

# EURO 2020 groupstage predictions: round of 16

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

24 June 2021

## Contents

<b>The statistical model (in brief)</b>	<b>1</b>
<b>Groupstage predictions: round of 16 (26-29 June)</b>	<b>1</b>

## 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,  $i = 1, \dots, n$ , respectively. `ranking` denotes the Coca-Cola FIFA ranking at May 27th, 2021, whereas `att` and `def` denote the attack and the defence abilities, respectively.

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

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

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

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

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

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

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

Lines (1)-(2) display the likelihood's equations (two Poisson distributions); lines (3)-(4) display the log-linear models for the scoring rates  $\lambda_1, \lambda_2$ ; 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), **UEFA Euro 2020** (groupstage matches).

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: round of 16 (26-29 June)

Posterior matches probabilities from the posterior predictive distribution of the model above are displayed in the table below. `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.

**Attention:** the matches probabilities below refer to the results within the regular 90 minutes.

home	away	home win	draw	away win	mlo
Wales	Denmark	0.370	0.287	0.344	0-0 (0.122)
Italy	Austria	0.661	0.212	0.127	1-0 (0.145)
Netherlands	Czech Republic	0.569	0.241	0.190	1-0 (0.132)
Belgium	Portugal	0.657	0.192	0.151	2-0 (0.107)
Croatia	Spain	0.327	0.269	0.403	1-1 (0.124)
France	Switzerland	0.626	0.224	0.150	1-0 (0.138)
England	Germany	0.614	0.219	0.167	1-0 (0.126)
Sweden	Ukraine	0.539	0.246	0.215	1-0 (0.129)

