# Biases in public draws for major sporting competitions

by


Jiahang Liu
(ID: 5647941)
Supervised by Prof. Gareth Roberts

Department of Statistics
University of Warwick
Coventry CV4 7AL
United Kingdom

Email: u5647941@live.warwick.ac.uk


6 September 2025

Report submitted in partial fulfillment of the requirements for the degree of MSc in Statistics in the University of Warwick

# Abstract

This dissertation examines fairness and probabilistic biases in the FIFA World Cup group stage draw. Existing sequential procedures used by FIFA and UEFA are shown to deviate from the uniform distribution whenever geographical restrictions are applied, even without host-nation privileges.

To overcome these limitations, we implement an improved version of Roberts' Multiple-Rejection Solution. Rather than enumerating all valid completions, which is computationally infeasible, our method estimates conditional assignment probabilities using rejection sampling at each partial stage of the draw. This modification maintains theoretical uniformity while making the procedure efficient enough for practical use.

The algorithm is implemented in Python and validated through extensive simulations. Results from 1,000,000 rejection-sampling draws and 10,000 multiple-rejection draws are statistically indistinguishable, confirming the correctness of the approach. With an appropriate choice of simulation size (N=50), the complete draw can be executed within seconds, making it suitable for live broadcasts where transparency and efficiency are essential.

We further apply the method to the 2018 World Cup and discuss the challenges and implications for the expanded 2026 format. These applications highlight the flexibility of the procedure and its potential as a fair, efficient, and transparent alternative to current tournament draw mechanisms.

# Acknowledgements

I would like to express my sincere gratitude to my supervisor, Prof. Gareth Roberts, for his invaluable guidance, encouragement, and insightful feedback throughout the process of writing this dissertation.

I am also deeply thankful to my parents, who not only supported me throughout my studies but also helped test the website implementation and provided valuable suggestions for improvements.

Finally, I acknowledge the assistance of AI tools, including ChatGPT and Grammarly, which I used to improve my writing clarity, polish the language, and learn how to use Streamlit to transform my Python code into a form that can be displayed on a webpage.

# Contents

# Chapter 1

# INTRODUCTION

Football competitions can generally be divided into two main categories: leagues and cup tournaments. A league is a long-term competition in which all teams in the league play against each other. Take, for example, the English Premier League. It consists of 20 different teams, and each team plays against the other 19 teams twice, once at home and once away, resulting in a total of 38 rounds of matches and 380 games in a season. The champion is determined based on the total number of final points (3 points for a win, 1 point for a draw, and none for a loss).

This system has the advantage of providing a more comprehensive assessment of each team's performance throughout a long season, thereby minimizing the impact of a single match on the final standings. However, it also means that the competition lasts for an extended period (typically around ten months) and may result in a lack of suspense toward the end of the season. For example, in the 2024–2025 season, Liverpool secured the championship with a ten-point lead over second-placed Arsenal, clinching the title with four rounds remaining.

In contrast, the core of cup tournaments lies in their knockout format, where advancement is decided based on the result of a single match at a neutral venue or two-legged home-and-away ties. This allows the tournament to be completed in a much shorter time frame (usually around one month) and introduces a high-stakes, suspenseful atmosphere. In this format, weaker teams can occasionally defeat more powerful opponents with a single outstanding performance. However, in a league, a weaker team must maintain consistent excellence throughout the entire season to outperform stronger teams - an outcome that is clearly less probable.

However, the knockout format also has its drawbacks. Popular, high-profile teams are more likely to be eliminated early by underdogs, potentially reducing audience interest.

Modern football cup competitions often seek a balance between these two formats. Whether at the national team level (e.g., the FIFA World Cup, UEFA European Championship, AFC Asian Cup, or Copa América) or the club level (e.g., UEFA Champions League, Copa Libertadores), most tournaments adopt a format that includes a group stage—similar to a mini-league—followed by knockout rounds. This structure reduces the likelihood of early elimination for strong teams, preserves suspense until the final stages, and offers underdogs more opportunities to achieve strong results.

Due to time constraints, cup tournaments cannot have every team play against all others as in a league. Therefore, they rely on draws to determine the matchups for the initial stage. To enhance spectacle and maintain transparency, these draws are often broadcast live and include various restrictions to increase viewer engagement. This makes the draw process more complex.

This article focuses on the group stage draw of the FIFA World Cup. It discusses potential fairness issues in the current draw procedures and proposes a brand new, fully transparent, and broadcast-friendly method to ensure fairness in the draw process.

## 1.1 Previous Literature

As one of the most popular sporting events in the world, the FIFA World Cup has continuously evolved toward a more fair, transparent, and balanced competition format—both in terms of team strength and geographic representation. Among all aspects of the tournament structure, the group stage draw is especially critical. Since the World Cup adopted a more mature format in 1982, the draw mechanism and procedures have become a frequent topic of discussion among fans. However, because the rules for group stage draws were newly established at the time, many issues remained in the procedures.

A classic example occurred during the 1990 World Cup draw (Csató, 2025).[2] FIFA released the detailed draw procedure 24 hours before the event, but the host nation, Italy, strongly opposed it. Within those 24 hours, FIFA revised the procedure. The change significantly reduced Italy's chances of being drawn against a stronger South American team from Pot 2, from 80

Numerous academic papers have examined various issues with the World Cup draw model from diverse perspectives. Before 2014, except for the seeded teams, the remaining teams were grouped into pots mainly based on geographic location. This was intended to support the principle of continental separation in the group stage (i.e., no group could have more than one team from the same continent, except for Europe, which was allowed a maximum of two teams per group). Guyon (2015) argued that this pot allocation method often led to competitive imbalance among groups. [5]He proposed a pot division based on FIFA World Rankings, a suggestion that was eventually adopted in the 2018 World Cup.

Regarding the 2022 World Cup draw, Csató (2023) questioned the allocation of teams into pots. Since some intercontinental playoffs had not yet been played before the draw, three tournament spots were yet to be determined. All the teams competing for those spots were placed in Pot 4, but Csató (2023) argued that this was an unreasonable approach.[1] Csató (2025) further pointed out that FIFA's current draw procedure leads to probabilistic imbalances, and he demonstrated that even slight imbalances can result in significant deviations in final group results.[2]

To address the issue of imbalance, Guyon (2015) proposed an improved draw method.[5] However, the new approach still yielded uneven results and required substantial computational resources, rendering it unsuitable for live broadcast demonstrations. Roberts (2024) proposed three alternative solutions, showing that all three could eliminate bias in the draw and could be implemented transparently and efficiently during live media coverage.[8] However, only the first two methods were actually demonstrated by Roberts.

This paper primarily explains the biases introduced by the current FIFA draw procedure and argues that even adopting the UEFA Champions League round-of-16 draw format, as used before the 2023–2024 season, does not solve the problem. It then introduces a third draw model proposed by Roberts, proves that this method avoids any bias, implements the procedure in code, and rigorously verifies its validity. Finally, we discuss several practical considerations related to the draw procedure presented in this paper.

## 1.2 The 2022 FIFA World Cup Group Draw

Since the 2018 FIFA World Cup, the group stage draw procedure has been further refined. The tournament features 32 national teams, including 31 that qualify through competition and the host nation, which qualifies automatically (Qatar in 2022). These 32 teams are allocated into eight groups of four, with the teams in each group playing one another in the group stage.

