Hierarchical models

Urša Zrimšek

Introduction

Our goal is to build a hierarchical model on basketball free throw shooting data that is suitable for answering the following two questions:

- With what certainty can you claim that the best guard (Player #20) is on average a better free throw shooter than the best forward (Player #3)?
- We take a random guard and a random forward, each of them shoots one free throw. What are the probabilities of A) the guard winning (guard makes the shot, forward misses the shot), B) the forward winning (guard misses the shot, forward makes the shot), and C) a tie (both players either miss the shot or hit the shot). We will also compute the probabilities for other combinations (guard vs. center and forward vs. center).

For purposes of this analysis we will only look at free throw shots made on the traditional rim size and discard variables Angle and ThrowNum. We are left with 2700 rows, corresponding to shots, each described with 3 columns: `PlayerID` - 45 distinct players, `Made` - column of zeros and ones, indicating if the player made that shot and `Position` - the position of the player, C for center, G for guard and F for forward.

Model

We decided to model **each player's shots** with **Bernoulli distribution**. The target variable is whether the i-th shot of player p, y_{pi}, was made, and the Bernoulli parameter θ_p will represent player p's probability that he makes the shot:

$$y_{pi}|\theta_p \sim Bernoulli(\theta_p).$$

We will combine the shot probabilities of players that play on the same position. We can model **means of each group** probabilities with **Beta distribution**. To do that, we will need to reparametrize α and β with mean probability $\mu_g = \frac{\alpha_g}{\alpha_g + \beta_g}$, g being one of the positions, G, C or F. We first tried setting $\tau_g = \alpha_g + \beta_g$, total count, as the other parameter, but the model diverged too often. That's why we took its reciprocal value $\lambda_g = \frac{1}{\tau_g}$ and with it as a parameter, the model converged nicely. We are not interested into inference of this value, so this change doesn't effect our further analysis. The main parameter we are interested in here (not directly for answering the questions, but it still carries important information) is the mean probability, μ_g, that represents the probability that a player on position g makes the shot. Let's write our model:

$$\theta_p|\mu_g, \lambda_g \sim Beta\left(\frac{\mu_g}{\lambda_g}, \frac{1-\mu_g}{\lambda_g}\right).$$

We decided to model p_g and λ_g separately for each position, as we can't say that players on different positions have the same accuracy or shot consistency (that would express in model's variance and consequentially in λ).

On this point we also tried adding an extra layer to the model, connecting all three positions, but it didn't change the model much, so we decided to stick with the simpler one, as we would otherwise just unnecessarily raise the model's complexity. We didn't use any priors, fitted the model on 4 chains with 1000 warm-up and 1000 sampling iterations and checked standard diagnostics.

Answering initial questions

To answer the first question, we can just compare inferred probabilities θ_p for players $p \in \{3, 20\}$. We can claim that the **best guard** (Player #20) is **on average a better free throw shooter** than the **best forward** (Player #3) with $\mathbf{0.921} \pm 0.005$ **certainty**. The probabilities are shown on Figure 1.

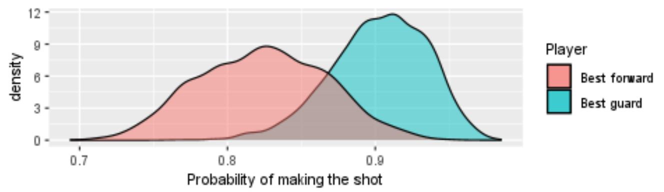

Figure 1. **Inferred probabilities of players #3 (best forward) and #20 (best guard).**

To answer the second question, we draw (1000×) a random player from each position, a probability from his chains and calculate means ± standard deviations of next probabilities:

$$P(p_1 \text{ wins}) = \theta_{p_1} \cdot (1 - \theta_{p_2}),$$
$$P(p_2 \text{ wins}) = (1 - \theta_{p_1}) \cdot \theta_{p_2},$$
$$P(\text{tie}) = \theta_{p_1} \cdot \theta_{p_2} + (1 - \theta_{p_1}) \cdot (1 - \theta_{p_2}).$$

The answers can be seen in Table I. We can see that in all combinations the **most probable outcome is tie**. For better explanation we showed the sampled probabilities on Figure 2.

Table I
Probabilities of outcomes of one free throw

	First wins	Second wins	**Tie**
G vs. F	0.205 ± 0.001	0.122 ± 0.002	$\mathbf{0.673} \pm 0.001$
G vs. C	0.267 ± 0.002	0.110 ± 0.002	$\mathbf{0.623} \pm 0.002$
F vs. **C**	0.167 ± 0.001	0.240 ± 0.002	$\mathbf{0.593} \pm 0.001$

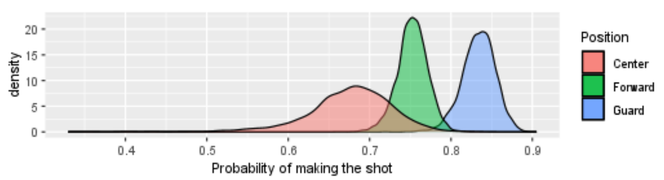

Figure 2. **Sampled probabilities of random players (θ_{p_i}) on different positions.**

Discussion

Even though our data consists of 2700 throws, they were only made by 45 players: 19 guards and forwards and **only 7 centers**. So when talking about our results, we must emphasize that they hold for these specific players, not all on those positions.

When deciding for parameters for Beta distribution, we let μ and λ be different for distinct classes. Since we have low number of players on center position, it could be better to set $\lambda_C = \lambda_G = \lambda_F = \lambda$. But that would be a further assumption that we didn't want to make without more domain knowledge.