



DOBLE GRADO EN  
MATEMÁTICAS Y ESTADÍSTICA

---

TRABAJO FIN DE GRADO

---

*Comparaciones Pareadas  
mediante modelos*

*Bradley Terry y su aplicación  
a datos en el fútbol*

---

Paola Cabrera Perza

Tutora: Ana M. Muñoz Reyes

Sevilla, Junio de 2022

*A Ana, estoy muy orgullosa de ti,  
sigue trabajando y luchando, todo llega. Te quiero.*

# Índice general

Agradecimientos . . . . .	III
Resumen . . . . .	V
Abstract . . . . .	VI
<b>1. Introducción</b>	<b>1</b>
1.1. Antecedentes en las comparaciones pareadas hasta Bradley-Terry . . . . .	1
<b>2. Modelo Bradley-Terry</b>	<b>5</b>
2.1. Historia del Modelo Bradley-Terry . . . . .	5
2.2. Aplicaciones del Modelo . . . . .	6
2.3. Modelo . . . . .	7
2.4. Axioma de Luce . . . . .	7
2.5. Estimación de los parámetros de máxima verosimilitud . . . . .	8
2.6. Modelo de Bradley-Terry como un modelo lineal generalizado . . . . .	9
<b>3. Modelo Bradley-Terry con empates</b>	<b>13</b>
3.1. Modelo . . . . .	13
3.2. Estimación de los parámetros de máxima verosimilitud . . . . .	14
3.3. Modelo Bradley-Terry con empates como un modelo lineal generalizado . . . . .	15
<b>4. Modelo Bradley-Terry con ventaja común de jugar de local</b>	<b>21</b>
4.1. Modelo . . . . .	21
4.2. Estimación de los parámetros de máxima verosimilitud para ventaja de local común . . . . .	22
4.3. Modelo Bradley-Terry con ventaja de jugar de local como un modelo lineal generalizado . . . . .	25
<b>5. Modelo Bradley-Terry con ventaja individual de jugar de local</b>	<b>29</b>
5.1. Modelo . . . . .	29
5.2. Estimación de los parámetros de máxima verosimilitud para ventaja individual de jugar de local . . . . .	29
5.3. Modelo Bradley-Terry con ventaja individual de jugar de local como un modelo lineal generalizado . . . . .	33
<b>6. Aplicación del Modelo Bradley Terry a datos de fútbol</b>	<b>37</b>
6.1. Clasificación mediante el modelo de Davidson . . . . .	37
6.2. Análisis de la ventaja de jugar de local común e individual para los distintos equipos de La Liga . . . . .	42
<b>A. Análisis de La Liga usando el modelo de Davidson en R</b>	<b>47</b>

A.1. Generar la matriz del modelo de Davidson . . . . .	47
A.2. Analisis para el final de temporada . . . . .	48
A.3. Análisis para los diferentes periodos de la temporada . . . . .	49
<b>B. Análisis de la ventaja de jugar de local común e individual para los distintos equipos de La Liga</b> . . . . .	<b>55</b>
B.1. Generar la matriz del modelo con ventaja común . . . . .	55
B.2. Generar la matriz del modelo con ventaja individual . . . . .	56
B.3. Creamos la tabla final con las ventajas individuales y comunes de todos los equipos a lo largos de las últimas 8 temporadas . . . . .	57
B.4. Creamos los gráficos usados . . . . .	61
B.5. Generamos la tabla para comparar los ajustes de los modelos . . . . .	62
<b>Bibliografía</b>	<b>66</b>

# Agradecimientos

Quiero empezar dando las gracias a los pilares fundamentales de mi vida, mis padres.

A mi madre por hacerme la vida tan fácil, cuando fallan las fuerzas siempre está su mano para levantarme, me da lecciones de fortaleza constantemente, y siempre serás mi ejemplo a seguir.

A mi padre por enseñarme que todo en la vida puede conseguirse si se trabaja para ello, por admirar mi camino tanto como yo admiro el tuyo, gracias siempre.

A mi hermano por ser mi apoyo en todo momento, mi cómplice durante toda la vida, el hombro en el que llorar y el compañero perfecto para reír, siempre serás mi debilidad.

A Ana, sin tu ayuda y apoyo no lo habría conseguido jamás, gran parte de esto es tuyo, gracias por tu fe ciega en mí, por darme siempre ese empujón que me ayuda a seguir, por las noches en vela estudiando, por los malos momentos superados y por todo lo que nos queda por superar juntas, por estar siempre con una sonrisa haciéndome la vida más bonita y por ser hogar.

A Gloria, por ser mi confidente cuando lo necesito, que siempre tengas esa libertad y esas ganas de volar que yo estaré para que ayudarte a ello.

A Alicia y Esther, gracias por confiar siempre en mí, más que yo misma, por darme siempre otro punto de vista que ha hecho que amplie mis fronteras y crezca como persona.

A Cristina, mi profesora de Matemáticas durante secundaria, gracias por enseñarme las Matemáticas de esa forma, me hicieron admirarlas mientras te admiraba a ti, sin duda gracias a esa forma de enseñar he llegado hasta aquí.

A mis abuelos José y Dolores, abuelo espero seguir siendo tu niña muchos años más, que me sigas recibiendo con esa sonrisa y que nos siga uniendo el amor que sentimos por la abuela, abuela espero que estés orgullosa de mi desde el cielo, no sabes como me duele tu ausencia, eras mi talón y lo sigues siendo, gracias por enseñarme a amar.

*Paola Cabrera Perza*



# Resumen

Este trabajo tiene como objetivo la clasificación mediante comparaciones pareadas a través del modelo de Bradley-Terry de individuos de un conjunto a comparar. Comienza con los antecedentes históricos de las comparaciones pareadas hasta llegar al modelo de Bradley-Terry (capítulo 1), después se desarrolla la historia del modelo, su aplicación a diferentes ámbitos de la vida, su definición, la estimación de los parámetros que aparecen en el modelo por máxima verosimilitud y la representación del modelo como un modelo lineal generalizado (capítulo 2). En el siguiente capítulo ampliaremos el modelo inicial incluyendo los empates ya que en el modelo inicial no se contemplaban, estimaremos los parámetros por máxima verosimilitud y lo representaremos como un modelo lineal generalizado (capítulo 3). En los capítulos 4 y 5 ampliaremos el modelo anterior y además de empates, tendremos en cuenta la ventaja común e individual de jugar de local, estimaremos también los parámetros por máxima verosimilitud y los representaremos como modelos lineales generalizados. En el último capítulo aplicaremos nuestro modelo a datos reales de fútbol en España en las últimas 8 temporadas, analizaremos primero si el ranking final ofrecido por el modelo de Davidson con los datos de La Liga 2020/2021 se parece al modelo de clasificación usado por La Liga, y analizaremos la ventaja de jugar de local tanto común como individual a lo largo de las últimas 8 temporadas.

# Abstract

The objective of this work is the classification by means of paired comparisons through the Bradley-Terry model of individuals of a set to be compared. It starts with the historical background of paired comparisons until the Bradley-Terry model (chapter 1), then the history of the model, its application to different areas of life, its definition, the estimation of the parameters appearing in the model by maximum likelihood and the representation of the model as a generalized linear model (chapter 2). In the next chapter we will extend the initial model to include ties, which were not considered in the initial model, estimate the parameters by maximum likelihood and represent it as a generalized linear model (chapter 3). In chapters 4 and 5 we will extend the previous model and in addition to ties, we will take into account the common and individual home field advantage, we will also estimate the parameters by maximum likelihood and we will represent them as generalized linear models. In the last chapter we will apply our model to real Spanish soccer data over the last 8 seasons, we will first analyze if the final ranking provided by Davidson's model with the 2020/2021 La Liga data resembles the ranking model used by La Liga, and we will analyze the common and individual home advantage over the last 8 seasons.



# Capítulo 1

## Introducción

### 1.1. Antecedentes en las comparaciones pareadas hasta Bradley-Terry

El análisis de experimentos que involucran comparaciones pareadas ha recibido bastante atención en la metodología estadística a lo largo de la historia. Podemos remontarnos a 1927 cuando Thurstone trató la Ley del Juicio Comparativo como una aplicación especial para obtener escalas de medida para variables no físicas, sino psicológicas, como la inteligencia, actitudes, sensaciones, etc. El objetivo principal de Thurstone era medir estímulos psicológicos y compararlos entre sí a través de comparaciones dos a dos, comparaciones pareadas. Su trabajo se centró en clasificar preferencias y actitudes psicológicas. Sugirió que cada estímulo genera una respuesta subjetiva en los sujetos cuestionados, y en consecuencia ellos asignarán un valor subjetivo al estímulo, ya que son estímulos cuyos valores no pueden ser obtenidos mediante mediciones físicas. Este valor podría también variar dependiendo si se cuestiona en otro instante al mismo sujeto. Como consecuencia de estas variaciones, en lugar de un solo proceso discriminatorio en relación al estímulo, se lleva a cabo una distribución de procesos discriminatorios. Thurstone supuso en su modelo que esa distribución es normal. Para ordenar los estímulos en una escala, Thurstone comparó dos a dos los estímulos a través de un conjunto de sujetos y propuso un modelo matemático para relacionar los valores numéricos que fueron atribuidos a los estímulos con las proporciones observadas. Si hay muchos estímulos a comparar, la aplicación del modelo llevaría mucho tiempo por lo que no es un método muy usado (Thurstone, 1927). Muchos años después Mosteller (1951) profundizó en el método de Thurstone y desarrolló un test  $\chi^2$  para este modelo (Bradley and Terry, 1952).

Más tarde Zermelo en 1929 investigó sobre la fortaleza de jugadores de ajedrez en torneos, y esto supuso una investigación importante en la historia de las comparaciones pareadas.

En los torneos con los que trabajó Zermelo cada participante jugaba un cierto número de veces, normalmente dos con orden alterno, con cada uno de los otros participantes, y más tarde se ordenaban los ganadores según el número de partidas ganadas. Los empates se contaban como media victoria para cada uno de los dos participantes. Este método no tenía en cuenta el orden en el que se producían las partidas, por lo que el azar apenas influía. Sin embargo, esto no funcionaba en los torneos en los que el número de partidas jugadas no era el mismo para todos los participantes (Zermelo, 1929). Para solucionar este

problema, Zermelo formuló un modelo probabilístico para predecir los resultados en juego en función de parámetros desconocidos que miden la fortaleza de los distintos jugadores. Este modelo asume que en las partidas de ajedrez no existen los empates, solo se gana o se pierde. Dados dos competidores  $i$  y  $j$ , el modelo viene dado por

$$P(i \text{ gane a } j) = \frac{\pi_i}{\pi_i + \pi_j}$$

donde  $\pi_i$  y  $\pi_j$  son valores desconocidos de la fortaleza de los jugadores  $i$  y  $j$  que se tienen que estimar. Bradley-Terry basan su modelo en éste (Glickman, 2011).

En 1940 Kendall y Smith publicaron “On the Method of Paired Comparisons”, en él buscan solucionar el problema que trae consigo el método de comparaciones pareadas mediante clasificación. Se dieron cuenta que si teníamos una serie de objetos A,B,C y queremos estudiarlos según los diferentes grados o niveles que presentan de una cualidad común a los tres, si la cualidad no es medible de alguna manera objetiva, tendríamos que recurrir a una manera subjetiva preguntando a diferentes observadores. Esto trae consigo el problema de que al tratarse por ejemplo de gustos ya sean musicales, gastronómicos o de otra índole el observador puede preferir el objeto A al B, el B al C, y finalmente el C al A, pero este caso sería imposible en los métodos de clasificación ya que si el observador prefiere el A al B, y el B al C, automáticamente aparece que prefiere entonces el A al C. Kendall y Smith proponen entonces un método más general para investigar este tipo de preferencias. Su método para analizar las comparaciones pareadas se centra en analizar las combinaciones que generan las preferencias de los observadores. La coherencia de los observadores fue medida en referencia a lo que denominaron “triadas circulares”, las triadas circulares eran las preferencias que podían representarse como un círculo cerrado, del tipo A->B, B->C,C->A (Kendall and Smith, 1940).

