

# Simulación de Juegos de Polo Acuático

Leonardo Artiles Montero

Edián Broche Castro

June 10, 2024

## Abstract

En este trabajo, se presenta una simulación detallada de un torneo de polo acuático, basada en datos reales de enfrentamientos entre equipos de diversos países en competencias recientes. El principal objetivo es predecir al ganador de las próximas olimpiadas a realizarse en París 2024. Se exploran tres enfoques distintos para la simulación usando la distribución de Poisson[1]:

1. Distribución de Poisson General: Utilizando la cantidad total de goles anotados por cada equipo, se aplica una distribución de Poisson para predecir los resultados de los partidos.
2. Distribución Conjunta de Poisson y KDE: se evalúa la distribución conjunta de goles anotados en enfrentamientos previos entre dos equipos específicos, utilizando tanto Poisson como Kernel Density Estimation (KDE).
3. Distribución de Poisson No Homogénea con B-Splines: Este enfoque, el más detallado, considera la variabilidad en la intensidad de los goles a lo largo del tiempo de juego. Utilizamos B-splines para modelar esta no homogeneidad y ajustar una distribución de Poisson no homogénea.

En los experimentos realizados, se descubre que el segundo enfoque produce resultados más cercanos a la realidad, mientras que el tercero ofrece un control adicional al capturar la variabilidad temporal en la intensidad del juego. Además, se desarrolló una plataforma web interactiva que permite realizar simulaciones y visualizar resultados. Los hallazgos realizados proporcionan una herramienta útil para el análisis y predicción en el ámbito del polo acuático, destacando la importancia de considerar la no homogeneidad en la intensidad del juego.

## 1. Introducción

Este trabajo de simulación se basa en la predicción del resultado de un torneo de polo acuático. Utilizando datos reales de enfrentamientos entre equipos de diversos países en competencias recientes, el objetivo principal es predecir el ganador de las próximas Olimpiadas en París 2024.

El polo acuático, también conocido como waterpolo, es un deporte acuático en el que dos equipos compiten por marcar goles en la portería del equipo contrario. A diferencia de otros deportes, los jugadores deben realizar esta tarea mientras nadan, lo que añade un desafío adicional.

### **Reglas Básicas del Waterpolo**

1. **Objetivo:** El objetivo principal del waterpolo es convertir la mayor cantidad de goles en la portería rival durante el partido. Los puntos acumulados determinan al vencedor.
2. **Duración del Partido:** Cada partido se divide en cuatro tiempos de juego efectivo, con ocho minutos de duración cada uno. Entre cada tiempo, hay descansos de dos minutos.
3. **Equipos:** Cada equipo está compuesto por siete jugadores en el campo. Uno de ellos es el portero, identificado por un gorro rojo. Además, hasta seis jugadores suplentes pueden entrar en la piscina en cualquier momento durante las competiciones oficiales.

Además, se busca responder a algunas preguntas específicas sobre las dinámicas de los partidos, tales como:

- ¿Cómo influye la intensidad del juego en diferentes momentos del partido?
- ¿Qué impacto tienen los enfrentamientos históricos entre equipos en los resultados futuros?
- ¿Es posible identificar patrones en la anotación de goles a lo largo del tiempo?

Para abordar estos objetivos, se han identificado y utilizado varias variables de interés, incluyendo el número de goles anotados por cada equipo, la distribución temporal de los goles, y los resultados de enfrentamientos históricos entre equipos específicos. El sistema desarrollado para esta simulación se basa en tres enfoques principales:

1. **Distribución de Goles por Equipo:** Utilizando una distribución de Poisson para modelar la cantidad de goles anotados por cada equipo en general.
2. **Distribución Conjunta de Goles:** Utilizando una distribución conjunta de Poisson para modelar los goles anotados en enfrentamientos específicos entre equipos, así como la estimación de densidad de kernel (KDE) como una alternativa.
3. **Distribución de Poisson No Homogénea:** Utilizando una distribución de Poisson no homogénea con B-splines para modelar la variabilidad de la intensidad del juego a lo largo del tiempo.

