

# A Bayesian approach for predicting the results of matches in the Brazilian Football Championship

*Gabriel Gomes Ribeiro<sup>1</sup>, Lilia Carolina Carneiro da Costa<sup>1</sup>, Paulo Henrique Ferreira<sup>1</sup>*

<sup>1</sup> Department of Statistics, Federal University of Bahia

**Abbreviated abstract:** With the aim of predicting the results of the matches of the 2021 Brazilian championship first division, We used two structures of Bayesian inference Poisson models, a hierarchical and a non-hierarchical model, proposed by Leonardo Blanco (2021). To assess the quality of the forecasts, the De Finetti measure is used.

## Related publications:

– Blanco *et al*, Axioms 2021, 10, 276.



[gabrielrbr76@gmail.com](mailto:gabrielrbr76@gmail.com)



4th Conference on  
**Statistics and  
Data Science**  
Salvador, Brazil (online)  
December 1-3, 2022

# Problem, Data, Previous Works

Football is considered one of the most unpredictable sports in the world, where often the team considered to be the weakest manages to overcome another team considered to be stronger and it is in the interest of many to be able to predict the result of a match.

The 2021 Brazilian championship first division was disputed by 20 teams that faced each other 2 times throughout 380 matches

842 goals were scored, 483 of which were made by the home team and 359 by the away team

		AG				
		0	1	2	3	4+
HG	0	44	32	16	6	5
	1	59	48	21	5	1
	2	31	37	20	5	1
	3	15	15	6	1	1
	4+	2	3	4	2	0

Table 1: Championship Most Frequent Scores  
HG: Home Goals; AG: Away Goals

Blanco (2021) uses a Bayesian Poisson model and proposes two model structures for predicting the results of the Chilean football championship in the year 2020. The structures are hierarchical and non-hierarchical models.



[gabrielrbr76@gmail.com](mailto:gabrielrbr76@gmail.com)



4th Conference on  
**Statistics and  
Data Science**  
Salvador, Brazil (online)  
December 1-3, 2022

# Methods

Two Poisson models were fitted, with a hierarchical and a non-hierarchical structure

Non-Hierarchical Model:

$$\log(\theta_{g1}) = \beta_{home} + \beta_{1(g)}att_{home(g)} + \beta_{2(g)}def_{away(g)}$$

$$\log(\theta_{g2}) = \beta_{3(g)}att_{away(g)} + \beta_{4(g)}def_{home(g)}$$

$$\beta_{home} \sim Normal(0, 0.0001),$$

$$\beta_{1(g)} \sim Normal(0, 0.0001),$$

$$\beta_{2(g)} \sim Normal(0, 0.0001),$$

$$\beta_{3(g)} \sim Normal(0, 0.0001),$$

$$\beta_{4(g)} \sim Normal(0, 0.0001).$$

Hierarchical Model:

$$\log(\theta_{g1}) = \beta_{home} + \beta_{att(g1)}att_{(g1)} + \beta_{def(g2)}def_{(g2)}$$

$$\log(\theta_{g2}) = \beta_{att(g2)}att_{(g2)} + \beta_{def(g1)}def_{(g1)}$$

$$\beta_{home} \sim Normal(0, 0.0001),$$

$$\beta_{att(g)} \sim Normal(\mu_{att}, \tau_{att}),$$

$$\beta_{def(g)} \sim Normal(\mu_{def}, \tau_{def}),$$

$$\mu_{att} \sim Normal(0, 0.0001),$$

$$\tau_{att} \sim Gama(0.1, 0.1),$$

$$\mu_{def} \sim Normal(0, 0.0001),$$

$$\tau_{att} \sim Gama(0.1, 0.1).$$



[gabrielrbr76@gmail.com](mailto:gabrielrbr76@gmail.com)



4th Conference on  
**Statistics and  
Data Science**  
Salvador, Brazil (online)  
December 1-3, 2022

# Results and Conclusions

## Non-Hierarchical Model:

- The first model presented the best results predicting the matches 311 to 320
- De Finetti mean measure in the 80 predicted matches – 0.637
- The model predicted correctly 41.25% of the games

## Hierarchical Model:

- The second model presented the best results predicting the matches 351 to 360
- De Finetti mean measure in the 80 predicted matches – 0.650
- The model predicted correctly 43.75% of the games

Home Team	Away Team	PH	PD	PA	Result	Correct	DeFinetti
AME	GRE	0.44475	0.3090	0.24625	3 - 1	YES	0.4644
BRA	FOR	0.39025	0.23625	0.3735	3 - 0	YES	0.5671
INT	CAP	0.44725	0.2425	0.31025	2 - 1	YES	0.46059
COR	CUI	0.35725	0.35050	0.299225	3 - 2	YES	0.6213
SAO	FLA	0.16050	0.26125	0.57825	0 - 4	YES	0.2718
FLU	PAL	0.2560	0.25275	0.49125	2 - 1	NO	0.8587
CEA	SPO	0.38225	0.38625	0.23150	2 - 1	NO	0.5843
CHA	JUV	0.27225	0.26100	0.46675	0 - 2	YES	0.4266
CAP	CAM	0.1555	0.2185	0.6260	0 - 1	YES	0.2118
GRE	BRA	0.20575	0.24450	0.54975	3 - 0	NO	0.9928

Table 2: Predict for Matches 311 to 320 - Non-Hierarchical Model  
PH: Prob. of victory for the home team; PD: Prob. of draw; PA: Prob. of victory for the away team

Home Team	Away Team	PH	PD	PA	Result	Correct	DeFinetti
JUV	BRA	0.27975	0.2880	0.43225	1 - 0	NO	0.7885
FLA	CEA	0.5195	0.25425	0.22625	2 - 1	YES	0.3467
AME	CHA	0.4115	0.3000	0.2885	3 - 0	YES	0.5196
CUI	PAL	0.27125	0.27475	0.4540	1 - 3	YES	0.4472
BAH	CAM	0.24825	0.2640	0.48775	2 - 3	YES	0.3737
GRE	SAO	0.37025	0.3180	0.31175	3 - 0	YES	0.5949
CAP	CUI	0.36475	0.3110	0.32425	1 - 0	YES	0.6054
CHA	ACG	0.3205	0.3215	0.3580	0 - 1	YES	0.6182
SPO	FLA	0.16775	0.2410	0.59125	1 - 1	NO	0.9538
FOR	JUV	0.38375	0.28225	0.3340	1 - 0	YES	0.5710

Table 3: Predict for Matches 351 to 360 - Hierarchical Model  
PH: Prob. of victory for the home team; PD: Prob. of draw; PA: Prob. of victory for the away team