Hasta ahora se había trabajado para comparar elementos individuales, Guttman da un paso más y en 1946 introduce por primera vez el hecho de que también podrían compararse combinaciones de estos elementos. Lo expone a través de un caso práctico en el cuál tenían que determinar el número de puntos que habría que asignar a cada variable para la tarjeta de puntos para la desmovilización del Ejército de Estados Unidos a través de las opiniones de los propios soldados. El estudio al respecto se basa en el uso de comparaciones pareadas de una forma más compleja que las que se conocían hasta ahora. En un encuesta se descubrió que los soldados creían que las variables a considerar en la tarjeta para determinar el orden de desmovilización eran cinco: duración del tiempo en el Ejército, duración del tiempo en el extranjero, cantidad de combates, edad y número de hijos. Ahora tenían que determinar qué peso le daban a cada variable para obtener las puntuaciones totales. Según las comparaciones por pares que se habían visto hasta ahora, se tendría que responder por ejemplo a la pregunta “¿Quién debería salir primero al terminar la guerra: el soldado que tiene dos hijos o el que ha estado en dos batallas?”, pero los soldados se negaban a responder este tipo de preguntas porque no sabían la experiencia que tenía en la guerra el primer soldado, ni la descendencia del segundo. Por lo tanto en la investigación real usaron combinaciones de estas variables del tipo : ¿Quién cree que debería salir primero: Un hombre soltero que ha combatido durante dos años, un hombre casado sin hijos que lleva un año combatiendo o un hombre con hijos que no ha combatido? Compararon cada variable con todas las demás de esta forma (Guttman, 1946).

Ralph Allan Bradley y Milton E.Terry en 1952 en su estudio sobre comparaciones

pareadas dicen que cuando en un experimento de clasificación sólo se comparan dos objetos, el test de hipótesis de igualdad entre ellos puede basarse en la distribución binomial. La estimación de las probabilidades de que los elementos sean considerados superiores en una determinada comparación pueden llevarse a cabo y estas estimaciones nos ofrecen un método de clasificación de los elementos. En su estudio obtuvieron una generalización del modelo y de la distribución binomial. Formularon un modelo matemático y las estimaciones de máxima verosimilitud de calificaciones de los objetos a comparar proporcionan una solución sencilla al problema de cuantificación que tenía Guttman. El procedimiento del modelo es flexible. En las pruebas subjetivas el experimentador podría suponer que las valoraciones son uniformes o que varían según el observador, el tiempo o ambos. Es decir, se podría considerar que las calificaciones de los objetos son constantes a lo largo del tiempo o que varían dependiendo del observador o del tiempo (Bradley and Terry, 1952). Existen muchas extensiones de este modelo, en particular, la popular extensión a las comparaciones múltiples llamada modelo Plackett-Luce (Luce 1959, Plackett 1975), define una distribución a priori sobre permutaciones y se ha utilizado para la clasificación de múltiples individuos (Caron and Doucet, 2012).



# Capítulo 2

## Modelo Bradley-Terry

### 2.1. Historia del Modelo Bradley-Terry

Ralph A. Bradley nació en Ontario, Canadá, el 28 de noviembre de 1923, aunque creció en Wellington, un pueblo a orillas del lago Ontario. Se graduó en Matemáticas y Física en la Universidad de Queen en el año 1944. Mientras estaba estudiando en Queen comenzó la Segunda Guerra Mundial (1939-1945) y se requería que todo varón sin discapacidad tomara entrenamiento militar en la Universidad por lo que una vez acabados sus estudios, Bradley estuvo en el ejército canadiense entre 1944 y 1945, regresó después a Queen para completar su especialización académica en Estadística. En 1946 decidió unirse al nuevo departamento de la Universidad de Carolina del Norte, Chapel Hill (UNC), pues tenía un nuevo programa de doctorado en Estadística teórica y allí consiguió doctorarse. Su primer trabajo académico fue en la Universidad McGill en 1949, pero sintió que en esta Universidad no avanzaría profesionalmente, así que, un año más tarde, se mudó a Virginia donde ejerció como profesor en el Instituto Politécnico de Virginia (VPI). Allí le encargaron un proyecto para el Departamento de Agricultura de EE.UU. que trataba de una investigación sobre pruebas de sabor. Milton Terry trabajó con él en este proyecto. Uno de los problemas que Bradley y Terry encontraron en las pruebas de sabor era que un catador no podía probar muchas muestras en una sola sesión, ya que podría aparecer lo que denominaron fatiga gustativa. También, hubo un problema de transferencia de sabores de una muestra a la siguiente. Estos inconvenientes le llevaron a pensar que necesitarían usar un diseño por bloques incompleto. El diseño por bloques incompleto más simple tendría una observación en cada uno de los dos tratamientos por bloque, esto les llevó a un diseño de pruebas de sabor con comparaciones pareadas. El famoso modelo Bradley-Terry. Aunque el modelo Bradley-Terry se introdujo por primera vez para modelar datos sobre preferencias, también hay referencias a torneos o juegos de competición. Por ejemplo, la primera referencia anterior a Bradley-Terry a este modelo, aparece en Zermello (1929) dónde se interesó en la clasificación de los ajedrecitas (Hollander, 2000).

## 2.2. Aplicaciones del Modelo

Ralph A. Bradley en una entrevista concedida a Hollander en 2001 (Hollander, 2000) explicó que su modelo nació a raíz de una investigación sobre evaluaciones sensoriales del sabor, el olor, el aspecto e incluso la sensación que producen en la boca los alimentos. Desde entonces, las aplicaciones de este modelo han sido numerosas y en muchos ámbitos diferentes.

Matthews y Morris en 1995 publicaron un estudio sobre la medición del dolor en pacientes. Con este estudio querían medir el dolor experimentado en la administración de inyecciones subcutáneas rutinarias en pacientes sometidos a hemodiálisis regularmente, para ello contaban con cuatro tratamientos diferentes, con estructura factorial de  $2 \times 2$ . Un primer factor comprendía dos preparados distintos de la sustancia correctora de la anemia, la eritropoyetina, que se creía que la inyección de uno de ellos era considerablemente más dolorosa que la del otro y el otro factor era la aplicación en el lugar de la inyección de una crema anestésica o de un placebo. En cada visita a la unidad de diálisis el paciente recibe 2 tratamientos de los 4 posibles y se le pide que los compare. Los cuatro tratamientos posibles son: 1) Inyección A con crema activa, 2) Inyección B con crema activa, 3) Inyección A con crema placebo y 4) Inyección B con crema placebo. Estas comparaciones pareadas fueron analizadas mediante el modelo de Bradley-Terry teniendo en cuenta el orden en que se suministran los tratamientos y el caso de empate entre ellos (Matthews and Morris, 1995).

En 1998 Atkinson, Wampold, Lowe, Matthews y Ahn publicaron un estudio sobre las características que la población asiático-americana querían que tuvieran sus consejeros. Para ello haciendo uso del modelo Bradley-Terry compararon por pares características muy dispares de los consejeros y concluyeron que en igualdad de condiciones preferían consejeros étnicamente similares a ellos (Atkinson et al., 1998).

En genética el estudio que llevaron a cabo Sham y Curtis en 1995 sobre la transmisión de los alelos de los padres a su descendencia se basa en el modelo Bradley-Terry en el que los “individuos” a comparar son en este caso los alelos (Sham and Curtis, 1995).

Agresti en 2002 utilizó este modelo para clasificar a los equipos que compiten en béisbol, y estudió el problema que podría generar el orden en que se presentan los individuos, ya que en contextos deportivos el efecto de “jugar en casa” generalmente supone una ventaja para los individuos o equipos que compitan (Agresti, 2002).

Mazzucchi, Linzey y Bruning en 2008 aplicaron el modelo Bradley-Terry a un problema de fiabilidad. Pidieron a diferentes expertos en cableado eléctrico que distinguieran cuál sería el escenario más arriesgado para ellos comparados por pares para determinar la probabilidad de fallo de los cables en función de los factores que influyen en el entorno de un avión (Cattelan, 2012).

Como hemos podido comprobar el modelo Bradley-Terry ha tenido a lo largo de la historia numerosas y diversas aplicaciones.

## 2.3. Modelo

Consideremos un conjunto de  $m$  elementos. Para comparar entre sí dos elementos  $i$  y  $j$  de este conjunto, Bradley y Terry sugirieron el siguiente modelo (Bradley and Terry, 1952) :

$$P(\text{el individuo } i \text{ le gane al individuo } j) = \frac{\pi_i}{\pi_i + \pi_j}$$

donde  $\pi_i$  es un parámetro de valor positivo. Si consideramos a los individuos como equipos deportivos,  $\pi_i$  representa la habilidad general del equipo  $i$  (Hunter, 2004). Por convenio supondremos que  $\sum \pi_i = 1$  (Bradley and Terry, 1952). Este modelo sólo contempla que un equipo pueda perder o ganar, no existen los empates.

## 2.4. Axioma de Luce

Existen muchas extensiones del modelo Bradley Terry, en particular, la popular extensión a las comparaciones múltiples, como consecuencia del axioma que estableció Luce(1959), ya que inicialmente el modelo solo se usaba para comparar dos elementos.

Cuando una persona elige entre varias alternativas en un conjunto, muy a menudo sus respuestas parecen guiarse por probabilidades condicionadas al conjunto de elección. Luce (Luce, 2008) defendió que en este caso esto no es necesariamente así y desarrolló su axioma de elección:

Supongamos que  $P_S(R)$  denota la probabilidad ( o más genérico, el peso) de que la elección de una persona individual entre un conjunto finito de  $S$  alternativas caiga dentro de un grupo  $R \subset S$ .

Por ejemplo, si  $S$  representa las seis caras de un dado y  $R=\{1,6\}$ , entonces  $P_S(R)$  es la probabilidad de que una persona piense que saldrá un 1 o un 6 al lanzar el dado.

Para  $S = \{x, y\}$  se usa la abreviatura  $P(x, y) := P_S = P_{\{x,y\}}(x)$  esto denota la probabilidad de elegir  $x$  antes que  $y$ .

Sea  $N$  un subconjunto finito tal que, para todo  $S \subset N$ ,  $P_S$  es definida como:

1. Si  $P(x, y) \neq 0, 1$  para todo  $x, y \in N$ , entonces  $R \subset S \subset N$  y

$$P_N(R) = P_S(R)P_N(S)$$

La probabilidad  $P_N(R)$ , que  $R$  sea elegida en  $N$ , es igual al producto de la probabilidad  $P_S(R)$  de que sea elegida  $R$  de  $S$  y la probabilidad  $P_N(S)$  de que sea elegida  $S$  de  $N$ .

2. Si  $P(x, y) = 0$ , para algún  $x, y \in N$ , entonces para todo  $S \subset N$ ,

$$P_N(S) = P_{N \setminus \{x\}}(S \setminus \{x\})$$

Por lo tanto llegamos a la conclusión de que la probabilidad de elegir un elemento del conjunto de todas las alternativas no se ve afectada por la ausencia o la presencia de otros elementos en los subconjuntos a comparar que consideremos.

Una de las consecuencias más interesantes del axioma de elección propuesto por Luce es el propio modelo de Bradley-Terry.

Luce expuso que si tenemos un conjunto  $N$  de elementos a comparar  $N = 1, \dots, n$  y, sea  $P_X(x)$  la probabilidad de elegir  $x$  del conjunto de elementos  $X$ . Existen valores positivos  $v_1, \dots, v_n$  tal que

$$P_S(i) = \frac{v_i}{\sum_{j \in S} v_j}$$

$\forall i \in S \subseteq N$ ,  $v_1, \dots, v_n$  únicos salvo multiplicación por constante.

Debería ser evidente que si tomamos  $S$  como un subconjunto de las comparaciones por pares de los elementos de  $N$ , llegamos al modelo de Bradley-Terry. Por lo tanto las comparaciones por pares siguiendo el modelo Bradley-Terry siguen una estructura clara, el denominado axioma de Luce.

Esto nos garantiza que al usar nuestro modelo Bradley-Terry no habrá discrepancias en las elecciones, lo que podría ser frecuente si tratamos con elección de gustos, no es frecuente cuando hablamos de competiciones deportivas.

## 2.5. Estimación de los parámetros de máxima verosimilitud

Supondremos que observamos un conjunto de enfrentamientos por pares entre  $m$  individuos o equipos y deseamos estimar los parámetros  $\pi_1, \dots, \pi_m$  usando para ello el método de máxima verosimilitud.