Estos enfoques permiten analizar y simular diferentes escenarios y estrategias en los partidos de polo acuático, proporcionando una herramienta robusta para la predicción y análisis de resultados. Adicionalmente, se ha desarrollado una plataforma web interactiva

que permite a los usuarios realizar simulaciones y experimentar con los modelos propuestos, ofreciendo una comprensión más profunda de las dinámicas del juego.

El resto del documento está estructurado de la siguiente manera: en la Sección 2 se describen los detalles de implementación; en la Sección 3 se presentan los resultados y experimentos; y en la Sección 4 se detalla el modelo matemático utilizado.

## **2. Detalles de Implementación**

### **2.1. Recolección de Datos**

Para llevar a cabo esta investigación, se recolectaron datos de las últimas seis Copas Anuales de Polo Acuático mediante técnicas de *web scraping* desde la página web World Aquatics. Estos datos incluyen información sobre los enfrentamientos de equipos de diferentes países, el número de goles anotados y recibidos, y otros detalles relevantes de los partidos.

### **2.2. Metodología de Simulación**

Con el objetivo de predecir los resultados del torneo, se emplearon tres enfoques distintos:

#### **2.2.1. Enfoque 1: Distribución de Poisson General**

El primer enfoque se basa en la distribución de Poisson para cada equipo de manera individual. Se tomó la cantidad total de goles anotados por cada equipo en todas las competencias y se utilizó una distribución de Poisson para modelar la probabilidad de anotar goles en un partido. Este enfoque permitió responder a la pregunta: ¿Es posible tener una simulación precisa utilizando solo la distribución general de goles por equipo?

#### **2.2.2. Enfoque 2: Distribución Conjunta de Poisson y KDE**

El segundo enfoque mejora la precisión al considerar la distribución conjunta de goles anotados en enfrentamientos anteriores entre dos equipos específicos. Para ello, se utilizaron tanto la distribución conjunta de Poisson como la estimación de densidad de kernel (KDE). Este método permitió modelar de manera más precisa la dinámica específica entre los equipos basándose en sus enfrentamientos previos.

#### **2.2.3. Enfoque 3: Distribución de Poisson No Homogénea con B-Splines**

En el tercer enfoque, se enfocó en la variabilidad de la intensidad del juego a lo largo del tiempo. A diferencia de la distribución de Poisson homogénea, que asume una tasa constante de goles, se investigó si la intensidad varía durante diferentes momentos del juego. Para ello, se utilizaron B-splines para modelar esta no homogeneidad en la distribución de Poisson. La implementación fue la siguiente:

1. Se dividió el tiempo del partido en cuatro períodos de 8 minutos.
2. Se discretizó cada período en segundos y se calculó la probabilidad de anotar un gol en cada segundo utilizando una distribución de Poisson no homogénea ajustada con B-splines.
3. Se sumaron los resultados de los cuatro períodos para obtener el resultado final del partido.
4. Se simuló el torneo completo aplicando las reglas del polo acuático.
5. Se realizaron múltiples simulaciones hasta alcanzar un nivel de confianza adecuado en los resultados.