The allocation of teams is subject to the following constraints. The 32 teams are first divided into four seeding pots. Pot 1 consists of the host nation together with the seven highest-ranked teams according to the official FIFA World Ranking. Pot 2 contains the following highest-ranked teams, and so on.

The 2022 FIFA World Cup presented some special circumstances: at the time of the draw, three teams had yet to be decided through various playoffs, and each of these undecided slots was placed in Pot 4. Furthermore, each team belongs to one of six continental confederations: UEFA (Europe; 13 qualifiers, hereafter EU), CONMEBOL (South America; 4 or 5 qualifiers, hereafter SA), CONCACAF (North and Central America; 3 or 4 qualifiers, hereafter NA), AFC (Asia; 5 or 6 qualifiers, hereafter AS), CAF (Africa; 5 qualifiers, hereafter Af), and OFC (Oceania; 0 or 1 qualifier, hereafter OC).

**Pot 1:** Qatar [AS], Belgium [EU], Brazil [SA], France [EU], Argentina [SA], England [EU], Portugal [EU], Spain [EU]

**Pot 2:** Denmark [EU], Netherlands [EU], Germany [EU], Switzerland [EU], Croatia [EU], Mexico [NA], United States [NA], Uruguay [SA]

**Pot 3:** IR Iran [AS], Serbia [EU], Japan [AS], Senegal [AF], Tunisia [AF], Poland [EU], Korea Republic [AS], Morocco [AF]

**Pot 4:** Wales/Scotland/Ukraine [EU], Peru/United Arab Emirates/Australia [SA, AS], Costa Rica/New Zealand [NA, OC], Saudi Arabia [AS], Cameroon [AF], Ecuador [SA], Canada [NA], Ghana [AF]

The three undecided team slots in Pot 4 arose due to delays caused by the COVID-19 pandemic and the war in Ukraine, meaning that not all teams were confirmed at the time of the draw. Placeholder entries were therefore used, two of which corresponded to different possible confederations; their allocation needed to comply with the geographic restrictions of both potential confederations.

Under these specifications, the group composition restrictions were that each group must contain exactly one team from each pot, together with one or two teams from UEFA, and no more than one team from any other confederation.

# Chapter 2

# DEFICIENCIES IN THE CURRENT DRAWING PROCESS

## 2.1 Uniform Simulation and rejection sampling

First, we need to define the desired draw result in this paper, namely, the so-called uniform draw simulation. In this paper, 'uniform' refers to the case in which all valid results occur with equal probability. Here, a 'result' refers to the complete result of the draw, that is, the allocation of the 32 teams (three of which are to be determined through intercontinental play-offs after the draw ceremony) into eight groups.

It is evident that if we only consider the seeding rules and disregard the confederation separation principle, teams within the same pot would have identical probabilities of facing any given team from another pot. This is equivalent to randomly shuffling all eight teams in each pot and then filling Groups A-H in order. However, due to the principle of confederation separation and the requirement that the host nation, Qatar, be placed in Group A, some results generated through such a random allocation are invalid. Consequently, teams within the same pot may have different probabilities of facing certain opponents from other pots.

For example, as the host nation of the 2022 FIFA World Cup, Qatar was automatically assigned to Pot 1. The other teams in Pot 1 consisted of the seven highest ranked teams among the 31 remaining finalists (based on the FIFA World Ranking). This meant that Qatar was significantly weaker than the other Pot 1 teams. For teams in other pots, being placed in Group A with Qatar was highly desirable, as it meant avoiding the other strong Pot 1 teams. Consider Pot 2, which contained the United States from CONCACAF and Uruguay from CONMEBOL. Uruguay could not be drawn into the same group as Brazil or Argentina from Pot 1 (both also from CONMEBOL). Therefore, Uruguay had only six possible Pot 1 opponents, whereas the United States could face all eight Pot 1 teams. As a result, the probability that the United States would be drawn into Group A was slightly lower than that of Uruguay. Although this may seem disadvantageous for the United States, it is reasonable given the confederation separation principle and consistent with the uniformity definition pursued in this paper.

It should be noted that, in this paper, uniform does not mean that any two teams eligible to be drawn into the same group have an equal probability of actually being placed together. Rather, it means that any valid and complete draw result has the same probability of occurrence.

To obtain a uniform result, the simplest approach is to adopt a method similar to rejection sampling.[7] As discussed above, randomly generated draw results that meet only the seeding rules will sometimes satisfy all additional restrictions and sometimes not. If a randomly generated result satisfies all restrictions, we accept it; otherwise, we reject it. It is well known (Devroye, 1986; Guyon, 2015) that this rejection sampler generates a sample uniformly distributed over all valid draws.[3][5] And it is straightforward to prove, by using the conditional probability formula, that the accepted results will occur with equal probability. Let $\Omega$ denote the set of all draw results that satisfy the seeding rules (possibly including invalid results). Let $V \subset \Omega$ denote the set of all valid draw results (those satisfying the seeding rules and all additional restrictions, such as the confederation separation principle and the fixed placement of the host nation). On $\Omega$, results are generated with equal probability, each result has probability $1/|\Omega|$.

The rejection sampling procedure is as follows:

Randomly and uniformly generate a result from $\Omega$.

If the result belongs to V, accept it; otherwise, reject it and repeat the process.

$$P[X = x \mid X \in V] = \frac{P[X = x \ \cap \ X \in V]}{P[X \in V]} = \frac{P[X = x\ ]}{P[X \in V]} = \frac{\frac{1}{|\Omega|}}{\frac{|V|}{|\Omega|}} = \frac{1}{|V|}.$$

This method has the following advantages. First, it is easy to implement computationally: we simply generate a random unrestricted draw result and then check whether it satisfies all restrictions. Moreover, it can be used in other tournaments; it only needs to change the acceptance criteria. However, compared to the current draw procedure, it also has certain drawbacks. Most importantly, it suffers from a very low acceptance rate (approximately 1/560, see Roberts 2024), meaning that to obtain one valid result, we would expect to simulate about 560 times.[8] This makes it virtually impossible to replicate the current draw ceremony, in which football celebrities manually conduct the draw in a live broadcast. If the draw were conducted entirely by computer and only the result announced, it could invite skepticism from fans—particularly from supporters of teams that receive less favorable results. Therefore, this method is suitable for research purposes, to generate baseline results satisfying the uniformity requirement, but not for replacing the existing public draw process.

## 2.2 FIFA's Sequential Draw and Its Limitations

The FIFA group stage draw for the 2022 World Cup, held on 1 April 2022, proceeded as follows (FIFA, 2022).[4] First, the host nation, Qatar, was automatically assigned to position 1 of Group A. The remaining teams in Pot 1 were then drawn sequentially, one at a time, and allocated to Groups B through H in order. Next, a team from Pot 2 was drawn at random and assigned to the next available group, defined as the first group in which its placement would not violate any geographic restrictions—either immediately or in a way that would later make it impossible to complete the draw validly. This process was then repeated for Pot 3 and subsequently for Pot 4 to complete the draw.