Para un conjunto de  $m$  individuos hay  $\binom{m}{2} = \frac{m(m-1)}{2}$  enfrentamientos posibles.

Sea  $n_{ij}$  el número de veces que el equipo  $i$  se ha enfrentado al equipo  $j$ , denotemos por  $w_{ij}$  el número de veces que el equipo  $i$  ha ganado a  $j$ , tenemos que  $n_{ij} = w_{ij} + w_{ji} = n_{ji}$  y el número de veces que ganó el equipo  $i$  será entonces  $w_i = \sum_{j=1, j \neq i}^m w_{ij}$  (Caron and Doucet, 2012).

Si suponemos independencias entre los enfrentamientos, la función de máxima verosimilitud basándose en el modelo de Bradley-Terry viene dada por:



$$\begin{aligned}
 L(\pi_1, \dots, \pi_m) &= \prod_{i < j} \prod_{i=1}^m \left( \frac{\pi_i}{\pi_i + \pi_j} \right)^{w_{ij}} \left( \frac{\pi_j}{\pi_i + \pi_j} \right)^{w_{ji}} \\
 &= \prod_{i < j} \prod_{i=1}^m \left( \frac{\pi_i^{w_{ij}}}{(\pi_i + \pi_j)^{w_{ij}}} \right) \left( \frac{\pi_j^{w_{ji}}}{(\pi_i + \pi_j)^{w_{ji}}} \right) \\
 &= \prod_{i < j} \prod_{i=1}^m \pi_i^{w_{ij}} \pi_j^{w_{ji}} \prod_{i < j} \prod_{i=1}^m \left( \frac{1}{(\pi_i + \pi_j)^{w_{ij} + w_{ji}}} \right) \\
 &= \prod_{i=1}^m \pi_i^{\sum_{j=i+1}^m w_{ij}} \prod_{j=1}^m \pi_j^{\sum_{i=1}^{j-1} w_{ji}} \prod_{i < j} \prod_{i=1}^m (\pi_i + \pi_j)^{-n_{ij}} \\
 &= \prod_{i=1}^m \pi_i^{\sum_{j=i+1}^m w_{ij}} \prod_{i=1}^m \pi_i^{\sum_{j=1}^{i-1} w_{ij}} \prod_{i < j} \prod_{i=1}^m (\pi_i + \pi_j)^{-n_{ij}} \\
 &= \prod_{i=1}^m \pi_i^{\sum_{j=1}^m w_{ij}} \prod_{i < j} \prod_{i=1}^m (\pi_i + \pi_j)^{-n_{ij}} \\
 &= \prod_{i=1}^m \pi_i^{w_i} \prod_{i < j} \prod_{i=1}^m (\pi_i + \pi_j)^{-n_{ij}}
 \end{aligned}$$

Para obtener el estimador de máxima verosimilitud de los parámetros  $\pi_1, \dots, \pi_m$ , vamos a escribir el logaritmo de la verosimilitud:  $l(\pi_1, \dots, \pi_m) = \ln(L(\pi_1, \dots, \pi_m))$ , a  $l(\pi_1, \dots, \pi_m)$  se le llama log-verosimilitud (Ibarrola and Perez, 2002).

$$\begin{aligned}
 l(\pi_1, \dots, \pi_m) &= \ln(L(\pi_1, \dots, \pi_m)) \\
 &= \ln\left(\prod_{i=1}^m \pi_i^{w_i} \prod_{i < j} \prod_{i=1}^m (\pi_i + \pi_j)^{-n_{ij}}\right) \\
 &= \ln\left(\prod_{i=1}^m \pi_i^{w_i}\right) + \ln\left(\prod_{i < j} \prod_{i=1}^m ((\pi_i + \pi_j)^{-n_{ij}})\right) \\
 &= \sum_{i=1}^m \ln(\pi_i^{w_i}) + \sum_{i < j} \sum_{i=1}^m \ln((\pi_i + \pi_j)^{-n_{ij}}) \\
 &= \sum_{i=1}^m w_i \ln(\pi_i) - \sum_{i < j} \sum_{i=1}^m n_{ij} \ln(\pi_i + \pi_j)
 \end{aligned}$$

Las ecuaciones de verosimilitud son entonces

$$\frac{\partial}{\partial \pi_i} l(\pi_1, \dots, \pi_m) = \frac{w_i}{\pi_i} - \sum_{j=1, j \neq i}^m \frac{n_{ij}}{\pi_i + \pi_j} = 0$$

## 2.6. Modelo de Bradley-Terry como un modelo lineal generalizado

Un modelo lineal generalizado consta de 3 componentes: Una funcion de distribución, un predictor lineal y una funcion de enlace. En nuestro caso tendremos un vector  $Y$  con  $N$

observaciones, cada  $Y_i$  será el número de éxitos en un número fijo de ensayos que siguen una distribución binomial,  $\mu$  el vector de las probabilidades de  $Y$ , un vector de parámetros desconocidos  $\beta$ , la matriz del modelo  $X$ , y la función denominada función de enlace  $g$ . La ecuación definitiva del modelo lineal generalizado es:

$$g(\mu) = X\beta$$

Representaremos el modelo Bradley-Terry como un modelo lineal generalizado, y estimaremos así los parámetros desconocidos  $\beta$  a través de métodos conocidos para los modelos lineales generalizados. Bradley (Bradley, 1965) estableció que el modelo Bradley-Terry podía ser expresado como un modelo de regresión logística tomando como predictor lineal  $X\beta = \ln(\pi_i) - \ln(\pi_j)$ . Entonces si tomamos que el número de éxitos, en nuestro caso el número de veces que el elemento  $i$  gana al elemento  $j$  en las comparaciones entre  $(i,j)$ , sigue una distribución binomial  $Y_{ij} \sim \text{Bin}(n_{ij}, p_{ij})$ , con  $n_{ij}$  en número de comparaciones del par  $(i,j)$  y  $p_{ij}$  la probabilidad de que el elemento  $i$  gane al elemento  $j$ , por lo que al seguir una distribución binomial podemos representarla como un modelo de regresión logística tomando como función de enlace  $\text{logit}$ . Con

$$\text{logit}(p_{ij}) = \ln\left(\frac{p_{ij}}{1-p_{ij}}\right) = X\beta$$

Para expresar  $p_{ij}$  en función de los parámetros desconocidos  $\beta$  tomando  $e$ .

$$\begin{aligned} e^{\ln\left(\frac{p_{ij}}{1-p_{ij}}\right)} &= e^{X\beta} \\ \frac{p_{ij}}{1-p_{ij}} &= e^{X\beta} \\ p_{ij} &= e^{X\beta} - p_{ij}e^{X\beta} \\ p_{ij} + p_{ij}e^{X\beta} &= e^{X\beta} \\ (1 + e^{-X\beta})p_{ij} &= 1 \end{aligned}$$

$$p_{ij} = \frac{1}{(1 + e^{-X\beta})}$$

Tomando como predictor lineal  $X\beta = \ln(\pi_i) - \ln(\pi_j)$ . Comprobamos que efectivamente llegamos a nuestro modelo Bradley-Terry.

$$p_{ij} = \frac{1}{(1 + e^{-(\ln(\pi_i) - \ln(\pi_j))})} = \frac{1}{(1 + \frac{\pi_j}{\pi_i})} = \frac{\pi_i}{\pi_i + \pi_j}$$

Equivalentemente, podemos relacionar el modelo Bradley-Terry con un modelo de regresión logística tomando

$$\text{logit}(p_{ij}) = \ln\left(\frac{p_{ij}}{1-p_{ij}}\right) = \ln\left(\frac{\frac{\pi_i}{\pi_i + \pi_j}}{1 - \frac{\pi_i}{\pi_i + \pi_j}}\right) = \ln\left(\frac{\pi_i}{\pi_j}\right) = \ln(\pi_i) - \ln(\pi_j) = \beta_i - \beta_j \quad (2.1)$$

tenemos que  $\beta_i = \ln(\pi_i)$  que serán los parámetros desconocidos para mi modelo de regresión logística.

Tenemos entonces un modelo de regresión logística (GLM) con: (Critchlow and Fliagner, 1991)

- **Componente aleatoria:** Para  $i < j$ , siendo  $Y_{ij}$  el número de veces que el elemento  $i$  gana al elemento  $j$  en las  $n_{ij}$  comparaciones entre  $(i,j)$ . Todas las observaciones para los  $t$  componentes a comparar nos da el vector  $Y = (Y_{12}, Y_{13}, \dots, Y_{1t}, Y_{23}, \dots, Y_{2t}, \dots, Y_{(t-1)t})$ , con un tamaño  $N = \sum_{i=1}^{t-1} i = \frac{t(t-1)}{2}$ , cada  $Y_{ij} \sim Bin(n_{ij}, p_{ij})$ , con  $n_{ij}$  el número de veces que se compara el elemento  $i$  con el  $j$  y  $p_{ij}$  la probabilidad de ganar el elemento  $i$  al elemento  $j$ . Tenemos que  $E[Y_{ij}] = n_{ij}p_{ij}$ . El vector de las probabilidades de  $Y$  será el vector  $\mu = (p_{12}, p_{13}, \dots, p_{1t}, p_{23}, \dots, p_{2t}, \dots, p_{(t-1)t})$  con  $p_{ij} = E[\frac{Y_{ij}}{n_{ij}}]$ .
- **Predictor lineal.** El predictor lineal  $X\beta$  está formado por un vector de parámetros desconocidos  $\beta$  y la matriz del modelo  $X$ . El vector de parámetros desconocidos  $\beta$  será el formado por los  $\beta_i, \beta = (\beta_1, \dots, \beta_t)$  con  $\beta_i = \ln(\pi_i)$ . Y la matriz del modelo tendrá dimensión  $N = \frac{t(t-1)}{2}$  filas x  $t$  columnas. Como sabemos por (2.1) la  $k$ -ésima entrada en la  $(i,j)$ -ésima fila será 1 si  $k=i$ , -1 si  $k=j$  o 0 en cualquier otro caso.

$$X = \begin{pmatrix} 1 & -1 & 0 & \dots & 0 & 0 \\ 1 & 0 & -1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \\ 0 & 0 & 0 & \dots & 1 & -1 \end{pmatrix} \quad (2.2)$$

Tenemos

$$X\beta = \begin{pmatrix} 1 & -1 & 0 & \dots & 0 & 0 \\ 1 & 0 & -1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \\ 0 & 0 & 0 & \dots & 1 & -1 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_t \end{pmatrix} \quad (2.3)$$

- **Función de enlace:** La función de enlace para representar el modelo Bradley-Terry como un modelo de regresión logística es la función logit. Tenemos entonces :

$$\text{logit} \begin{pmatrix} p_{12} \\ p_{13} \\ \vdots \\ p_{t-1,t} \end{pmatrix} = X\beta = \begin{pmatrix} \beta_1 - \beta_2 \\ \beta_1 - \beta_3 \\ \vdots \\ \beta_{t-1} - \beta_t \end{pmatrix} \quad (2.4)$$



# Capítulo 3

## Modelo Bradley-Terry con empates

### 3.1. Modelo

El modelo Bradley-Terry tal y como lo conocemos hasta ahora no contempla que un elemento  $i$  del conjunto de  $m$  elementos a comparar no presente diferencias con otro elemento  $j$  del conjunto. Si los dos elementos no presentan diferencias entre ellos, expresar la preferencia del elemento  $i$  sobre el elemento  $j$  es poco realista.

Rao y Kupper (Rao and Kupper, 1967) consideraron una modificación del modelo Bradley-Terry mediante la introducción de un parámetro adicional, denominado parámetro de umbral que permitió hablar de “empates” entre elementos en el modelo.

Pocos años después, Davidson (Davidson, 1970) formuló su extensión del modelo Bradley-Terry en la que podían existir “empates” al expresar la preferencia entre dos elementos a comparar. Comprobó que su modelo cumplía el Axioma de Luce al igual que el modelo Bradley-Terry original.