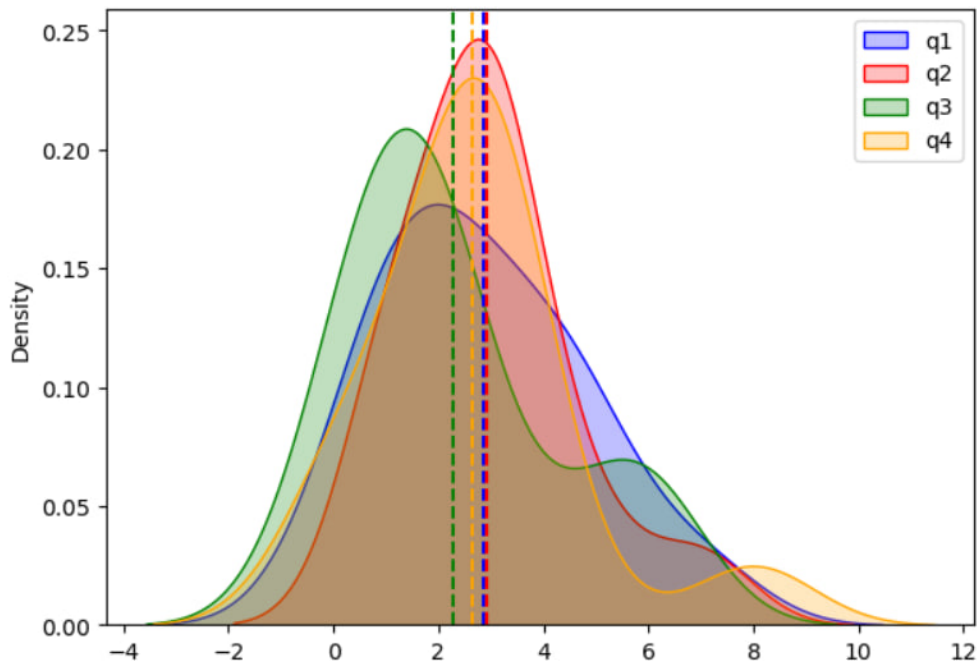


Figure 2.1: Ejemplo de la distribución de la probabilidad de gol en cada cuarto de un partido

### 2.3. Plataforma Web Interactiva

Se desarrolló una plataforma web interactiva que permite a los usuarios realizar sus propias simulaciones y experimentar con los diferentes enfoques. La plataforma visualiza los resultados y permite ajustar los parámetros de la simulación, proporcionando una herramienta educativa y analítica para entender mejor las dinámicas del polo acuático.

## 2.4. Evaluación de Resultados

Cada enfoque fue evaluado en términos de precisión y realismo comparando los resultados de las simulaciones con los resultados históricos de torneos recientes. En el primer enfoque, se observó que equipos con pocos partidos pero con buenos resultados tendían a sobrevalorarse. El segundo enfoque mostró una mayor precisión, alineándose mejor con las posiciones históricas de los equipos más fuertes. El tercer enfoque proporcionó un control adicional sobre la variabilidad temporal, permitiendo ajustes más finos en la simulación.

En resumen, los experimentos demostraron que considerar la no homogeneidad en la intensidad del juego es crucial para obtener simulaciones más precisas y realistas en el contexto del polo acuático.

## 3. Resultados y Experimentos

### 3.1. Enfoque 1: Distribución de Poisson General

El primer enfoque, basado en la distribución de Poisson general para cada equipo, no resultó óptimo para obtener resultados precisos. En particular, cuando la cantidad de datos disponibles era limitada, los equipos que habían jugado pocos partidos pero obtenido buenos resultados tendían a sobrevalorarse. Por ejemplo, equipos como Australia y Japón aparecían consistentemente como los más fuertes, lo cual no reflejaba la realidad competitiva.

### 3.2. Enfoque 2: Distribución Conjunta de Poisson y KDE

El segundo enfoque, que utiliza la distribución conjunta de Poisson y la estimación de densidad de kernel (KDE) para modelar los enfrentamientos entre equipos específicos, demostró ser significativamente más preciso. Los resultados de las simulaciones eran consistentes con los resultados históricos, con equipos como Croacia, España, Italia y Grecia posicionándose en los primeros lugares, reflejando su desempeño en torneos recientes.

### 3.3. Enfoque 3: Distribución de Poisson No Homogénea con B-Splines

El tercer enfoque, que incorpora la distribución de Poisson no homogénea ajustada con B-splines, proporcionó un control adicional sobre la variabilidad temporal en la intensidad del juego. Este enfoque permitió modelar de manera más precisa la dinámica de los partidos, reflejando variaciones en la intensidad de juego a lo largo de los cuatro períodos de 8 minutos de cada partido.