Each random selection was conducted by a distinguished football personality, who would uniformly at random select one ball from a bowl containing all teams in the given pot that had not yet been allocated. The ball was then opened to reveal the team's identity. After a team was assigned to a group, another celebrity would draw a ball from a separate bowl containing the four position numbers for that group, without replacement, to determine the team's position within the group. For Pot 1, this draw necessarily yielded the single red ball among three white balls, corresponding to the first position in the group. Since the assignment of positions within a group affects only the match schedule and not the group composition, it is disregarded in the present analysis.

Before explaining the discrepancy between FIFA's procedure and the uniform distribution, we first need to introduce a concept. Regardless of whether the result is generated by rejection sampling to obtain the uniform distribution, or by FIFA's official procedure, for teams in the same pot from the same continent (except for the host nation), we call them symmetric. For symmetric teams, their associated probabilities—for example, the probability of being assigned to a certain group or being drawn into the same group with another team—are equal. This is straightforward to prove: the restrictions in the draw depend only on pots and confederations, not on the identity of individual teams. If two teams are symmetric, exchanging them yields another valid result, and this substitution does not alter the expected probabilities of subsequent assignments.

It is easy to see that the results generated by the FIFA procedure deviate from the uniform distribution we desire. For example, when filling Pot 2 into Group A, since Qatar as host is fixed in Group A and does not conflict with any Pot 2 team (because Pot 2 contains no AFC team), it is equally likely that the eight Pot 2 teams are drawn into Group A under the FIFA procedure, each with probability 12.5%.

However, as explained above, Uruguay (CONMEBOL) and the United States (CONCACAF) both face continental restrictions elsewhere, since Pot 1 contains two CONMEBOL teams (Brazil and Argentina) but no CONCACAF teams. Therefore, the true uniform probabilities should differ. Using rejection sampling experiments, we find that the probabilities of Uruguay and the United States being placed into Group A are approximately 13.66% (95%CI=[13.00%,15.72%]) for Uruguay and 9.06% (95%CI=[9.00%,9.11%]) for USA.

To illustrate why FIFA's procedure produces bias, consider a simplified model: suppose there are only two pots and four groups. Pot 1: Qatar (Q) [Af], France (F) [Eu], England(E)[Eu], Argentina (A) [SA] Pot 2: Mexico (M) [NA], United States(USA) [NA], Germany(G)[EU] , Uruguay (U) [SA]

Following World Cup rules (ignoring the requirement that each group must contain at least one UEFA team, which is impossible here, since there are four groups but only three UEFA teams), the only restriction is that Argentina and Uruguay cannot be in the same group and Qatar as host is fixed in Group A. Thus, Uruguay can only be grouped with Qatar, France or England. Once Uruguay is paired with one of these three teams, the other six teams can be allocated at random. Hence, under the uniform distribution, the probability that Uruguay is grouped with each of Qatar, France, or England is equal:

$$P_u[U \sim Q] = P_U[U \sim E] = P_U[U \sim F] = \frac{1}{3}$$

However, according to the sequential procedure of FIFA, each Pot 2 team has an equal probability of $\frac{1}{4}$ being placed in Group A. For France and England, which are symmetric, their probabilities of being grouped with Uruguay must be equal. A more detailed case analysis shows that each equals $\frac{3}{8}$. Thus Qatar, as host in Group A, faces Uruguay with reduced probability compared to the uniform benchmark, while France and England have their probabilities increased equally.

One might suspect that this bias arises mainly from the host restriction (Qatar being fixed in Group A). To check this, we ignore the host restriction in the simplified model. In that case, Qatar, France, and England are symmetric, since only Argentina and Uruguay trigger geographical constraints. Then, without the host restriction, the procedure indeed produces uniform results.

However, if we add further constraints, for example, forbidding UEFA teams from being in the same group, the problem reappears. In this modified example, Germany (UEFA) from Pot 2 can only be grouped with Qatar or Argentina. Let us examine the probability that Germany is paired with Qatar. Under the uniform distribution, if Germany is with Qatar, then Uruguay must be grouped with one of France or England, while the remaining teams can be freely allocated. There will be $2 \times (4!) \times 2 \times 2$ valid results. And if Germany is with Argentina, other six teams will never break geographical constraints, which means there will be $2 \times (4!) \times 3 \times 2$ valid results. Thus, under the uniform distribution,

$$P_U[G \sim Q] = \frac{2 \times (4!) \times 2 \times 2}{2 \times (4!) \times 2 \times 2 + 2 \times (4!) \times 3 \times 2} = \frac{2}{5}.$$

By contrast, under FIFA's sequential procedure, the calculation is more involved. By placing France and England before or after Qatar and Argentina in the draw order, we compute

$$P_{FIFA}[G \sim Q] = \frac{31}{72} \approx 43.06\%.$$

This matches well with computer simulations (one million trials yielding 43.05%).

Thus, once the situation becomes even slightly more complex, simply removing the host restriction does not restore uniformity. The bias arises not primarily from the host rule, but from the sequential structure of FIFA's procedure itself.

## 2.3 UEFA's Sequential Draw and Its Limitations

Before the 2023-2024 season, the UEFA Champions League Round of 16 draw was subject to constraints similar to those of the FIFA World Cup group stage draw. The objective was to pair the eight group winners with the eight group runners-up such that teams from the same group or the same domestic league would not face each other. This setup resembles a draw with only two pots and constraints analogous to the continental restrictions seen in the World Cup group draw.

However, UEFA uses a fundamentally draw procedure (UEFA, 2023).[9] Similar to FIFA, the draw was conducted by football celebrities who randomly selected balls from bowls containing the candidate teams. The key distinction was that a runner-up team was first drawn at random. Then, the set of possible opponents for that runner-up, group winners who were neither from the same group nor from the same national league, were identified, and the corresponding balls representing these teams were placed into a separate bowl for a subsequent random draw. The team drawn in this second step was then paired with the runner-up.

Interpreted through the lens of the FIFA group draw process, this approach can be seen as first randomly assigning a team from Pot 2 (runners-up) to Group A, then randomly selecting an eligible team from Pot 1 (group winners) to pair with it in Group A. This process is repeated sequentially for each remaining group. Such a method avoids the issue that can arise in the FIFA draw where a drawn team cannot be placed in the earliest available group, ensuring that all drawn pairings are valid at the time of selection.

Klößner (2013) has shown the bias of UEFA's procedure.[6] And if we compare UEFA's procedure for the Champions League draw with Roberts'(2024) Multiple Balls Method (different numbers of balls representing different teams are placed in the bowl, instead of putting one ball for each team, in order to reflect the varying probabilities of teams being selected), we notice a critical difference.[8] In the UEFA procedure, when placing balls in the bowl to represent the potential opponents of a team, the rule is not to assign a distinct ball to each possible opponent. Instead, multiple balls representing the same opponent are inserted to reflect the different probabilities with which that opponent may be drawn. In effect, UEFA assumes that all potential opponents are equally likely to be drawn. However, once the constraints are taken into account, this assumption is clearly inaccurate. For instance, in the simplified model described in Section 2.2, we have already shown that under the joint constraints preventing Uruguay and Argentina from being drawn into the same group, and simultaneously requiring that Germany cannot be drawn into the same group as either England or France, the probabilities of Germany being grouped with its two possible opponents—Qatar and Argentina—are not equal. In fact, the probability of Germany being placed in the same group as Qatar is $\frac{2}{5}$, while the probability of being grouped with Argentina is $\frac{3}{5}$.