Entonces Davidson expuso que:

Sea  $\sum \pi_i = 1$ , con  $\pi_i$  es un parámetro de valor positivo, asociado a la habilidad del elemento  $i$ .

$$P(\text{el individuo } i \text{ le gane al individuo } j) = \frac{\pi_i}{\pi_i + \pi_j + v\sqrt{\pi_i\pi_j}}$$
$$P(\text{el individuo } j \text{ le gane al individuo } i) = \frac{\pi_j}{\pi_i + \pi_j + v\sqrt{\pi_i\pi_j}}$$
$$P(\text{el individuo } i \text{ empate con el individuo } j) = \frac{v\sqrt{\pi_i\pi_j}}{\pi_i + \pi_j + v\sqrt{\pi_i\pi_j}}$$

donde la constante  $v \geq 0$  refleja la tendencia con la que se producen los empates que no depende ni de  $i$  ni de  $j$ .

Sea  $p_{ij}(i)$  la probabilidad de que gane el elemento  $i$  al  $j$ ,  $p_{ij}(j)$  la probabilidad de que gane el elemento  $j$  al  $i$  y  $p_{ij}(0)$  la probabilidad de que empaten ambos elementos, se cumple que  $p_{ij}(i)+p_{ij}(j)+p_{ij}(0)=1$ .

Como podemos ver, si  $v = 0$  tenemos el modelo Bradley-Terry original que no contemplaba empates.

## 3.2. Estimación de los parámetros de máxima verosimilitud

Supondremos que observamos un conjunto de enfrentamientos por pares entre  $m$  individuos de un conjunto y deseamos estimar los parámetros  $\pi_1, \dots, \pi_m$  y  $v$  usando para ello el método de máxima verosimilitud.

Sea  $n_{ij}$  el número de veces que el individuo  $i$  se ha enfrentado al individuo  $j$ , denotemos por  $w_{ij}$  el número de veces que el individuo  $i$  ha ganado a  $j$ ,  $w_{ji}$  el número de veces que el individuo  $j$  ha ganado a  $i$  y  $t_{ij}$  el número de veces que el individuo  $i$  ha empatado con  $j$ , tenemos que  $n_{ij} = w_{ij} + w_{ji} + t_{ij} = n_{ji}$ , el número de veces que ganó el individuo  $i$  será entonces  $w_i = \sum_{j=1, j \neq i}^m w_{ij}$  y el número de veces que empata el individuo  $i$  será  $t_i = \sum_{j=1, j \neq i}^m t_{ij}$ , siendo  $T$  el número total de empates en mi conjunto con  $T = \sum \sum_{i < j}^m t_{ij}$ , es evidente que  $t_{ij} = t_{ji}$ .

Si suponemos independencias entre los enfrentamientos, la función de máxima verosimilitud basándose en la extensión del modelo de Bradley-Terry expuesta por Davidson viene dada por:

$$\begin{aligned}
L(\pi_1, \dots, \pi_m; v) &= \prod_{i < j}^m \left( \frac{\pi_i}{\pi_i + \pi_j + v\sqrt{\pi_i\pi_j}} \right)^{w_{ij}} \left( \frac{\pi_j}{\pi_i + \pi_j + v\sqrt{\pi_i\pi_j}} \right)^{w_{ji}} \left( \frac{v\sqrt{\pi_i\pi_j}}{\pi_i + \pi_j + v\sqrt{\pi_i\pi_j}} \right)^{t_{ij}} \\
&= \prod_{i < j}^m \left( \frac{\pi_i^{w_{ij}}}{(\pi_i + \pi_j + v\sqrt{\pi_i\pi_j})^{w_{ij}}} \right) \left( \frac{\pi_j^{w_{ji}}}{(\pi_i + \pi_j + v\sqrt{\pi_i\pi_j})^{w_{ji}}} \right) \left( \frac{(v\sqrt{\pi_i\pi_j})^{t_{ij}}}{(\pi_i + \pi_j + v\sqrt{\pi_i\pi_j})^{t_{ij}}} \right) \\
&= \prod_{i < j}^m \pi_i^{w_{ij}} \pi_j^{w_{ji}} (v\sqrt{\pi_i\pi_j})^{t_{ij}} \prod_{i < j}^m \left( \frac{1}{(\pi_i + \pi_j + v\sqrt{\pi_i\pi_j})^{w_{ij} + w_{ji} + t_{ij}}} \right) \\
&= \prod_{i=1}^m \pi_i^{\sum_{j=i+1}^m w_{ij}} \prod_{j=1}^m \pi_j^{\sum_{i=1}^{j-1} w_{ji}} \prod_{i < j}^m v^{t_{ij}} \prod_{i < j}^m (\pi_i\pi_j)^{\frac{t_{ij}}{2}} \prod_{i < j}^m (\pi_i + \pi_j + v\sqrt{\pi_i\pi_j})^{-n_{ij}} \\
&= \prod_{i=1}^m \pi_i^{\sum_{j=i+1}^m w_{ij}} \prod_{i=1}^m \pi_i^{\sum_{j=1}^{i-1} w_{ji}} v^{\sum \sum_{i < j}^m t_{ij}} \prod_{i=1}^m \pi_i^{\sum_{j=i+1}^m \frac{t_{ij}}{2}} \prod_{j=1}^m \pi_j^{\sum_{i=1}^{j-1} \frac{t_{ji}}{2}} \\
&\quad \prod_{i < j}^m (\pi_i + \pi_j + v\sqrt{\pi_i\pi_j})^{-n_{ij}} \\
&= \prod_{i=1}^m \pi_i^{\sum_{j=1}^m w_{ij}} v^T \prod_{i=1}^m \pi_i^{\sum_{j=i+1}^m \frac{t_{ij}}{2}} \prod_{i=1}^m \pi_i^{\sum_{j=1}^{i-1} \frac{t_{ij}}{2}} \prod_{i < j}^m (\pi_i + \pi_j + v\sqrt{\pi_i\pi_j})^{-n_{ij}} \\
&= \prod_{i=1}^m \pi_i^{\sum_{j=1}^m w_{ij}} v^T \prod_{i=1}^m \pi_i^{\sum_{j=1}^m \frac{t_{ij}}{2}} \prod_{i < j}^m (\pi_i + \pi_j + v\sqrt{\pi_i\pi_j})^{-n_{ij}} \\
&= \prod_{i=1}^m \pi_i^{w_i} v^T \prod_{i=1}^m \pi_i^{\frac{1}{2}t_i} \prod_{i < j}^m (\pi_i + \pi_j + v\sqrt{\pi_i\pi_j})^{-n_{ij}} \\
&= \frac{v^T \prod_{i=1}^m \pi_i^{w_i + \frac{1}{2}t_i}}{\prod_{i < j}^m (\pi_i + \pi_j + v\sqrt{\pi_i\pi_j})^{n_{ij}}}
\end{aligned}$$

Para estimar por máxima verosimilitud, tomando logaritmo:

$$\begin{aligned}
 l(\pi_1, \dots, \pi_m; v) &= \ln(L(\pi_1, \dots, \pi_m; v)) \\
 &= \ln\left(v^T \prod_{i=1}^m \pi_i^{w_i + \frac{1}{2}t_i} \prod_{i < j}^m (\pi_i + \pi_j + v\sqrt{\pi_i\pi_j})^{-n_{ij}}\right) \\
 &= \ln(v^T) + \ln\left(\prod_{i=1}^m \pi_i^{w_i + \frac{1}{2}t_i}\right) + \ln\left(\prod_{i < j}^m (\pi_i + \pi_j + v\sqrt{\pi_i\pi_j})^{-n_{ij}}\right) \\
 &= T\ln(v) + \sum_{i=1}^m \ln(\pi_i^{w_i + \frac{1}{2}t_i}) + \sum_{i < j}^m \ln(\pi_i + \pi_j + v\sqrt{\pi_i\pi_j})^{-n_{ij}} \\
 &= T\ln(v) + \sum_{i=1}^m \left(w_i + \frac{1}{2}t_i\right) \ln(\pi_i) - \sum_{i < j}^m n_{ij} \ln(\pi_i + \pi_j + v\sqrt{\pi_i\pi_j})
 \end{aligned}$$

Las ecuaciones de verosimilitud son entonces:

$$\begin{aligned}
 \frac{\partial}{\partial \pi_i} l(\pi_1, \dots, \pi_m; v) &= \frac{w_i + \frac{1}{2}t_i}{\pi_i} - \sum_{j=1, j \neq i}^m \frac{n_{ij}}{2} \frac{2 + v\sqrt{\frac{\pi_j}{\pi_i}}}{\pi_i + \pi_j + v\sqrt{\pi_i\pi_j}} = 0 \\
 \frac{\partial}{\partial v} l(\pi_1, \dots, \pi_m; v) &= \frac{T}{v} - \sum_{i < j}^m \frac{n_{ij}\sqrt{\pi_i\pi_j}}{\pi_i + \pi_j + v\sqrt{\pi_i\pi_j}} = 0
 \end{aligned}$$

En el caso especial en el que  $m = 2$ , cuando mi conjunto sólo tenga dos elementos a comparar, las ecuaciones anteriores de verosimilitud pueden resolverse explícitamente. Sin embargo cuando  $m > 2$  la solución de las ecuaciones no puede obtenerse directamente y, por tanto, hay que emplear un proceso iterativo (Davidson, 1970).

### 3.3. Modelo Bradley-Terry con empates como un modelo lineal generalizado

Un modelo lineal generalizado consta de 3 componentes: Una función de distribución, un predictor lineal y una función de enlace.

La ecuación definitiva del modelo lineal generalizado es:

$$g(\mu) = X\theta$$

Con  $\mu$  vector de medias, un vector de parámetros desconocidos  $\theta$ , la matriz del modelo  $X$ , y la función denominada función de enlace  $g$ .

Representaremos el modelo Bradley-Terry extendido por Davidson incluyendo los empates como un modelo lineal generalizado, y estimaremos así los parámetros desconocidos  $\theta$  a través de métodos conocidos para los modelos lineales generalizados.

Antes de identificar cómo será el vector de parámetros  $\theta$  y la matriz del modelo,  $X$ , es necesario discutir cómo ajustaremos la distribución multinomial que tenemos a un modelo lineal generalizado, ya que la distribución multinomial no se modela como un modelo lineal generalizado directamente. De acuerdo al Th. 13.4-1 que desarrollaron Bishop, Fienberg y

Holland (Bishop et al., 1975), una distribución multinomial puede modelarse como un vector de variables aleatorias de Poisson independientes, introduciendo restricciones adecuadas en el modelo (Critchlow and Fligner, 1991).

Podríamos ver que cada enfrentamiento entre el individuo  $i$  y el individuo  $j$  sigue una distribución multinomial, en este caso trinomial con probabilidades,

P(el individuo  $i$  le gane al individuo  $j$ )

$$p_{ij1} = \frac{\pi_i}{\pi_i + \pi_j + v\sqrt{\pi_i\pi_j}}$$

P(el individuo  $j$  le gane al individuo  $i$ )

$$p_{ij2} = \frac{\pi_j}{\pi_i + \pi_j + v\sqrt{\pi_i\pi_j}}$$

P(el individuo  $i$  empate con el individuo  $j$ )

$$p_{ij3} = \frac{v\sqrt{\pi_i\pi_j}}{\pi_i + \pi_j + v\sqrt{\pi_i\pi_j}}$$

$$Y_{ij} = [Y_{ij1}|Y_{ij2}|Y_{ij3}] \sim Mult_3(n_{ij}, p_{ij} = [p_{ij1}|p_{ij2}|p_{ij3}])$$

Multiplicando estas probabilidades por  $n_{ij}$ , el número de comparaciones del par  $(i,j)$ , obtendremos el valor esperado del número de veces que gana  $i$ , el número de veces que gana  $j$  y el número de veces que empatan. Por lo que  $\mu = E[Y_{ijk}] = n_{ij}p_{ijk}$

La transformación Multinomial-Poisson (Baker, 1994) simplifica la estimación de los parámetros del modelo por máxima verosimilitud.

Tomaremos cada  $Y_{ijk}$  como el número de veces que al comparar el par  $(i,j)$  ocurre  $k$ , con  $k=1$  el número de veces que gana el individuo  $i$ , con  $k=2$  el número de veces que gana el individuo  $j$ , y con  $k=3$  el número de veces que empatan el individuo  $i$  con el individuo  $j$ . Entonces  $Y_{ijk} \sim Poi(n_{ij}p_{ijk})$  con  $E[Y_{ijk}] = n_{ij}p_{ijk} = \mu_{ijk}$ .

