expected excitement was 2.55, the highest expectation among all the games in the championship.

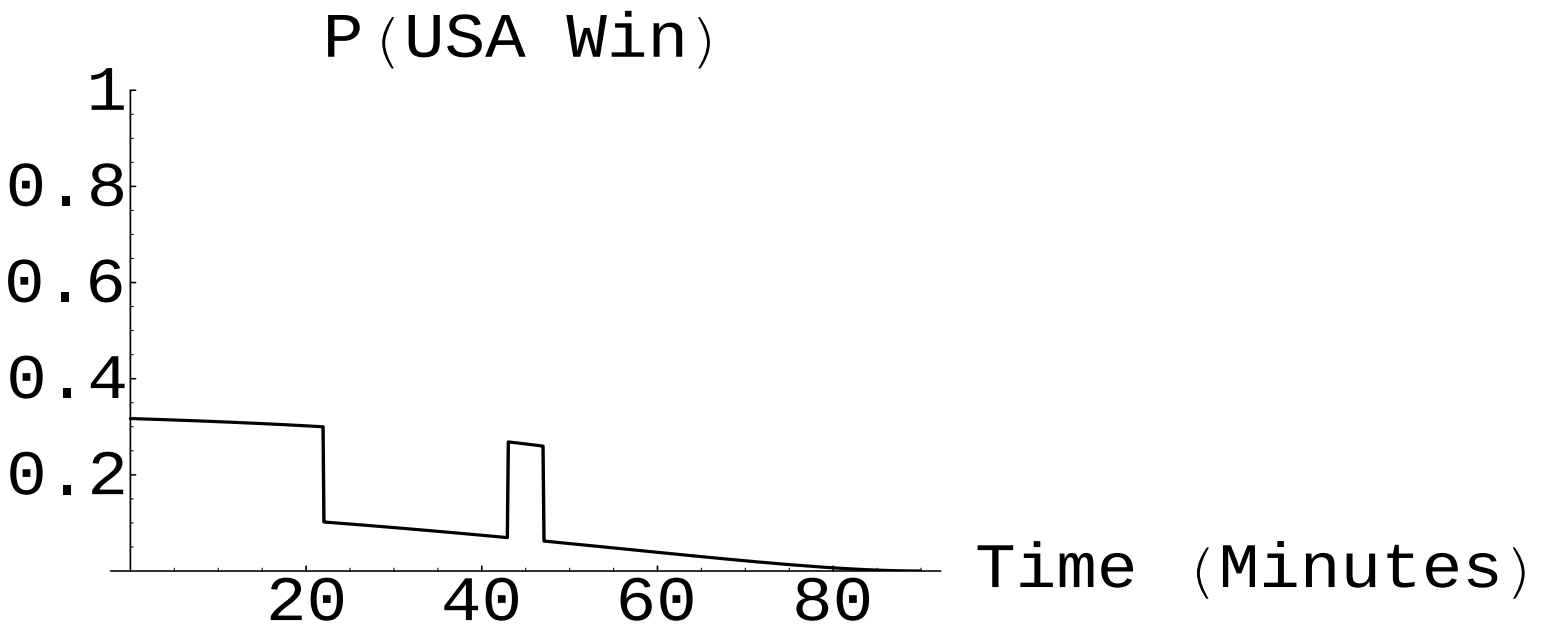
Draw



Ghana



USA



Game	λ_1	λ_2	Score	Win	Loss	Draw	Total	Model	Exp
Spain – France	1.16	0.92	1 - 3	1.48	1.16	1.41	4.05	4.53	2.42
Italy – Australia	1.65	0.62	1 - 0	1.57	0.35	1.68	3.60	3.52	2.11
Germany * – Argentina	1.22	1.06	1 - 1	0.62	1.27	1.47	3.36	3.62	2.47
Argentina * – Mexico	2.07	0.57	1 - 1	1.11	0.91	0.78	2.80	2.73	1.87
Italy * – France	0.97	0.82	1 - 1	0.97	0.93	0.87	2.77	2.77	2.36
Portugal – Holland	1.15	0.98	1 - 0	1.06	0.58	0.44	2.08	1.40	2.44
Brazil – France	1.56	0.78	0 - 1	0.57	0.82	0.59	1.98	2.18	2.28
England – Portugal *	0.78	1.22	0 - 0	0.46	0.58	0.61	1.65	1.40	2.33
Germany – Portugal	1.56	0.93	3 - 1	0.72	0.28	0.60	1.60	1.76	2.40
England – Ecuador	1.73	0.64	1 - 0	0.69	0.17	0.64	1.50	1.82	2.08
Portugal – France	0.77	1.13	0 - 1	0.73	0.33	0.42	1.48	1.48	2.34
Switzerland – Ukraine *	0.90	1.20	0 - 0	0.25	0.45	0.68	1.38	1.41	2.43
Germany – Italy *	1.06	0.91	0 - 0	0.66	0.39	0.30	1.35	1.38	2.39
Italy – Ukraine	1.43	0.60	3 - 0	0.44	0.15	0.28	0.86	0.91	2.16
Germany – Sweden	1.72	0.74	2 - 0	0.39	0.15	0.25	0.79	0.83	2.20
Brazil – Ghana	0.50	2.21	3 - 0	0.22	0.07	0.17	0.46	0.57	1.71

Table: Elimination round games ordered by the excitement level, regulation time plus injury time. Draw was a possible outcome. Games marked by a star ended up with draw, and went into overtime.