# Chapter 3

# A MULTIPLE-REJECTION SOLUTION

## 3.1 The Original Multiple-Rejection Method and Its Limitations

In addition to the Multiple Balls Method discussed above, Roberts (2024) also introduced an alternative approach called the Multiple-Rejection Solution. This method is of particular interest in the present study, and we implement a modified version that ensures the generated draws follow a uniform distribution while reducing computational complexity and improving practical feasibility.[8]

We first review Roberts' (2024) Multiple-Rejection Solution.[8] Consider a partial draw and let $n_j$ denote the number of valid ways to complete the draw if team $j$ is assigned to the next position $p$. Define the set of feasible teams for position $p$ as

$$T_p = \{j \mid n_j > 0\},$$

and let

$$m_p = \max_{j \in T_p} n_j.$$

For each team $j \in T_p$, we then generate a random variable.

$$G_j \sim \text{Geom}\left(\frac{n_j}{m_p}\right).$$

The assignment of a team to position $p$ proceeds according to the following steps:

1. Select a team uniformly at random from $T_p$ (for example, corresponding to drawing a ball from a bowl).
2. If the selected team $j$ has already been drawn $G_j$ times, assign it to position $p$ and proceed to the next position $p+1$.
3. Otherwise, return to Step 1 and select another team uniformly at random.

It is clear that the computation of $n_j$ forms the foundation of the entire method. However, as Roberts (2024) emphasized, this task can be extremely challenging. According to his analysis (Roberts, 2024), for the 2022 FIFA World Cup draw, there exist approximately $5.9 \times 10^{14}$ valid results.[8] When assigning the teams in Pot 1, these $n_j$ values are identical. The reason is that for any valid result, if we exchange all four teams between any two groups other than Group A (for instance, by swapping the four teams of Group C with those of Group D), the resulting allocation remains valid. This is because the group-stage constraints (apart from the host nation being fixed in Group A) only apply within individual groups. Consequently, after partially completing the allocation of Pot 1, when attempting to place the next Pot 1 team $t_1$ into position $p$, every valid completion of the draw involving $t_1$ corresponds one-to-one with another valid completion involving any other remaining Pot 1 team. Thus, all the $n_j$ for the remaining Pot 1 teams are equal. This observation implies that the Pot 1 allocation can be conducted entirely through FIFA's standard sequential procedure without compromising uniformity.

In contrast, once Pot 1 has been fully allocated, difficulties arise in the assignment of Pot 2 teams. When attempting to place one of the eight Pot 2 teams into Group A, under a uniform distribution, the probabilities of assignment differ across teams. In other words, the values of $n_j$ are no longer equal, and explicit computation becomes necessary. However, this requires an enormous computational effort: given that approximately.

$$\frac{5.9 \times 10^{14}}{7!} \approx 1.17 \times 10^{11}$$

valid results remain after the allocation of Pot 1, one would need to enumerate all such results and count, for each Pot 2 team, the number of instances in which it is placed into Group A. Clearly, this calculation is impossible in real time during a live draw. Even if one were able to find one million valid results per second, it would still take more than 27 hours to exhaustively generate all valid results. Moreover, after this enumeration, it would still be necessary to record, for each result, which Pot 2 team was assigned to Group A.

## 3.2 Estimating the Numbers of Valid Results

Although obtaining the exact values of $n_j$ is computationally infeasible, we can attempt to avoid these calculations. Since the ratio $n_j/m_p$ is what is ultimately required, we can instead make use of the probability

$$p_j = \frac{n_j}{\sum_{k \in T_p} n_k},$$

which represents the conditional probability that team $j$ is placed into position $p$ given the current partial draw. If we let $q_p$ denote the largest among the nonzero probabilities $\{p_j\}$, then

$$\frac{p_j}{q_p} = \frac{\frac{n_j}{\sum_k n_k}}{\frac{m_p}{\sum_k n_k}} = \frac{n_j}{m_p}.$$

Although the exact computation of $p_j$ still depends on the intractable evaluation of $n_j$, it is possible to estimate $p_j$. Following a strategy similar to that proposed by Roberts (2024) in the Multiple Balls Method, we can generate $N$ valid results under the uniform distribution by applying rejection sampling to the current partial draw, and then compute the empirical frequency $\hat{p}_j$, which serves as an estimator for $p_j$.[8] Provided that $N$ is chosen to be sufficiently large yet computationally manageable, this approach is straightforward to implement. Moreover, it is readily verified that the estimators $\hat{p}_j$ are unbiased and satisfy $\sum_j \hat{p}_j = 1$. Therefore, this method significantly enhances the practical feasibility of implementing the draw procedure in real-world settings.

## 3.3 Complete Draw Procedure

Combining the previous ideas together gives the following algorithm for choosing the team in the next position P of a uniform group draw with distribution U, by using multi-rejection solution, as follows:

1. Select a positive integer valued algorithm parameter $N$.
2. Based on the current partial draw, generate $N$ valid results using rejection sampling.
3. From these results, compute the empirical frequencies $\{\hat{p}_j\}$ for all teams that appear in position $p$, with $\hat{p}_j > 0$, and determine $q_p = \max_j \hat{p}_j$.
4. For each candidate team $j$, generate a random variable $G_j \sim \mathrm{Geom}(\hat{p}_j/q_p)$.
5. Uniformly at random (e.g., by drawing a ball from the bowl), select one team among all teams that have appeared in position $p$ in the $N$ simulated results.
6. If the selected team $j$ has already been chosen exactly $G_j$ times, then assign team $j$ to position $p$ and proceed to the next position.
7. Otherwise, return to Step 5 and repeat the uniform random selection.

In the following, we will prove that applying this procedure at each position can get a uniform distribution.

**Proposition.** At a fixed position $p$, let the candidate set be $T_p$. Suppose the target probability that each candidate team $j \in T_p$ is selected is $p_j$, with $\sum_{j \in T_p} p_j = 1$. Under the multiple-rejection method, the probability that team $j$ is eventually chosen is exactly $p_j$.

**Proof.** First, define

$$q_p = \max_{j \in T_p} p_j, \quad r_j = \frac{p_j}{q_p}, \quad j \in T_p.$$

For each team $j$, we independently generate a geometric random variable $G_j \sim \mathrm{Geom}(r_j)$. A geometric random variable means that we perform independent trials, each succeeding with probability $r_j$, and $G_j$ records the round in which the first success occurs.

In the algorithm, in each round we choose one team $J^{(t)}$ uniformly at random from $T_p$. If the chosen team is $j$, we check whether it has already been selected $G_j$ times. If yes, then the algorithm stops and $j$ is chosen; otherwise, the process continues to the next round.

This mechanism is equivalent to an acceptance–rejection sampling procedure: in each round, a team $j$ is chosen uniformly at random with probability $\frac{1}{|T_p|}$, and if selected, it is accepted with probability $r_j$. If not accepted, the process repeats in the next round.