En los experimentos realizados con este enfoque, se observó que la intensidad del juego variaba significativamente, con más goles anotados en los primeros y segundos períodos. Esto sugiere que los primeros dos períodos pueden ser determinantes en el resultado final del partido.

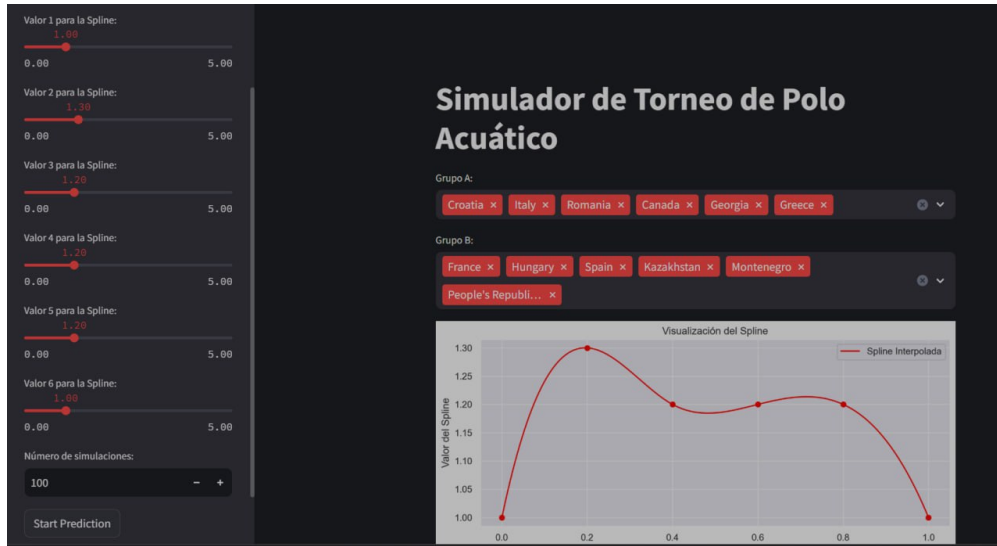


Figure 3.1: Interfaz de la plataforma web interactiva desarrollada para realizar simulaciones.

### 3.4. Manejo de Datos Insuficientes

Un experimento adicional interesante fue la simulación de partidos entre equipos que no tenían datos de enfrentamientos previos. Inicialmente, se utilizó una selección aleatoria entre los equipos, lo cual resultó ineficaz y produjo resultados poco realistas. Posteriormente, se consideraron métricas como la cantidad total de goles anotados para inferir la fortaleza relativa de los equipos. Sin embargo, estas ideas más complejas, como el uso de la descomposición de valores singulares (SVD) para el completamiento de la matriz de probabilidades, o el completamiento del DAG (grafo de dominación), no se incluyeron en el análisis final por su complejidad y falta de relación directa con el enfoque principal del trabajo.

### 3.5. Evaluación Comparativa de los Enfoques

Enfoque	Precisión	Realismo
Poisson General	Baja	Baja
Poisson Conjunta y KDE	Alta	Media
Poisson No Homogénea	Alta	Alta

Table 3.1: Comparación de precisión y realismo entre los distintos enfoques de simulación.

Como se muestra en la Tabla 3.1, el enfoque de Poisson general resultó ser el menos preciso y realista. El enfoque de distribución conjunta de Poisson y KDE mostró una alta precisión y realismo, alineándose con los resultados históricos. El enfoque de Poisson no homogénea proporcionó el mayor nivel de precisión y realismo, permitiendo un modelado detallado de la intensidad variable del juego.

En resumen, los experimentos demostraron que considerar la no homogeneidad en la intensidad del juego es crucial para obtener simulaciones más precisas y realistas en el contexto del polo acuático. La plataforma web desarrollada proporciona una herramienta útil para explorar estos modelos y sus implicaciones en escenarios competitivos reales.