Game	Score	T1 Advance	T2 Advance	Total
England – Portugal *	0 - 0	1.47	1.53	3.00
Portugal – Holland	1 - 0	1.24	1.16	2.40
Italy * – France	1 - 1	1.10	1.16	2.26
Switzerland – Ukraine *	0 - 0	1.14	1.02	2.16
Germany * – Argentina	1 - 1	1.06	1.04	2.10
Spain – France	1 - 3	1.00	0.92	1.92
Italy – Australia	1 - 0	0.74	0.76	1.50
Brazil – France	0 - 1	0.71	0.70	1.41
Argentina – Mexico	2 - 1	0.67	0.55	1.22
Germany – Italy	0 - 2	0.57	0.56	1.13
Germany – Portugal	3 - 1	0.40	0.49	0.89
Portugal – France	0 - 1	0.48	0.40	0.88
England – Ecuador	1 - 0	0.41	0.38	0.79
Italy – Ukraine	3 - 0	0.26	0.26	0.52
Germany – Sweden	2 - 0	0.22	0.20	0.42
Brazil – Ghana	3 - 0	0.11	0.18	0.29

Table: Elimination round games ordered by the excitement level, including overtime. Draw was not an option. Games marked by a star went into penalty shootouts.

Expected Excitement - Poisson Model

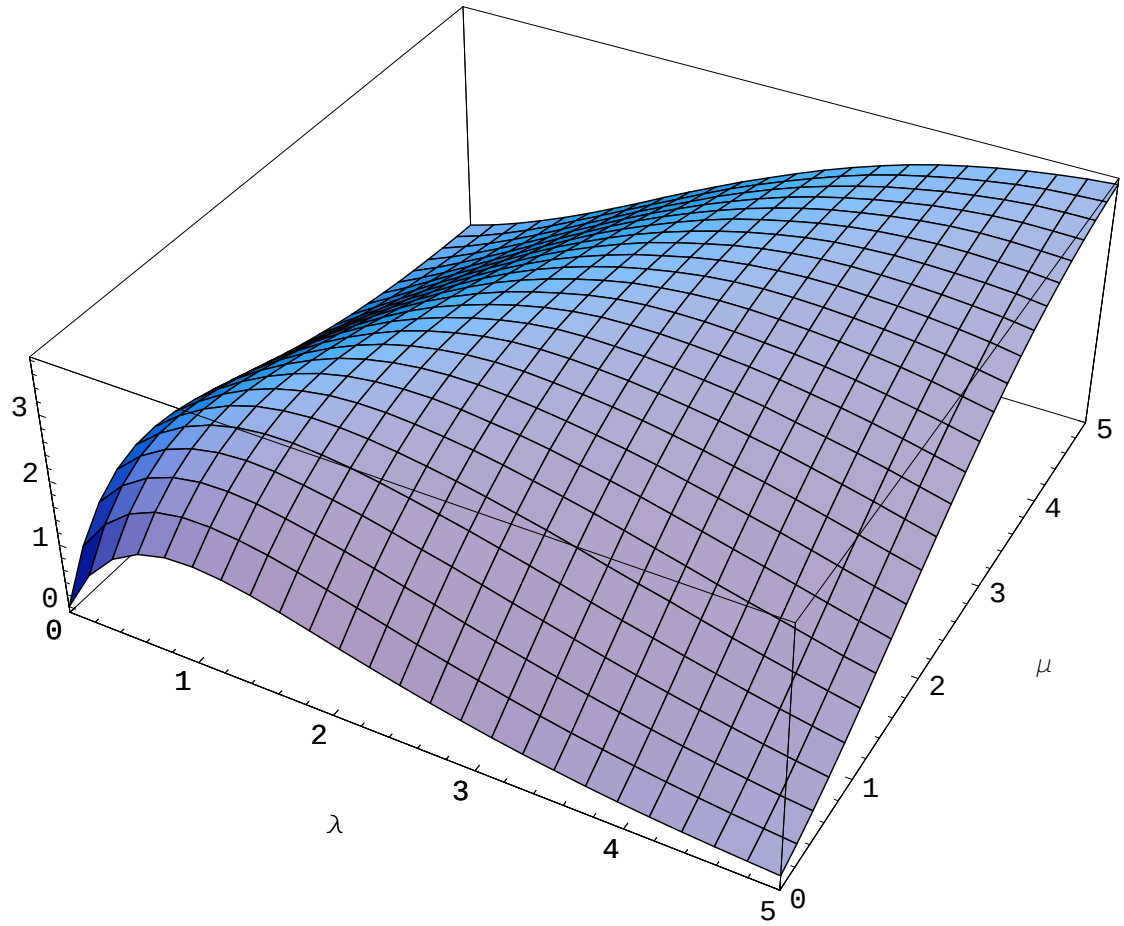


Figure: Expected excitement as a function of intensities of scoring.