The probability that in a given round team $j$ is both chosen and accepted is

$$\mathbb{P}(\text{choose and accept } j) = \frac{1}{|T_p|} \cdot r_j.$$

The probability that any team is accepted in this round is

$$\mathbb{P}(\text{accept}) = \sum_{k \in T_p} \frac{1}{|T_p|} \cdot r_k = \frac{1}{|T_p|} \sum_{k \in T_p} \frac{p_k}{q_p} = \frac{1}{|T_p|} \cdot \frac{1}{q_p}.$$

Therefore, conditional on the event that acceptance occurs in a round, the probability that the accepted team is $j$ equals

$$\frac{\frac{1}{|T_p|} \cdot r_j}{\frac{1}{|T_p|} \cdot \frac{1}{q_p}} = \frac{r_j}{1/q_p} = p_j.$$

Since the rounds are independent and identically distributed, and the process is guaranteed to stop eventually (because the geometric distribution ensures success in finite time), the final probability of selecting team $j$ equals the above conditional probability. Hence, the selection probability of team $j$ is exactly $p_j$. Considering that estimators $\hat{p}_j$ are unbiased, the results is also uniform if we use $\hat{p}_j$ to replace real $p_j$.

## 3.4 Managing N

We observe that in the aforementioned method, the choice of the number of simulations $N$ at each step is crucial. Selecting a very large $N$ allows us to obtain more accurate estimates $\hat{p}_j$ of $p_j$, but it also increases computational complexity, particularly the time required to generate the number of draws $G$ for each potential team to be placed in the next position. This contradicts the purpose of using estimated probabilities instead of exact values. Conversely, if $N$ is too small, some teams that could be placed in the next position may not appear in these $N$ simulated draws. Although this does not affect the uniformity of the final result, it resembles an artificial restriction, preventing certain teams from being included in the candidate pool for the next position, which could be perceived as unfair or manipulative by fans.

Next, we discuss the relationship between $N$ and the probability of such an occurrence. It is clear that for a team with a lower probability of being placed in the next position, the probability that it does not appear in any of the $N$ simulations is higher, and this probability depends on the true value of $p_j$, which we cannot directly compute. To simplify the model, suppose that in the next step of the draw, all eight candidate teams can be assigned to the next position with equal probability 1/8. Using the inclusion-exclusion principle, we can calculate the probability that after $N$ simulations, all eight teams appear at least once in the next position.

Define the events:

$$A_i : \text{the } i\text{-th outcome does not appear.}$$

We are interested in:

$$P(\text{each outcome appears at least once}) = 1 - P\left(\bigcup_{i=1}^{8} A_i\right).$$

By the principle of inclusion-exclusion:

$$P\Big(\bigcup_{i=1}^{8} A_i\Big) = \sum_{k=1}^{8} (-1)^{k+1} \left( \sum_{1 \le i_1 < \cdots < i_k \le 8} P(A_{i_1} \cap \cdots \cap A_{i_k}) \right).$$

$$P(A_i) = \left(\frac{7}{8}\right)^N,$$

since each draw selects one of the other 7 outcomes.

For $k$ outcomes not appearing simultaneously (e.g., $i_1, i_2, \ldots, i_k$):

$$P(A_{i_1} \cap \cdots \cap A_{i_k}) = \left(\frac{8-k}{8}\right)^N,$$

as only the remaining $8 - k$ outcomes can be drawn.

There are $\binom{8}{k}$ ways to choose these outcomes.

Hence:

$$P\Big(\bigcup_{i=1}^{8} A_i\Big) = \sum_{k=1}^{8} (-1)^{k+1} \binom{8}{k} \left(\frac{8-k}{8}\right)^N.$$

Finally, the probability that all 8 outcomes appear at least once is:

$$P(\text{all 8 outcomes appear at least once}) = 1 - \sum_{k=1}^{8} (-1)^{k+1} \binom{8}{k} \left(\frac{8-k}{8}\right)^N.$$

The corresponding relationship is shown in Figure 3.1. We see that when $N = 50$, this probability is already very close to 1 (specifically, 98.99%). It should be noted that since the actual selection probabilities of the teams differ, for the position with eight candidate teams (e.g., Pot 2 in Group A), the true probability of including all teams is slightly lower than the figure shows. Moreover, increasing $N$ beyond 50 does not substantially improve this probability, but the computational burden grows with $N$, approximately in a linear fashion. We conducted ten independent runs of the draw simulation for each value of N = 10, 20, 30, 40, 50, 60, 70, 80, 90, and 100, and recorded the total execution time for each set of ten simulations. As illustrated in the Figure 3.2, a strong linear relationship between the simulation time and the value of N is clearly discernible, notwithstanding the presence of an outlier (at N = 70) that can be attributed to the limited number of repetitions. Therefore, in our final simulations, we set $N = 50$.

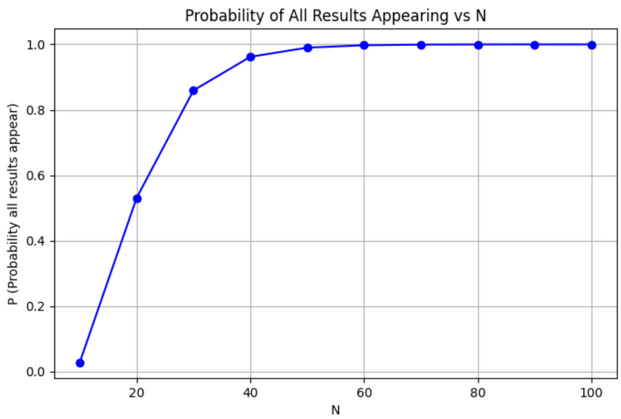


Figure 3.1: Probability that all candidate teams appear at least once in $N$ simulations

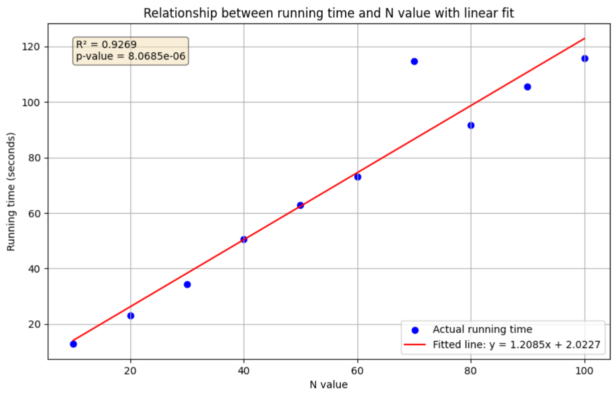


Figure 3.2: Running time for 10 simulations versus $N$

# Chapter 4

# Result Presentation And Comparison

Using the method described in Section 3, I implemented both the rejection sampling and the multi-rejection solution in Python, and conducted 1,000,000 and 10,000 draws respectively to validate the correctness of my code.

## 4.1 Rejection Sampling

First, we verify whether the results obtained via rejection sampling follow a uniform distribution. As discussed in Section 3, this verification is crucial because rejection sampling not only serves as a benchmark to check whether the results obtained from the multi-rejection solution are indeed uniform, but also forms the basis for the multi-rejection solution itself.

In Roberts' (2024) article, they performed the same 1,000,000 draws via rejection sampling and reported some conditional probabilities as follows:

- The probability that England is in the same group as Germany should be 10.53% (95% CI = [10.47%, 10.59%]) under $U$.
- The probability that Germany is in the same group as Qatar should be 13.74% (95% CI = [13.67%, 13.80%]) under $U$.
- The probability that Canada is in the same group as Qatar should be 15.53% (95% CI = [15.46%, 15.60%]) under $U$.