Y sabemos que podemos expresar una distribución de Poisson como un modelo lineal generalizado tomando como función de enlace el logaritmo neperiano por lo que resulta:

$$\ln(\mu) = X\theta$$

con  $\mu = (\mu_{121}, \mu_{122}, \mu_{123}, \dots, \mu_{(t-1)t1}, \mu_{(t-1)t2}, \mu_{(t-1)t3})$ .

Tenemos que  $\mu_{ijk} = n_{ij}p_{ijk}$ , sustituyendo:

$$\begin{aligned} \ln(\mu_{ij1}) &= \ln(n_{ij}p_{ij1}) \\ &= \ln(n_{ij}) + \ln(p_{ij1}) \\ &= \ln(n_{ij}) + \ln\left(\frac{\pi_i}{\pi_i + \pi_j + v\sqrt{\pi_i\pi_j}}\right) \\ &= \ln(n_{ij}) + \ln(\pi_i) - \ln(\pi_i + \pi_j + v\sqrt{\pi_i\pi_j}) \\ &= \ln(\pi_i) + \ln\left(\frac{n_{ij}}{\pi_i + \pi_j + v\sqrt{\pi_i\pi_j}}\right) \end{aligned}$$



$$\begin{aligned}
\ln(\mu_{ij2}) &= \ln(n_{ij}p_{ij2}) \\
&= \ln(n_{ij}) + \ln(p_{ij2}) \\
&= \ln(n_{ij}) + \ln\left(\frac{\pi_j}{\pi_i + \pi_j + v\sqrt{\pi_i\pi_j}}\right) \\
&= \ln(n_{ij}) + \ln(\pi_j) - \ln(\pi_i + \pi_j + v\sqrt{\pi_i\pi_j}) \\
&= \ln(\pi_j) + \ln\left(\frac{n_{ij}}{\pi_i + \pi_j + v\sqrt{\pi_i\pi_j}}\right)
\end{aligned}$$

$$\begin{aligned}
\ln(\mu_{ij3}) &= \ln(n_{ij}p_{ij3}) \\
&= \ln(n_{ij}) + \ln(p_{ij3}) \\
&= \ln(n_{ij}) + \ln\left(\frac{v\sqrt{\pi_i\pi_j}}{\pi_i + \pi_j + v\sqrt{\pi_i\pi_j}}\right) \\
&= \ln(n_{ij}) + \ln(v\sqrt{\pi_i\pi_j}) - \ln(\pi_i + \pi_j + v\sqrt{\pi_i\pi_j}) \\
&= \ln(v\sqrt{\pi_i\pi_j}) + \ln\left(\frac{n_{ij}}{\pi_i + \pi_j + v\sqrt{\pi_i\pi_j}}\right) \\
&= \ln(v) + \ln(\sqrt{\pi_i\pi_j}) + \ln\left(\frac{n_{ij}}{\pi_i + \pi_j + v\sqrt{\pi_i\pi_j}}\right) \\
&= \ln(v) + \frac{1}{2}\ln(\pi_i\pi_j) + \ln\left(\frac{n_{ij}}{\pi_i + \pi_j + v\sqrt{\pi_i\pi_j}}\right) \\
&= \ln(v) + \frac{1}{2}(\ln(\pi_i) + \ln(\pi_j)) + \ln\left(\frac{n_{ij}}{\pi_i + \pi_j + v\sqrt{\pi_i\pi_j}}\right)
\end{aligned}$$

donde  $\beta_i = \ln(\pi_i)$ ,  $\beta = \ln(v)$  y  $\alpha_{ij} = \ln\left(\frac{n_{ij}}{\pi_i + \pi_j + v\sqrt{\pi_i\pi_j}}\right)$  que será una constante de normalización. Tendremos  $\frac{t(t-1)}{2}$   $\alpha_{ij}$  incluidas en el vector de parámetros  $\theta$ , lo que hará que la matriz del modelo aumente  $\frac{t(t-1)}{2}$  columnas. Estas columnas introducirán una variable de clasificación, clasificando nuestras observaciones de Poisson con la distribución multinomial a la que corresponden. Por lo que la columna asociada a la constante de normalización  $\alpha_{ij}$  tendrá un 1 en las filas correspondientes a  $Y_{ij1}$ ,  $Y_{ij2}$ ,  $Y_{ij3}$  y 0 en las demás filas.

Tenemos que

$$\ln(\mu_{ij1}) = \beta_i + \alpha_{ij} \quad (3.1)$$

$$\ln(\mu_{ij2}) = \beta_j + \alpha_{ij} \quad (3.2)$$

$$\ln(\mu_{ij3}) = \beta + \frac{1}{2}(\beta_i + \beta_j) + \alpha_{ij} \quad (3.3)$$

donde  $\beta_i = \ln(\pi_i)$ ,  $\beta = \ln(v)$  y  $\alpha_{ij} = \ln\left(\frac{n_{ij}}{\pi_i + \pi_j + v\sqrt{\pi_i\pi_j}}\right)$  serán los parámetros desconocidos para mi modelo de regresión de Poisson.

Tenemos entonces un modelo de regresión de Poisson con:

- Componente aleatoria: Para  $i < j$ , siendo  $Y_{ijk}$  el número de veces que el elemento  $i$  gana ( $k=1$ ) al elemento  $j$ , el número de veces que el elemento  $j$  gana ( $k=2$ ) al elemento  $i$  o el número de veces que el elemento  $i$  empata ( $k=3$ ) con el elemento  $j$  en las  $n_{ij}$  comparaciones entre  $(i,j)$ . Todas las observaciones para los  $t$  componentes a comparar nos da el vector  $Y=(Y_{121}, Y_{122}, Y_{123}, \dots, Y_{(t-1)t1}, Y_{(t-1)t2}, Y_{(t-1)t3})$ , con un tamaño  $N = 3 \sum_{i=1}^{t-1} i = 3 \frac{t(t-1)}{2}$ , cada  $Y_{ijk} \sim Poi(n_{ij}p_{ijk})$ , con  $n_{ij}$  el número de veces que se compara el elemento  $i$  con el  $j$  y  $p_{ijk}$  la probabilidad de ganar ( $k=1$ ) el elemento  $i$  al elemento  $j$ , la probabilidad de ganar ( $k=2$ ) el elemento  $j$  al elemento  $i$  o la probabilidad de empatar ( $k=3$ ) el elemento  $i$  con el elemento  $j$ . Tenemos que  $E[Y_{ijk}] = n_{ij}p_{ijk}$ . El vector de las probabilidades de  $Y$  será el vector  $p_{ijk} = (p_{121}, p_{122}, p_{123}, \dots, p_{(t-1)t1}, p_{(t-1)t2}, p_{(t-1)t3})$  con  $p_{ijk} = E[\frac{Y_{ijk}}{n_{ij}}]$ .
- Predictor lineal. El predictor lineal  $X\theta$  está formado por un vector de parámetros desconocidos  $\theta$  y la matriz del modelo  $X$ . El vector de parámetros desconocidos  $\theta$  será el formado por los  $\beta_i, \beta$  y  $\alpha_{ij}$ , con  $\theta = (\beta_1, \dots, \beta_t, \beta, \alpha_{12}, \dots, \alpha_{(t-1)t})$ . Y la matriz del modelo tendrá dimensión  $N = 3 \frac{t(t-1)}{2}$  filas x  $t+1 + \frac{t(t-1)}{2}$  columnas.

Sea  $X$  la matriz del modelo, formada por  $X = [X_1 | X_2]$ .

La matriz  $X_1$  tendrá  $N=3 \frac{t(t-1)}{2}$  filas x  $(t+1)$  columnas. Como sabemos por (3.1), (3.2), (3.3) la  $m$ -ésima entrada en la  $(i,j,k)$ -ésima fila de la matriz  $X_1$  si  $k=1$  será 1 si  $m=i$  y 0 en caso contrario, si  $k=2$  será 1 si  $m=j$  y 0 en caso contrario, y si  $k=3$  será  $\frac{1}{2}$  si  $m=i$  o  $m=j$  y 1 si  $m=t+1$  y 0 en caso contrario.

La matriz  $X_2$  tendrá  $N=3 \frac{t(t-1)}{2}$  filas x  $\frac{t(t-1)}{2}$  columnas. Cada columna se asocia a la constante de normalización  $\alpha_{ij}$ , por lo que tendrá un 1 en las filas correspondientes a  $Y_{ij1}, Y_{ij2}, Y_{ij3}$  y 0 en las demás filas.

$$X = \begin{pmatrix} 1 & 0 & \dots & 0 & 0 & 0 & 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & 0 & 0 & 1 & 0 & \dots & 0 \\ \frac{1}{2} & \frac{1}{2} & \dots & 0 & 0 & 1 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 & 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & \dots & 0 & 1 & 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & \dots & \frac{1}{2} & \frac{1}{2} & 1 & 0 & 0 & \dots & 1 \end{pmatrix} \quad (3.4)$$

Tenemos

$$X\theta = \begin{pmatrix} 1 & 0 & \dots & 0 & 0 & 0 & 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & 0 & 0 & 1 & 0 & \dots & 0 \\ \frac{1}{2} & \frac{1}{2} & \dots & 0 & 0 & 1 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 & 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & \dots & 0 & 1 & 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & \dots & \frac{1}{2} & \frac{1}{2} & 1 & 0 & 0 & \dots & 1 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_{t-1} \\ \beta_t \\ \beta \\ \alpha_{12} \\ \alpha_{13} \\ \vdots \\ \alpha_{(t-1)t} \end{pmatrix} \quad (3.5)$$

- Función de enlace: La función de enlace para representar el modelo Bradley-Terry con empates como un modelo de regresión de Poisson es la función  $\ln$ .

Tenemos entonces :

$$\ln \begin{pmatrix} \mu_{121} \\ \mu_{122} \\ \mu_{123} \\ \vdots \\ \mu_{(t-1)t1} \\ \mu_{(t-1)t2} \\ \mu_{(t-1)t3} \end{pmatrix} = X\theta = \begin{pmatrix} \beta_1 + \alpha_{12} \\ \beta_2 + \alpha_{12} \\ \beta + \frac{1}{2}(\beta_1 + \beta_2) + \alpha_{12} \\ \vdots \\ \beta_{t-1} + \alpha_{(t-1)t} \\ \beta_t + \alpha_{(t-1)t} \\ \beta + \frac{1}{2}(\beta_{t-1} + \beta_t) + \alpha_{(t-1)t} \end{pmatrix} \quad (3.6)$$

(Critchlow and Fligner, 1991)



# Capítulo 4

## Modelo Bradley-Terry con ventaja común de jugar de local

### 4.1. Modelo

En las comparaciones entre pares de elementos de un conjunto de  $m$  elementos puede tener influencia el orden en que se presenten los elementos a comparar. Por ejemplo en pruebas de preferencia de sabores puede tener influencia el probar un sabor en primera o segunda posición, ya que el sabor del segundo puede estar condicionado por el sabor del primero, o por ejemplo en torneos deportivos de deportes como el fútbol tiene influencia que un equipo juegue como local o como visitante. Por ello en 1977 Davidson y Beaver (Davidson and Beaver, 1977) formularon una generalización del Modelo Bradley-Terry que contemplaba tanto los empates del modelo expuesto por Davidson como el efecto de orden entre los elementos a comparar. En el ámbito deportivo a esto se le conoce como “ventaja de jugar en casa”.

Sean  $i, j$  dos individuos del conjunto de  $m$  elementos a comparar y manteniendo que  $\sum \pi_i = 1$ , con  $\pi_i$  parámetro de valor positivo, asociado a la habilidad del elemento  $i$ , Davidson y Beaver sugirieron el siguiente modelo:

$$P(\text{el individuo } i \text{ le gane al individuo } j \text{ en casa}) = \frac{\gamma\pi_i}{\gamma\pi_i + \pi_j + v\sqrt{\gamma\pi_i\pi_j}}$$
$$P(\text{el individuo } j \text{ le gane al individuo } i \text{ fuera de casa}) = \frac{\pi_j}{\gamma\pi_i + \pi_j + v\sqrt{\gamma\pi_i\pi_j}}$$
$$P(\text{el individuo } i \text{ empate con el individuo } j \text{ en casa}) = \frac{v\sqrt{\gamma\pi_i\pi_j}}{\gamma\pi_i + \pi_j + v\sqrt{\gamma\pi_i\pi_j}}$$

donde la constante  $v \geq 0$  es el parámetro de empate y el nuevo parámetro  $\gamma \geq 0$  representa la ventaja de jugar en casa que tienen todos los equipos en común (Heather Turner and Robertson, 2013).