## 4. Modelo Matemático

### 4.1. Distribución de Poisson

El primer enfoque de simulación utiliza la distribución de Poisson para modelar el número de goles anotados por cada equipo en un partido. La distribución de Poisson es adecuada para este propósito ya que modela eventos discretos que ocurren de manera independiente a una tasa constante. La función de probabilidad de la distribución de Poisson está dada por:

$$P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!} \quad (4.1)$$

donde  $P(X = k)$  es la probabilidad de que ocurran  $k$  goles, y  $\lambda$  es la tasa promedio de goles anotados por partido.

### 4.2. Distribución Conjunta de Poisson

Para mejorar la precisión, el segundo enfoque considera la distribución conjunta de goles anotados en enfrentamientos específicos entre dos equipos. La distribución conjunta de Poisson se puede modelar utilizando una matriz de intensidades  $\Lambda$ , donde cada entrada  $\lambda_{ij}$  representa la tasa de goles del equipo  $i$  contra el equipo  $j$ . La función de probabilidad conjunta está dada por:

$$P(X = k, Y = m) = \frac{\lambda_{ij}^k e^{-\lambda_{ij}}}{k!} \cdot \frac{\lambda_{ji}^m e^{-\lambda_{ji}}}{m!} \quad (4.2)$$

donde  $\lambda_{ij}$  y  $\lambda_{ji}$  son las tasas de goles de los equipos  $i$  y  $j$ , respectivamente.

### 4.3. Distribución de Poisson No Homogénea con B-Splines

El tercer enfoque introduce la distribución de Poisson no homogénea, que permite la variabilidad en la intensidad de goles a lo largo del tiempo del partido. Esta variabilidad se modela utilizando B-splines, que proporcionan una forma flexible de aproximar funciones complejas. La intensidad de la tasa de goles  $\lambda(t)$  en un momento  $t$  se define como:

$$\lambda(t) = \sum_{i=1}^n \beta_i B_i(t) \quad (4.3)$$

donde  $B_i(t)$  son las funciones base de B-splines y  $\beta_i$  son los coeficientes correspondientes. La función de probabilidad para la distribución de Poisson no homogénea está dada por:

$$P(N(t) = k) = \frac{\left(\int_0^t \lambda(u) du\right)^k e^{-\int_0^t \lambda(u) du}}{k!} \quad (4.4)$$

donde  $\int_0^t \lambda(u) du$  es la integral de la tasa de goles a lo largo del tiempo  $t$ .

#### 4.4. Criterio de Parada de la Simulación

Para determinar cuándo detener las simulaciones, se utilizó un criterio basado en la convergencia de los resultados. Después de cada simulación, se calcula la distribución de las posiciones finales de los equipos. Si la variabilidad de las posiciones no cambia significativamente después de un número predefinido de simulaciones, se considera que los resultados han convergido. La medida de variabilidad se define como la desviación estándar de las posiciones de los equipos en las simulaciones.

El criterio de parada se puede formalizar como:

$$\text{Si } \frac{1}{N} \sum_{i=1}^N \sigma_i < \epsilon \quad \text{entonces detener simulaciones} \quad (4.5)$$

donde  $\sigma_i$  es la desviación estándar de las posiciones de los equipos en la  $i$ -ésima simulación,  $N$  es el número total de simulaciones realizadas, y  $\epsilon$  es un umbral pequeño predefinido.

En resumen, el modelo matemático desarrollado en este trabajo combina técnicas de simulación estocástica con modelado avanzado de tasas de eventos, proporcionando una herramienta robusta para la predicción y análisis de resultados en torneos de polo acuático.

## 5. Referencias

### References

- [1] R. Greenhough, P. Birch, S. Chapman, and G. Rowlands, *Soccer: Is scoring goals a predictable Poissonian process?*, EPL (Europhysics Letters), vol. 89, no. 3, Mar. 2010. Available: <https://iopscience.iop.org/article/10.1209/0295-5075/89/38007/meta>
- [2] A. N. Kolmogorov and V. I. Arnold, *Kolmogorov-Arnold Networks*, Journal of Nonlinear Dynamics, vol. 5, no. 2, 1998.

## Acknowledgments

Thanks to molcito por sentarse a depurar el cosito de Latex.