- The probability that the United States is in the same group as Qatar should be 9.06% (95% CI = [9.00%, 9.11%]) under $U$.

I also computed these probabilities from my 1,000,000 draws via rejection sampling, obtaining the following results:

Table 4.1: Comparison of probabilities from Roberts and our rejection sampling results

| Team Pair | Roberts' Probability (95% CI) | Our Result (%) | Within CI? |
|---|---|---|---|
| England–Germany | 10.53 (10.47–10.59) | 10.56 | Yes |
| Germany–Qatar | 13.74 (13.67–13.80) | 13.67 | Yes |
| Canada–Qatar | 15.53 (15.46–15.60) | 15.51 | Yes |
| United States–Qatar | 9.06 (9.00–9.11) | 9.08 | Yes |

It is clear that all four probabilities fall within the confidence intervals reported in the article, providing strong evidence that our rejection sampling correctly produces a uniform distribution.

## 4.2 The Multi-Rejection Solution

Since performing a single multi-rejection solution requires substantially more time than a simple rejection sampling, we conducted only 10,000 draws. We focused on a question of particular interest to fans: the probabilities that teams from Pot 2 are drawn into the same group (Group A) as Qatar, whose strength is considerably lower than that of the other Pot 1 teams. The probabilities for the eight Pot 2 teams being assigned to Group A under the two drawing methods are shown below.

It can be observed that the probabilities obtained from the two methods are very close. To verify whether the multi-rejection method indeed produces uniform results, we performed a goodness-of-fit test on the two sets of sampling results. It should be noted that, according to the symmetry concept introduced earlier, the five European teams in Pot 2 (Denmark, Netherlands, Switzerland, Germany, Croatia) are theoretically symmetric, and thus their probabilities of being drawn into Group A should be identical. Similarly, the two North American teams (Mexico and the United States) are also symmetric and should have the same probability of being drawn into Group A.

From the simulation results, the probabilities for these symmetric teams being drawn into Group A are very close, although slight differences remain.

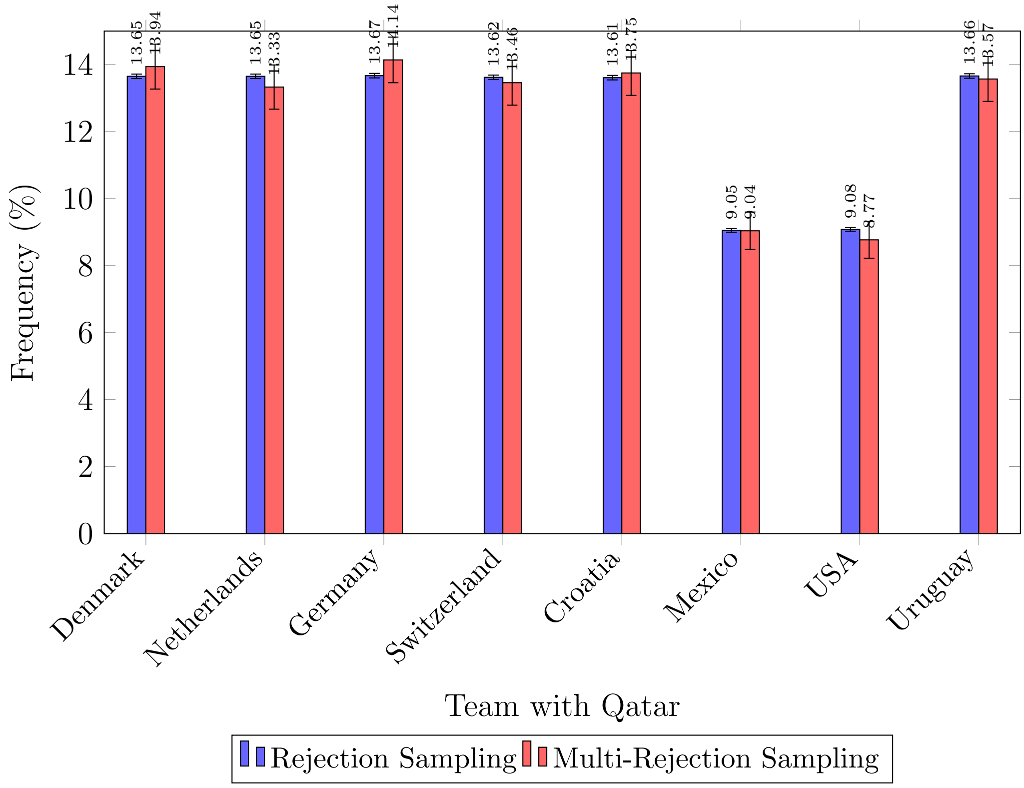


Figure 4.1: Comparison of the frequencies with which each team is drawn into the same group as Qatar using Rejection Sampling (n=1,000,000) and Multi-Rejection Sampling (n=10,000). Error bars represent 95% confidence intervals.

Therefore, for the goodness-of-fit test, we aggregated the probabilities for the five European teams, as well as the probabilities for the two North American teams. In practice, we performed the test for three categories: (i) European teams from Pot 2 assigned to Group A, (ii) North American teams from Pot 2 assigned to Group A, and (iii) South American teams from Pot 2 assigned to Group A. The resulting Chi-squared statistic is $\chi^2 = 0.8706$ with a p-value of 0.6471, indicating that the frequencies obtained from the multi-rejection draws differ very little from the theoretical expected frequencies (obtained via rejection sampling), and thus closely match the target distribution.

We also implemented a web-based version of our code using Streamlit. If you wish to try the multi-rejection method by yourself, the web application can be accessed at `https://world-cup-draw-f8lz7ru2mqzjtzmklnizba.streamlit.app/`. By clicking `Start`, the draw begins with the automatic assignment of the eight Pot 1 teams. Next, by clicking `Next`, the multi-rejection procedure is applied to the subsequent position to be filled. The process is displayed in another table with three columns: the first column lists all candidate teams that could potentially be drawn, the second column shows the number of draws required for each team to be assigned to the position (i.e., $G_j$ as described in Section 3), and the third column displays the cumulative number of times each candidate team has been drawn so far. In each step, the team drawn in that iteration is highlighted in red, and once a team reaches the required number of draws and is successfully assigned to the position, it is highlighted in green.

For convenience, if you think drawing one position at a time by clicking `Next` is too slow, you can click `Finish Draw` to rapidly complete the draw using the multi-rejection method and obtain the final result. The `Start` button can be used at any time to restart the entire process.

# Chapter 5

# Applications To Other Tournaments

We can observe that the Multiple-Rejection Method is fundamentally built upon rejection sampling. For any draw where uniform results can theoretically be obtained via rejection sampling, the Multiple-Rejection Method can also, in principle, produce the desired outcomes. As discussed in Section 2.1, rejection sampling is highly versatile: it can be applied to draws with a wide variety of constraints while still yielding uniform results. This implies that, by appropriately modifying the acceptance criteria in the rejection sampling step, the Multiple-Rejection Method can be adapted to different tournaments, with varying numbers of participating teams and distinct rules.