## 4.2. Estimación de los parámetros de máxima verosimilitud para ventaja de local común

Supondremos que observamos un conjunto de enfrentamientos por pares entre  $m$  individuos de un conjunto y deseamos estimar los parámetros  $\pi_1, \dots, \pi_m$ ,  $v$  y  $\gamma$  usando para ello el método de máxima verosimilitud.

Sea  $n_{ij}$  el número de veces que el individuo  $i$  se ha enfrentado al individuo  $j$  en casa de  $i$ . El número total de comparaciones entre individuos en el conjunto será

$$N = \sum_{i=1, i \neq j}^m \sum_{j=1, j \neq i}^m (n_{ij} + n_{ji})$$

Sea  $w_{ij}(1)$  el número de veces que el individuo  $i$  ha ganado a  $j$  jugando  $i$  como local  
 $w_{ij}(2)$  el número de veces que el individuo  $j$  ha ganado a  $i$  jugando  $i$  como local  
 $t_{ij}$  el número de veces que el individuo  $i$  ha empatado con  $j$  jugando  $i$  como local  
 $t_{ji}$  el número de veces que el individuo  $i$  ha empatado con  $j$  jugando  $j$  como local,  
entonces se tiene que

$$n_{ij} = w_{ij}(1) + w_{ij}(2) + t_{ij}$$

Entonces  $w_{iL}$  el número de veces que gana  $i$  jugando como local vendrá dado por

$$w_{iL} = \sum_{j=1, j \neq i}^m w_{ij}(1)$$

$w_{iV}$  el número de veces que gana  $i$  jugando como visitante vendrá dado por

$$w_{iV} = \sum_{j=1, j \neq i}^m w_{ji}(2)$$

$t_{iL}$  el número de veces que empata  $i$  jugando como local vendrá dado por

$$t_{iL} = \sum_{j=1, j \neq i}^m t_{ij}$$

$t_{iV}$  el número de veces que empata  $i$  jugando como visitante vendrá dado por

$$t_{iV} = \sum_{j=1, j \neq i}^m t_{ji}(2)$$

El número total de victorias locales en el conjunto de enfrentamientos será

$W_L = \sum_{i=1, i \neq j}^m \sum_{j=1, j \neq i}^m w_{ij}(1)$  y el número total de empates locales será

$T_L = \sum_{i=1, i \neq j}^m \sum_{j=1, j \neq i}^m t_{ij}$  (Davidson and Beaver, 1977).

Si suponemos independencias entre los enfrentamientos, la función de máxima verosimilitud basándose en la extensión del modelo de Bradley-Terry expuesta por Davidson y Beaver viene dada por:

$$\begin{aligned}
 L(\pi_1, \dots, \pi_m; v; \gamma) &= \prod_{\substack{i=1 \\ i \neq j}}^m \prod_{\substack{j=1 \\ j \neq i}}^m \left( \frac{\gamma \pi_i}{\gamma \pi_i + \pi_j + v \sqrt{\gamma \pi_i \pi_j}} \right)^{w_{ij}(1)} \left( \frac{\pi_j}{\gamma \pi_i + \pi_j + v \sqrt{\gamma \pi_i \pi_j}} \right)^{w_{ij}(2)} \left( \frac{v \sqrt{\gamma \pi_i \pi_j}}{\gamma \pi_i + \pi_j + v \sqrt{\gamma \pi_i \pi_j}} \right)^{t_{ij}} \\
 &= \prod_{\substack{i=1 \\ i \neq j}}^m \prod_{\substack{j=1 \\ j \neq i}}^m \frac{(\gamma \pi_i)^{w_{ij}(1)}}{(\gamma \pi_i + \pi_j + v \sqrt{\gamma \pi_i \pi_j})^{w_{ij}(1)}} \frac{\pi_j^{w_{ij}(2)}}{(\gamma \pi_i + \pi_j + v \sqrt{\gamma \pi_i \pi_j})^{w_{ij}(2)}} \frac{(v \sqrt{\gamma \pi_i \pi_j})^{t_{ij}}}{(\gamma \pi_i + \pi_j + v \sqrt{\gamma \pi_i \pi_j})^{t_{ij}}} \\
 &= \prod_{\substack{i=1 \\ i \neq j}}^m \prod_{\substack{j=1 \\ j \neq i}}^m (\gamma \pi_i)^{w_{ij}(1)} \pi_j^{w_{ij}(2)} (v \sqrt{\gamma \pi_i \pi_j})^{t_{ij}} \prod_{\substack{i=1 \\ i \neq j}}^m \prod_{\substack{j=1 \\ j \neq i}}^m \frac{1}{(\gamma \pi_i + \pi_j + v \sqrt{\gamma \pi_i \pi_j})^{w_{ij}(1)+w_{ij}(2)+t_{ij}}} \\
 &= \prod_{\substack{i=1 \\ i \neq j}}^m \prod_{\substack{j=1 \\ j \neq i}}^m \gamma^{w_{ij}(1)} \prod_{\substack{i=1 \\ i \neq j}}^m \prod_{\substack{j=1 \\ j \neq i}}^m \pi_i^{w_{ij}(1)} \prod_{\substack{i=1 \\ i \neq j}}^m \prod_{\substack{j=1 \\ j \neq i}}^m \pi_j^{w_{ij}(2)} \prod_{\substack{i=1 \\ i \neq j}}^m \prod_{\substack{j=1 \\ j \neq i}}^m v^{t_{ij}} \prod_{\substack{i=1 \\ i \neq j}}^m \prod_{\substack{j=1 \\ j \neq i}}^m (\gamma \pi_i \pi_j)^{\frac{t_{ij}}{2}} \\
 &\quad \prod_{\substack{i=1 \\ i \neq j}}^m \prod_{\substack{j=1 \\ j \neq i}}^m \frac{1}{(\gamma \pi_i + \pi_j + v \sqrt{\gamma \pi_i \pi_j})^{n_{ij}}} \\
 &= \gamma^{\sum_{\substack{i=1 \\ i \neq j}}^m \sum_{\substack{j=1 \\ j \neq i}}^m w_{ij}(1)} \prod_{\substack{i=1 \\ i \neq j}}^m \pi_i^{\sum_{\substack{j=1 \\ j \neq i}}^m w_{ij}(1)} \prod_{\substack{i=1 \\ i \neq j}}^m \prod_{\substack{j=1 \\ j \neq i}}^m \pi_j^{w_{ij}(2)} v^{\sum_{\substack{i=1 \\ i \neq j}}^m \sum_{\substack{j=1 \\ j \neq i}}^m t_{ij}} \gamma^{\sum_{\substack{i=1 \\ i \neq j}}^m \sum_{\substack{j=1 \\ j \neq i}}^m \frac{t_{ij}}{2}} \\
 &\quad \prod_{\substack{i=1 \\ i \neq j}}^m \pi_i^{\sum_{\substack{j=1 \\ j \neq i}}^m \frac{t_{ij}}{2}} \prod_{\substack{i=1 \\ i \neq j}}^m \prod_{\substack{j=1 \\ j \neq i}}^m \pi_j^{\frac{t_{ij}}{2}} \prod_{\substack{i=1 \\ i \neq j}}^m \prod_{\substack{j=1 \\ j \neq i}}^m \frac{1}{(\gamma \pi_i + \pi_j + v \sqrt{\gamma \pi_i \pi_j})^{n_{ij}}} \\
 &= \gamma^{W_L} \prod_{\substack{i=1 \\ i \neq j}}^m \pi_i^{w_{iL}} \prod_{\substack{i=1 \\ i \neq j}}^m \prod_{\substack{j=1 \\ j \neq i}}^m \pi_j^{w_{ij}(2)} v^{T_L} \gamma^{\frac{T_L}{2}} \prod_{\substack{i=1 \\ i \neq j}}^m \pi_i^{\frac{t_{iL}}{2}} \prod_{\substack{i=1 \\ i \neq j}}^m \prod_{\substack{j=1 \\ j \neq i}}^m \pi_j^{\frac{t_{ij}}{2}} \prod_{\substack{i=1 \\ i \neq j}}^m \prod_{\substack{j=1 \\ j \neq i}}^m \frac{1}{(\gamma \pi_i + \pi_j + v \sqrt{\gamma \pi_i \pi_j})^{n_{ij}}} \\
 &= \frac{\gamma^{(W_L + \frac{T_L}{2})} v^{T_L} \prod_{\substack{i=1 \\ i \neq j}}^m \pi_i^{(w_{iL} + \frac{t_{iL}}{2})} \prod_{\substack{i=1 \\ i \neq j}}^m \prod_{\substack{j=1 \\ j \neq i}}^m \pi_j^{(w_{ij}(2) + \frac{t_{ij}}{2})}}{\prod_{\substack{i=1 \\ i \neq j}}^m \prod_{\substack{j=1 \\ j \neq i}}^m (\gamma \pi_i + \pi_j + v \sqrt{\gamma \pi_i \pi_j})^{n_{ij}}}
 \end{aligned}$$

Para estimar por máxima verosimilitud, tomando logaritmo:

$$\begin{aligned}
 l(\pi_1, \dots, \pi_m; v; \gamma) &= \ln(L(\pi_1, \dots, \pi_m; v; \gamma)) \\
 &= \ln(\gamma^{(W_L + \frac{T_L}{2})} v^{T_L} \prod_{\substack{i=1 \\ i \neq j}}^m \pi_i^{(w_{iL} + \frac{t_{iL}}{2})} \prod_{\substack{i=1 \\ i \neq j}}^m \prod_{\substack{j=1 \\ i \neq j}}^m \pi_j^{(w_{ij}(2) + \frac{t_{ij}}{2})} \prod_{\substack{i=1 \\ i \neq j}}^m \prod_{\substack{j=1 \\ i \neq j}}^m (\gamma \pi_i + \pi_j + v \sqrt{\gamma \pi_i \pi_j})^{-n_{ij}}) \\
 &= \ln(\gamma^{(W_L + \frac{T_L}{2})}) + \ln(v^{T_L}) + \sum_{\substack{i=1 \\ i \neq j}}^m \ln(\pi_i^{(w_{iL} + \frac{t_{iL}}{2})}) + \sum_{\substack{i=1 \\ i \neq j}}^m \sum_{\substack{j=1 \\ i \neq j}}^m \ln(\pi_j^{(w_{ij}(2) + \frac{t_{ij}}{2})}) \\
 &\quad + \sum_{\substack{i=1 \\ i \neq j}}^m \sum_{\substack{j=1 \\ i \neq j}}^m \ln(\gamma \pi_i + \pi_j + v \sqrt{\gamma \pi_i \pi_j})^{-n_{ij}} \\
 &= (W_L + \frac{T_L}{2}) \ln(\gamma) + T_L \ln(v) + \sum_{\substack{i=1 \\ i \neq j}}^m (w_{iL} + \frac{1}{2} t_{iL}) \ln(\pi_i) + \sum_{\substack{i=1 \\ i \neq j}}^m \sum_{\substack{j=1 \\ i \neq j}}^m (w_{ij}(2) + \frac{t_{ij}}{2}) \ln(\pi_j) \\
 &\quad + \sum_{\substack{i=1 \\ i \neq j}}^m \sum_{\substack{j=1 \\ i \neq j}}^m -n_{ij} \ln(\gamma \pi_i + \pi_j + v \sqrt{\gamma \pi_i \pi_j})
 \end{aligned}$$

