EURO 2020 groupstage predictions: 1st match-day

Leonardo Egidi - DEAMS, University of Trieste, Italy. Mail: legidi@units.it

8 June 2021

Contents

The statistical model (in brief)

We use a **double Poisson model with dynamic team-specific abilities** for the attack and the defence. Let (X_i, Y_i) denote the random number of goals scored by the home and the away team in the i-th game, $i = 1, \ldots, n$, respectively. ranking denotes the Coca-Cola FIFA ranking at May 27th, 2021, whereas att and def denote the attack and the defence abilities, respectively.

$$X_i | \lambda_{1i} \sim \text{Poisson}(\lambda_{1i}), \tag{1}$$

$$Y_i | \lambda_{2i} \sim \text{Poisson}(\lambda_{2i}), \tag{2}$$

$$\log(\lambda_{1i}) = \text{ home} + \text{att}_{h_i,t} + \text{def}_{a_i,t} + \frac{\gamma}{2}(\text{ranking}_{h_i} - \text{ranking}_{a_i}) \tag{3}$$

$$\log(\lambda_{2i}) = \text{att}_{a_i,t} + \text{def}_{h_i,t} - \frac{\gamma}{2}(\text{ranking}_{h_i} - \text{ranking}_{a_i}), \quad i = 1, \ldots, n \text{ (matches)}, \tag{4}$$

$$\text{att}_{k,t} \sim \mathcal{N}(\text{att}_{k,t-1}, \sigma^2), \tag{5}$$

$$\text{def}_{k,t} \sim \mathcal{N}(\text{def}_{k,t-1}, \sigma^2), \tag{6}$$

$$\sum_{k=1}^{n_t} \text{att}_{k,} = 0, \ \sum_{k=1}^{n_t} \text{def}_{k,} = 0, \quad k = 1, \ldots n_t \text{ (teams)}, \ t = 1, \ldots, T \text{ (times)}. \tag{7}$$

Lines (1)-(2) display the likelihood's equations (two Poisson distributions); lines (3)-(4) display the log-linear models for the scoring rates λ_1, λ_2; lines (5)-(6) display the dynamic prior distributions for the attack and the defence parameters, respectively; line (7) displays the sum-to-zero identifiability constraints. Model fitting has been obtained through the Hamiltonian Monte Carlo sampling, 2000 iterations, 4 chains (`rstan` package). The historical data used to fit the models come from: **Nations' League** (2019-2020), **Euro UEFA Qualifiers** (2020-2021), **World Cup UEFA Qualifiers** (2021).

The idea is to provide a dynamic predictive scenario: at the end of each match-day, the model will be refitted to predict the remaining matches.

Groupstage predictions: 1st match-day (11-15 June)

Posterior matches probabilities from the posterior predictive distribution of the model above are displayed in the table below (home win, draw, away win probabilities). **mlo** denotes the most likely exact outcome (in

parenthesis, the corresponding posterior probability). Darker regions in the plots below denote more likely outcomes: on the x-axis the home goals, on the y-axis the away goals.

home	away	home win	draw	away win	mlo
Italy	Turkey	0.624	0.226	0.151	1-0 (0.145)
Wales	Switzerland	0.404	0.289	0.307	1-0 (0.138)
Denmark	Finland	0.748	0.166	0.086	2-0 (0.134)
Russia	Belgium	0.100	0.155	0.745	0-2 (0.104)
England	Croatia	0.662	0.199	0.139	2-0 (0.124)
Austria	FYR Macedonia	0.626	0.213	0.161	1-0 (0.124)
Netherlands	Ukraine	0.587	0.230	0.183	1-0 (0.125)
Scotland	Czech Republic	0.363	0.280	0.357	1-1 (0.124)
Poland	Slovakia	0.592	0.228	0.180	1-0 (0.126)
Spain	Sweden	0.601	0.224	0.175	1-0 (0.127)
Hungary	Portugal	0.221	0.244	0.535	0-1 (0.126)
Germany	France	0.287	0.258	0.455	1-1 (0.118)

Posterior match probabilities