To test the generality of the Multi-Rejection Method, we applied it to previous World Cups. The 2018 World Cup, being the first to employ pots based on team strength, had similar constraint rules and the same number of participating teams, although some teams differed from 2022. According to László Csató (2025), there were significantly more valid group combinations in 2018 (approximately $2.05 \times 10^{15}$, compared to $5.9 \times 10^{14}$ in 2022).[2] While this would further increase the computational cost for the original Multiple-Rejection Method, the improved version described in this work benefits from being built on rejection sampling: a larger number of valid results corresponds to a higher acceptance rate, thereby increasing the efficiency of the draw.

This illustrates another advantage of the improved Multiple-Rejection Method compared to the original version. The time required for each draw step does not scale directly with the total number of valid outcomes or the total number of possible random permutations; rather, it depends on the ratio between these two quantities (i.e., the acceptance rate). The easier it is to satisfy the constraints, the higher the acceptance rate and the more efficient the draw becomes, even when more teams are added.

The pot allocations for 2018 were as follows:

- **Pot 1**: Russia [EU], Germany [EU], Brazil [SA], Portugal [EU], Argentina [SA], Belgium [EU], Poland [EU], France [EU].
- **Pot 2**: Spain [EU], Peru [SA], Switzerland [EU], England [EU], Colombia [SA], Mexico [NA], Uruguay [SA], Croatia [EU].
- **Pot 3**: Denmark [EU], Iceland [EU], Costa Rica [NA], Sweden [EU], Tunisia [AF], Egypt [AF], Senegal [AF], Iran [AS].
- **Pot 4**: Serbia [EU], Nigeria [AF], Australia [AS], Japan [AS], Morocco [AF], Panama [NA], Korea Republic [AS], Saudi Arabia [AS].

Apart from the host nation Russia, all other teams in the 2018 FIFA World Cup were allocated into four pots strictly according to their FIFA World Rankings at that time. It is worth noting that 2018 was the first edition in which FIFA adopted a ranking-based allocation system, rather than the previous geography-based approach, following Guyon's (2015) criticism of the earlier seeding rules. This modification not only helped balance the competitive pressure across groups but also rendered the draw process more standardized.

Nevertheless, the FIFA rankings did not perfectly reflect the actual strengths of the teams. For instance, Japan, which was placed in Pot 4, was arguably underestimated. Among the eight Pot 4 teams, Japan was the only one to advance to the Round of 16. Although they were eventually eliminated by Belgium (a Pot 1 team), Japan once held a surprising 2–0 lead, forcing the eventual third-place team to stage a dramatic comeback. This clearly illustrates that those drawn into the same group as Japan faced a significantly tougher challenge than expected when matched with a supposedly "weak" Pot 4 opponent.

In fact, in 2018, it was Poland from Pot 1 who had the misfortune of being grouped with Japan. Despite being the highest-ranked team in their group, Poland not only failed to qualify for the knockout stage but even finished bottom of their group. This outcome may indicate that Poland's strength was overestimated, but it also highlights the substantial role of randomness in shaping the difficulty of each team's path.

This naturally raises an interesting question: what was the probability that a strong team from Pot 1 or Pot 2 would be drawn into the same group as Japan? More importantly, can the multiple-rejection draw correctly capture such "unlucky" probabilities in a uniform and unbiased way? By applying rejection sampling with 1,000,000 draws and multi-rejection sampling with 10,000 draws, we simulated the 2018 FIFA World Cup group stage draw and compared the frequencies of higher-ranked teams being placed in the same group as Japan.

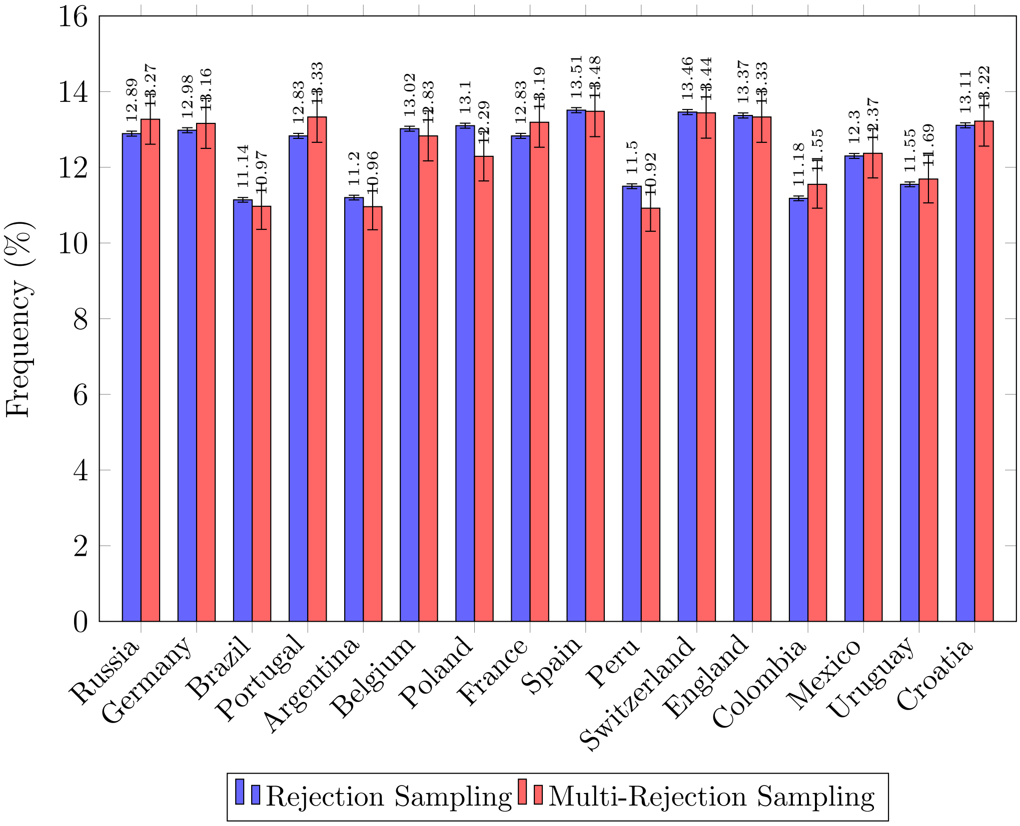


Figure 5.1: Comparison of the frequencies with which Japan is drawn in the same group as each team using Rejection Sampling (n=1,000,000) and Multi-Rejection Sampling (n=10,000). Error bars represent 95% confidence intervals.

From Figure 5.1, we can clearly observe that the probabilities of the top 15 ranked teams, together with the host nation Russia, being drawn into the same group as Japan are almost identical under both sampling methods. This demonstrates that the differences in probabilities among these sixteen teams are mainly caused by the structural constraints of the draw itself. Furthermore, if we take into account the symmetry among certain teams and average their probabilities of facing Japan in the group stage, the results obtained from the two methods become even closer. To provide a more intuitive demonstration that both drawing procedures yield equivalent and uniform results, we conducted a goodness-of-fit test on the data categorized by symmetry, as previously mentioned. The results show that $\chi^2 = 2.0470$ with a p-value of 0.8426. This strongly indicates that the outcomes from the multi-rejection method align well with theoretical expectations and are statistically indistinguishable from those obtained through rejection sampling. No evidence of any systematic bias was detected. This indicates that, even when the set of participating teams changes, the multiple-rejection sampling method not only remains applicable but also continues to yield results that follow a uniform distribution, fully satisfying the requirements of fairness.