Las ecuaciones de verosimilitud son entonces:

$$\begin{aligned}
 \frac{\partial}{\partial \pi_i} l(\pi_1, \dots, \pi_m; v; \gamma) &= \frac{w_{iL} + \frac{1}{2} t_{iL}}{\pi_i} + \frac{\partial}{\partial \pi_i} \sum_{\substack{j=1 \\ i \neq j}}^m (w_{ij}(1) + \frac{1}{2} t_{ji}) \ln(\pi_i) \\
 &\quad - \frac{\partial}{\partial \pi_i} \sum_{\substack{i=1 \\ i \neq j}}^m \sum_{\substack{j=1 \\ i \neq j}}^m (n_{ij} \ln(\gamma \pi_i + \pi_j + v \sqrt{\gamma \pi_i \pi_j}) + n_{ji} \ln(\gamma \pi_j + \pi_i + v \sqrt{\gamma \pi_i \pi_j})) \\
 &= \frac{w_{iL} + \frac{1}{2} t_{iL}}{\pi_i} + \frac{w_{iL} + \frac{1}{2} t_{iV}}{\pi_i} - \frac{1}{2} \left( \sum_{\substack{j=1 \\ i \neq j}}^m \frac{n_{ij} (2\gamma + v \sqrt{(\frac{\gamma \pi_j}{\pi_i})})}{\gamma \pi_i + \pi_j + v \sqrt{\gamma \pi_i \pi_j}} + \sum_{\substack{j=1 \\ i \neq j}}^m \frac{n_{jj} (2 + v \sqrt{(\frac{\gamma \pi_j}{\pi_i})})}{\gamma \pi_j + \pi_i + v \sqrt{\gamma \pi_i \pi_j}} \right) \\
 &= \frac{1}{2} \frac{4w_{iL} + t_{iL} + t_{iV}}{\pi_i} - \frac{1}{2} \left( \sum_{\substack{j=1 \\ i \neq j}}^m \frac{n_{ij} (2\gamma + v \sqrt{(\frac{\gamma \pi_j}{\pi_i})})}{\gamma \pi_i + \pi_j + v \sqrt{\gamma \pi_i \pi_j}} + \sum_{\substack{j=1 \\ i \neq j}}^m \frac{n_{ji} (2 + v \sqrt{(\frac{\gamma \pi_j}{\pi_i})})}{\gamma \pi_j + \pi_i + v \sqrt{\gamma \pi_i \pi_j}} \right) = 0
 \end{aligned}$$

$$\frac{\partial}{\partial v} l(\pi_1, \dots, \pi_m; v; \gamma) = \frac{T_L}{v} - \sum_{\substack{i=1 \\ i \neq j}}^m \sum_{\substack{j=1 \\ i \neq j}}^m \frac{n_{ij} \sqrt{\gamma \pi_i \pi_j}}{\gamma \pi_i + \pi_j + v \sqrt{\gamma \pi_i \pi_j}} = 0$$

$$\begin{aligned}
 \frac{\partial}{\partial \gamma} l(\pi_1, \dots, \pi_m; v; \gamma) &= \frac{W_L + \frac{T_L}{2}}{\gamma} - \sum_{\substack{i=1 \\ i \neq j}}^m \sum_{\substack{j=1 \\ i \neq j}}^m \frac{n_{ij} (\pi_i + \frac{v \pi_i \pi_j}{2 \sqrt{\gamma \pi_i \pi_j}})}{\gamma \pi_i + \pi_j + v \sqrt{\gamma \pi_i \pi_j}} \\
 &= \frac{1}{2} \frac{(2W_L + T_L)}{\gamma} - \sum_{\substack{i=1 \\ i \neq j}}^m \sum_{\substack{j=1 \\ i \neq j}}^m \frac{n_{ij}}{2} \frac{(2\pi_i + v \sqrt{\frac{\pi_i \pi_j}{\gamma}})}{\gamma \pi_i + \pi_j + v \sqrt{\gamma \pi_i \pi_j}} = 0
 \end{aligned}$$



### 4.3. Modelo Bradley-Terry con ventaja de jugar de local como un modelo lineal generalizado

Un modelo lineal generalizado consta de 3 componentes: Una función de distribución, un predictor lineal y una función de enlace.

La ecuación definitiva del modelo lineal generalizado es:

$$g(\mu) = X\theta$$

Con  $\mu$  vector de medias, un vector de parámetros desconocidos  $\theta$ , la matriz del modelo  $X$ , y la función denominada función de enlace  $g$ .

Representaremos el modelo Bradley-Terry extendido por Davidson y Beaver incluyendo los empates y la ventaja de jugar como local como un modelo lineal generalizado, y estimaremos así los parámetros desconocidos  $\theta$  a través de métodos conocidos para los modelos lineales generalizados.

Antes de identificar cómo será el vector de parámetros  $\theta$  y la matriz del modelo,  $X$ , es necesario discutir cómo ajustaremos la distribución multinomial que tenemos a un modelo lineal generalizado, ya que la distribución multinomial no se modela como un modelo lineal generalizado directamente. De acuerdo al Th. 13.4-1 que desarrollaron Bishop, Fienberg y Holland (Bishop et al., 1975), una distribución multinomial puede modelarse como un vector de variables aleatorias de Poisson independientes, introduciendo restricciones adecuadas en el modelo (Critchlow and Fliener, 1991).

Podríamos ver que cada enfrentamiento entre el individuo  $i$  y el individuo  $j$  sigue una distribución multinomial, en este caso trinomial con probabilidades,

P(el individuo  $i$  le gane al individuo  $j$  jugando  $i$  como local)

$$p_{ij1} = \frac{\gamma\pi_i}{\gamma\pi_i + \pi_j + v\sqrt{\gamma\pi_i\pi_j}}$$

P(el individuo  $j$  le gane al individuo  $i$  jugando  $i$  como local)

$$p_{ij2} = \frac{\pi_j}{\gamma\pi_i + \pi_j + v\sqrt{\gamma\pi_i\pi_j}}$$

P(el individuo  $i$  empate con el individuo  $j$  jugando  $i$  como local)

$$p_{ij3} = \frac{v\sqrt{\gamma\pi_i\pi_j}}{\gamma\pi_i + \pi_j + v\sqrt{\gamma\pi_i\pi_j}}$$

$$Y_{ij} = [Y_{ij1}|Y_{ij2}|Y_{ij3}] \sim Mult_3(n_{ij}, p_{ij} = [p_{ij1}|p_{ij2}|p_{ij3}])$$

Multiplicando estas probabilidades por el número de comparaciones del par  $(i,j)$  siendo local  $i, n_{ij}$ , , obtendremos el valor esperado del número de veces que gana  $i$  a  $j$  siendo local  $i$ , el número de veces que gana  $j$  a  $i$  siendo local  $i$  y el número de veces que empatan siendo local  $i$ . Por lo que  $\mu = E[Y_{ijk}] = n_{ij}p_{ijk}$

La transformación Multinomial-Poisson (Baker, 1994) simplifica la estimación de los parámetros del modelo por máxima verosimilitud.

Tomaremos cada  $Y_{ijk}$  como el número de veces que al comparar el par  $(i,j)$  siendo local  $i$  ocurre  $k$ , con  $k=1$  el número de veces que gana el individuo  $i$  siendo local  $i$ , con  $k=2$  el número de veces

que gana el individuo  $j$  siendo local  $i$ , y con  $k=3$  el número de veces que empatan el individuo  $i$  con el individuo  $j$  siendo local  $i$ . Entonces  $Y_{ijk} \sim \text{Poisson}(n_{ij}p_{ijk})$  con  $E[Y_{ijk}] = n_{ij}p_{ijk} = \mu_{ijk}$ .

Y sabemos que podemos expresar una distribución de Poisson como un modelo lineal generalizado tomando como función de enlace el logaritmo neperiano por lo que resulta:

$$\ln(\mu) = X\theta$$

con  $\mu = (\mu_{121}, \mu_{122}, \mu_{123}, \dots, \mu_{t(t-1)1}, \mu_{t(t-1)2}, \mu_{t(t-1)3})$ .

Tenemos que  $\mu_{ijk} = n_{ij}p_{ijk}$ , sustituyendo:

$$\begin{aligned} \ln(\mu_{ij1}) &= \ln(n_{ij}p_{ij1}) \\ &= \ln(n_{ij}) + \ln(p_{ij1}) \\ &= \ln(n_{ij}) + \ln\left(\frac{\gamma\pi_i}{\gamma\pi_i + \pi_j + v\sqrt{\gamma\pi_i\pi_j}}\right) \\ &= \ln(n_{ij}) + \ln(\gamma\pi_i) - \ln(\gamma\pi_i + \pi_j + v\sqrt{\gamma\pi_i\pi_j}) \\ &= \ln(\gamma) + \ln(\pi_i) + \ln\left(\frac{n_{ij}}{\gamma\pi_i + \pi_j + v\sqrt{\gamma\pi_i\pi_j}}\right) \end{aligned}$$

$$\begin{aligned} \ln(\mu_{ij2}) &= \ln(n_{ij}p_{ij2}) \\ &= \ln(n_{ij}) + \ln(p_{ij2}) \\ &= \ln(n_{ij}) + \ln\left(\frac{\pi_j}{\gamma\pi_i + \pi_j + v\sqrt{\gamma\pi_i\pi_j}}\right) \\ &= \ln(n_{ij}) + \ln(\pi_j) - \ln(\gamma\pi_i + \pi_j + v\sqrt{\gamma\pi_i\pi_j}) \\ &= \ln(\pi_j) + \ln\left(\frac{n_{ij}}{\gamma\pi_i + \pi_j + v\sqrt{\gamma\pi_i\pi_j}}\right) \end{aligned}$$

$$\begin{aligned} \ln(\mu_{ij3}) &= \ln(n_{ij}p_{ij3}) \\ &= \ln(n_{ij}) + \ln(p_{ij3}) \\ &= \ln(n_{ij}) + \ln\left(\frac{v\sqrt{\gamma\pi_i\pi_j}}{\gamma\pi_i + \pi_j + v\sqrt{\gamma\pi_i\pi_j}}\right) \\ &= \ln(n_{ij}) + \ln(v\sqrt{\gamma\pi_i\pi_j}) - \ln(\gamma\pi_i + \pi_j + v\sqrt{\gamma\pi_i\pi_j}) \\ &= \ln(v\sqrt{\gamma\pi_i\pi_j}) + \ln\left(\frac{n_{ij}}{\gamma\pi_i + \pi_j + v\sqrt{\gamma\pi_i\pi_j}}\right) \\ &= \ln(v) + \ln(\sqrt{\gamma\pi_i\pi_j}) + \ln\left(\frac{n_{ij}}{\gamma\pi_i + \pi_j + v\sqrt{\gamma\pi_i\pi_j}}\right) \\ &= \ln(v) + \frac{1}{2}\ln(\gamma\pi_i\pi_j) + \ln\left(\frac{n_{ij}}{\gamma\pi_i + \pi_j + v\sqrt{\gamma\pi_i\pi_j}}\right) \\ &= \ln(v) + \frac{1}{2}(\ln(\gamma) + \ln(\pi_i) + \ln(\pi_j)) + \ln\left(\frac{n_{ij}}{\gamma\pi_i + \pi_j + v\sqrt{\gamma\pi_i\pi_j}}\right) \end{aligned}$$

donde  $\beta_i = \ln(\pi_i)$ ,  $\beta = \ln(v)$ ,  $\alpha_{ij} = \ln\left(\frac{n_{ij}}{\gamma\pi_i + \pi_j + v\sqrt{\gamma\pi_i\pi_j}}\right)$ ,  $\alpha_{ij} \neq \alpha_{ji}$  que será una constante de normalización y  $\psi = \ln(\gamma)$ . Tendremos  $t(t-1)$   $\alpha_{ij}$  incluidas en el vector de parámetros  $\theta$ , lo

que hará que la matriz del modelo aumente  $t(t-1)$  columnas. Estas columnas introducirán una variable de clasificación, clasificando nuestras observaciones de Poisson con la distribución multinomial a la que corresponden. Por lo que la columna asociada a la constante de normalización  $\alpha_{ij}$  tendrá un 1 en las filas correspondientes a  $Y_{ij1}, Y_{ij2}, Y_{ij3}$  y 0 en las demás filas.

$$\ln(\mu_{ij1}) = \beta_i + \psi + \alpha_{ij} \quad (4.1)$$

$$\ln(\mu_{ij2}) = \beta_j + \alpha_{ij} \quad (4.2)$$

$$\ln(\mu_{ij3}) = \beta + \frac{1}{2}(\beta_i + \beta_j + \psi) + \alpha_{ij} \quad (4.3)$$

donde  $\beta_i = \ln(\pi_i)$ ,  $\beta = \ln(v)$ ,  $\alpha_{ij} = \ln\left(\frac{n_{ij}}{\gamma\pi_i + \pi_j + v\sqrt{\gamma\pi_i\pi_j}}\right)$  y  $\psi = \ln(\gamma)$  serán los parámetros desconocidos para mi modelo de regresión de Poisson.