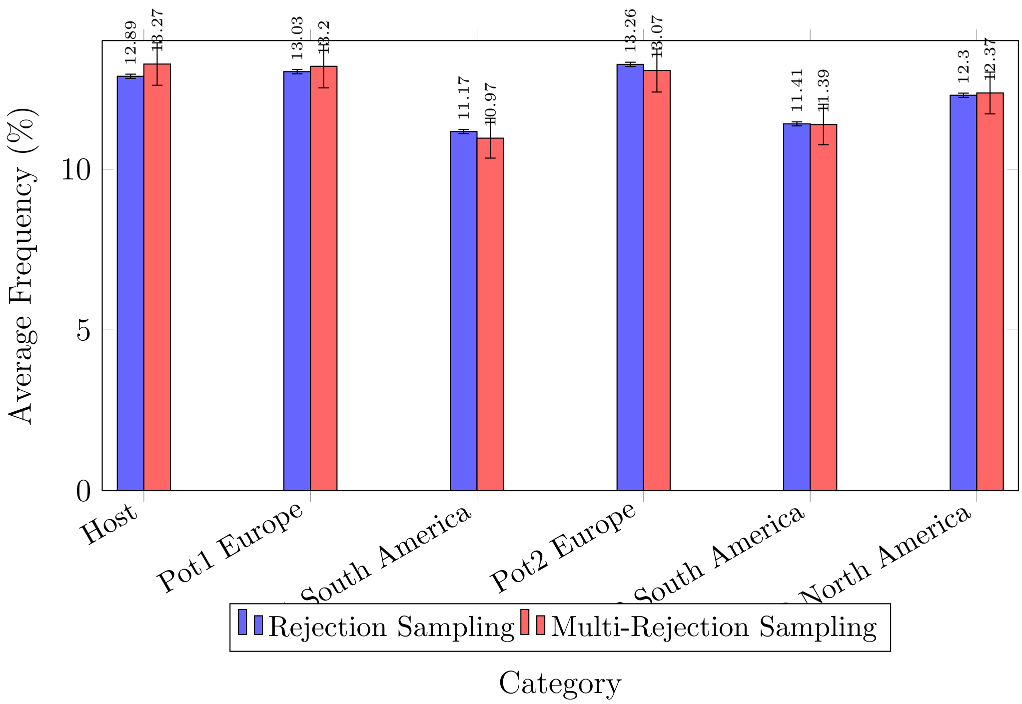


Figure 5.2: Average probabilities of being drawn into the same group as Japan, aggregated by category, comparing Rejection Sampling (n=1,000,000) and Multi-Rejection Sampling (n=10,000). Error bars represent 95% confidence intervals.

The 2026 FIFA World Cup will expand from the traditional 32 teams to 48 teams, and will include three host nations: the United States, Mexico, and Canada. These novel changes undoubtedly pose a significant challenge for the multiple-rejection method. However, as of now, the qualification stage is still ongoing, and only 13 teams, including the three hosts, have secured their places in the tournament. Moreover, this will be the first World Cup in history to be hosted by three countries simultaneously (the previous record being two hosts in 2002, when Japan and South Korea co-hosted). How the three host nations will be placed in the draw, and whether there will be specific rules concerning their allocation, remain unanswered questions, as the detailed regulations have not yet been released. Therefore, although we are highly interested in simulating the draw for the 2026 World Cup, the absence of crucial information makes it impossible to conduct such a simulation at this stage. Once the official rules are announced, however, our method can be directly applied to evaluate the fairness and uniformity of the draw under the new format.

# Chapter 6

# CONCLUSION

This paper aims to investigate and address the fairness and probabilistic biases inherent in the current FIFA World Cup group stage draw procedure. Through an analysis of the existing FIFA draw process, we identify that its sequential drawing mechanism contains intrinsic biases. Even after removing special restrictions, such as the host country privilege, unfairness persists as long as the geographical constraints are sufficiently complex. Therefore, a method that guarantees a uniform distribution of draw outcomes from first principles is necessary.

In this context, we successfully implement the *Multiple-Rejection Solution* proposed by Roberts (2024), which previously existed only as a theoretical construct, and transform it into a practical procedure. We introduce a key modification to the original method: based on a given partial draw, we estimate the probability of each team being assigned to the next position using a small number of rejection sampling iterations. This avoids the enormous computational cost of enumerating all valid completions. This innovative improvement greatly enhances the feasibility of the method for practical use, making it fully compatible with the transparency and efficiency requirements of live draws. Our implemented draw procedure can complete the entire draw within an acceptable time frame (e.g., a few tens of seconds), which is of significant practical importance.

To verify the correctness of our implementation, we conducted 10,000 multi-rejection draw simulations and compared the results with standard uniform distribution data generated via rejection sampling. Statistical analysis shows that the probability distributions produced by the two methods are indistinguishable. This not only confirms the theoretical correctness of the multiple-rejection solution but also further validates the accuracy of our implemented procedure.

Moreover, to ensure transparency and avoid public skepticism, we examined the number of valid uniform results required at each step of the draw (the parameter $N$). We ultimately set $N = 50$, a choice that balances computational efficiency with the likelihood that all potential teams are included in the candidate pool, thereby preserving the perceived fairness of the draw.

In summary, this work not only successfully translates an advanced theoretical draw method into a practical application but also demonstrates, through rigorous simulation and statistical validation, that this draw procedure represents a fair, efficient, and transparent alternative to the current FIFA draw mechanism.

# Bibliography


[1] László Csató. Group draw with unknown qualified teams: A lesson from the 2022 fifa world cup draw. *International Journal of Sports Science & Coaching*, 18(2):539–551, 2023.

[2] László Csató. The fairness of the group draw for the fifa world cup. *International Journal of Sports Science & Coaching*, 20(2):554–567, 2025.

[3] Luc Devroye. Sample-based non-uniform random variate generation. In *Proceedings of the 18th conference on Winter simulation*, pages 260–265, 1986.

[4] FIFA. Fifa world cup qatar 2022 draw procedures. `https://digitalhub.fifa.com/m/2ef762dcf5f577c6/original/Portrait-Master-Template.pdf`, 2022.

[5] Julien Guyon. Rethinking the fifa world cup™ final draw. *Journal of Quantitative Analysis in Sports*, 11(3):169–182, 2015.

[6] Stefan Klößner and Martin Becker. Odd odds: the uefa champions league round of 16 draw. *Journal of Quantitative Analysis in Sports*, 9(3):249–270, 2013.

[7] Christian P Robert, George Casella, and George Casella. *Monte Carlo statistical methods*, volume 2. Springer, 1999.

[8] Gareth O Roberts and Jeffrey S Rosenthal. Football group draw probabilities and corrections. *Canadian Journal of Statistics*, 52(3):659–677, 2024.

[9] UEFA. Uefa champions league round of 16 draw. `https://www.uefa.com/uefachampionsleague/news/0283-1879486c017e-68061468b1a7-1000--uefa-champions-league-round-of-16-draw/`, 2023.