Tenemos entonces un modelo de regresión de Poisson con:

- **Componente aleatoria:** Para  $i \neq j$ , siendo  $Y_{ijk}$  el número de veces que el elemento  $i$  gana ( $k=1$ ) al elemento  $j$  siendo local  $i$ , el número de veces que el elemento  $j$  gana ( $k=2$ ) al elemento  $i$  siendo local  $i$  o el número de veces que el elemento  $i$  empata ( $k=3$ ) con el elemento  $j$  siendo local  $i$  en las  $n_{ij}$  comparaciones entre  $(i,j)$  siendo local  $i$ . Todas las observaciones para las  $t$  componentes a comparar nos da el vector  $Y=(Y_{121}, Y_{122}, Y_{123}, \dots, Y_{t(t-1)1}, Y_{t(t-1)2}, Y_{t(t-1)3})$ , con un tamaño  $N = 3t(t-1)$ , cada  $Y_{ijk} \sim Poi(n_{ij}p_{ijk})$ , con  $n_{ij}$  el número de veces que se compara el elemento  $i$  con el  $j$  siendo  $i$  local y  $p_{ijk}$  la probabilidad de ganar ( $k=1$ ) el elemento  $i$  al elemento  $j$  siendo  $i$  local, la probabilidad de ganar ( $k=2$ ) el elemento  $j$  al elemento  $i$  siendo  $i$  local o la probabilidad de empatar ( $k=3$ ) el elemento  $i$  con el elemento  $j$  siendo  $i$  local. Tenemos que  $E[Y_{ijk}] = n_{ij}p_{ijk}$ . El vector de las probabilidades de  $Y$  será el vector  $p_{ijk} = (p_{121}, p_{122}, p_{123}, \dots, p_{t(t-1)1}, p_{t(t-1)2}, p_{t(t-1)3})$  con  $p_{ijk} = E\left[\frac{Y_{ijk}}{n_{ij}}\right]$ .
- **Predictor lineal.** El predictor lineal  $X\theta$  está formado por un vector de parámetros desconocidos  $\theta$  y la matriz del modelo  $X$ . El vector de parámetros desconocidos  $\theta$  será el formado por los  $\beta_i, \beta, \psi, \alpha_{ij}$ , con  $\theta = (\beta_1, \dots, \beta_t, \psi, \beta, \alpha_{12}, \dots, \alpha_{t(t-1)})$ . Y la matriz del modelo tendrá dimensión  $N = 3t(t-1)$  filas x  $(t+2)+t(t-1)$  columnas.

Sea  $X$  la matriz del modelo, formada por  $X = [X_1|X_2]$ .

La matriz  $X_1$  tendrá  $N=3t(t-1)$  filas x  $(t+2)$  columnas.

Como sabemos por (4.1), (4.2), (4.3) la  $m$ -ésima entrada en la  $(i,j,k)$ -ésima fila de la matriz  $X_1$  si  $k=1$  será 1 si  $m=i, m=t+1$  y 0 en caso contrario, si  $k=2$  será 1 si  $m=j$  y 0 en caso contrario, y si  $k=3$  será  $\frac{1}{2}$  si  $m=i, m=j$  o  $m=t+1$  y 1 si  $m=t+2$  y 0 en caso contrario.

La matriz  $X_2$  tendrá  $N=3t(t-1)$  filas x  $t(t-1)$  columnas. Cada columna se asocia a la constante de normalización  $\alpha_{ij}$ , por lo que tendrá un 1 en las filas correspondientes a  $Y_{ij1}, Y_{ij2}, Y_{ij3}$  y 0 en las demás filas.

$$X = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & 0 & 1 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 & 0 & 0 & 1 & 0 & \cdots & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & \cdots & 0 & 0 & \frac{1}{2} & 1 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 & 1 & 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 1 & 0 & 0 & 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & 1 & 0 & 0 & \cdots & 1 \end{pmatrix} \quad (4.4)$$

Tenemos

$$X\theta = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & 0 & 1 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 & 0 & 0 & 1 & 0 & \cdots & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & \cdots & 0 & 0 & \frac{1}{2} & 1 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 & 1 & 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 1 & 0 & 0 & 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & 1 & 0 & 0 & \cdots & 1 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \vdots \\ \beta_{t-1} \\ \beta_t \\ \psi \\ \beta \\ \alpha_{12} \\ \alpha_{13} \\ \vdots \\ \alpha_{t(t-1)} \end{pmatrix} \quad (4.5)$$

- Función de enlace: La función de enlace para representar el modelo Bradley-Terry con empates y ventaja de local como un modelo de regresión de Poisson es la función ln.

Tenemos entonces :

$$\ln \begin{pmatrix} \mu_{121} \\ \mu_{122} \\ \mu_{123} \\ \vdots \\ \mu_{t(t-1)1} \\ \mu_{t(t-1)2} \\ \mu_{t(t-1)3} \end{pmatrix} = X\theta = \begin{pmatrix} \beta_1 + \psi + \alpha_{12} \\ \beta_2 + \alpha_{12} \\ \beta + \frac{1}{2}(\beta_1 + \beta_2 + \psi) + \alpha_{12} \\ \vdots \\ \beta_t + \psi + \alpha_{t(t-1)} \\ \beta_{t-1} + \alpha_{t(t-1)} \\ \beta + \frac{1}{2}(\beta_{t-1} + \beta_t + \psi) + \alpha_{t(t-1)} \end{pmatrix} \quad (4.6)$$

(Critchlow and Fligner, 1991)

# Capítulo 5

## Modelo Bradley-Terry con ventaja individual de jugar de local

### 5.1. Modelo

Como sabemos, en deportes como el fútbol la ventaja de jugar como local depende del equipo ya que hay equipos que por diversos factores como el público, la familiaridad con el entorno, los desplazamientos u otros factores tienen mayor ventaja al jugar como local que otros que tengan unas circunstancias diferentes. Es por ello que estudiaremos también el parámetro de ventaja de jugar como local individualizado para cada equipo y así podremos ver si existen diferencias entre ellos. Por lo que obtenemos el siguiente modelo: (Davidson and Beaver, 1977) (Jamil, 2010)

$$P(\text{el individuo } i \text{ le gane al individuo } j \text{ en casa}) = \frac{\gamma_i \pi_i}{\gamma_i \pi_i + \pi_j + v \sqrt{\gamma_i \pi_i \pi_j}}$$
$$P(\text{el individuo } j \text{ le gane al individuo } i \text{ fuera de casa}) = \frac{\pi_j}{\gamma_i \pi_i + \pi_j + v \sqrt{\gamma_i \pi_i \pi_j}}$$
$$P(\text{el individuo } i \text{ empate con el individuo } j \text{ en casa}) = \frac{v \sqrt{\gamma_i \pi_i \pi_j}}{\gamma_i \pi_i + \pi_j + v \sqrt{\gamma_i \pi_i \pi_j}}$$

con el parámetro  $\gamma_i \geq 0$  representando la ventaja de jugar en casa que tienen cada elemento  $i$  del conjunto de  $m$  elementos a comparar.

### 5.2. Estimación de los parámetros de máxima verosimilitud para ventaja individual de jugar de local

Supondremos que observamos un conjunto de enfrentamientos por pares entre  $m$  individuos de un conjunto y deseamos estimar los parámetros  $\pi_1, \dots, \pi_m$ ,  $v$  y  $\gamma_1, \dots, \gamma_m$  usando para ello el método de máxima verosimilitud.

Sea  $n_{ij}$  el número de veces que el individuo  $i$  se ha enfrentado al individuo  $j$  en casa de  $i$ .

El número total de comparaciones entre individuos en el conjunto será

$$N = \sum_{i=1, i \neq j}^m \sum_{j=1, j \neq i}^m (n_{ij} + n_{ji})$$

Sea  $w_{ij}(1)$  el número de veces que el individuo  $i$  ha ganado a  $j$  jugando  $i$  como local

$w_{ij}(2)$  el número de veces que el individuo  $j$  ha ganado a  $i$  jugando  $i$  como local

$t_{ij}$  el número de veces que el individuo  $i$  ha empatado con  $j$  jugando  $i$  como local

$t_{ji}$  el número de veces que el individuo  $i$  ha empatado con  $j$  jugando  $j$  como local, entonces se tiene que

$$n_{ij} = w_{ij}(1) + w_{ij}(2) + t_{ij}$$

Entonces  $w_{iL}$  el número de veces que gana  $i$  jugando como local vendrá dado por

$$w_{iL} = \sum_{j=1, j \neq i}^m w_{ij}(1)$$

$w_{iV}$  el número de veces que gana  $i$  jugando como visitante vendrá dado por

$$w_{iV} = \sum_{j=1, j \neq i}^m w_{ji}(2)$$

$t_{iL}$  el número de veces que empata  $i$  jugando como local vendrá dado por

$$t_{iL} = \sum_{j=1, j \neq i}^m t_{ij}$$

$t_{iV}$  el número de veces que empata  $i$  jugando como visitante vendrá dado por

$$t_{iV} = \sum_{j=1, j \neq i}^m t_{ji}(2)$$

El número total de victorias locales en el conjunto de enfrentamientos será

$W_L = \sum_{i=1, i \neq j}^m \sum_{j=1, j \neq i}^m w_{ij}(1)$  y el número total de empates locales será

$T_L = \sum_{i=1, i \neq j}^m \sum_{j=1, j \neq i}^m t_{ij}$  (Davidson and Beaver, 1977).

Si suponemos independencias entre los enfrentamientos, la función de máxima verosimilitud basándose en la extensión del modelo de Bradley-Terry expuesta por Davidson y Beaver viene dada por:

$$\begin{aligned}
L(\pi_1, \dots, \pi_m; v; \gamma_1, \dots, \gamma_m) &= \prod_{\substack{i=1 \\ i \neq j}}^m \prod_{\substack{j=1 \\ i \neq j}}^m \left( \frac{\gamma_i \pi_i}{\gamma_i \pi_i + \pi_j + v \sqrt{\gamma_i \pi_i \pi_j}} \right)^{w_{ij}(1)} \left( \frac{\pi_j}{\gamma_i \pi_i + \pi_j + v \sqrt{\gamma_i \pi_i \pi_j}} \right)^{w_{ij}(2)} \left( \frac{v \sqrt{\gamma_i \pi_i \pi_j}}{\gamma_i \pi_i + \pi_j + v \sqrt{\gamma_i \pi_i \pi_j}} \right)^{t_{ij}} \\
&= \prod_{\substack{i=1 \\ i \neq j}}^m \prod_{\substack{j=1 \\ i \neq j}}^m \frac{(\gamma_i \pi_i)^{w_{ij}(1)}}{(\gamma_i \pi_i + \pi_j + v \sqrt{\gamma_i \pi_i \pi_j})^{w_{ij}(1)}} \frac{\pi_j^{w_{ij}(2)}}{(\gamma_i \pi_i + \pi_j + v \sqrt{\gamma_i \pi_i \pi_j})^{w_{ij}(2)}} \frac{(v \sqrt{\gamma_i \pi_i \pi_j})^{t_{ij}}}{(\gamma_i \pi_i + \pi_j + v \sqrt{\gamma_i \pi_i \pi_j})^{t_{ij}}} \\
&= \prod_{\substack{i=1 \\ i \neq j}}^m \prod_{\substack{j=1 \\ i \neq j}}^m (\gamma_i \pi_i)^{w_{ij}(1)} \pi_j^{w_{ij}(2)} (v \sqrt{\gamma_i \pi_i \pi_j})^{t_{ij}} \prod_{\substack{i=1 \\ i \neq j}}^m \prod_{\substack{j=1 \\ i \neq j}}^m \frac{1}{(\gamma_i \pi_i + \pi_j + v \sqrt{\gamma_i \pi_i \pi_j})^{w_{ij}(1)+w_{ij}(2)+t_{ij}}} \\
&= \prod_{\substack{i=1 \\ i \neq j}}^m \prod_{\substack{j=1 \\ i \neq j}}^m (\gamma_i \pi_i)^{w_{ij}(1)} \prod_{\substack{i=1 \\ i \neq j}}^m \prod_{\substack{j=1 \\ i \neq j}}^m \pi_j^{w_{ij}(2)} \prod_{\substack{i=1 \\ i \neq j}}^m \prod_{\substack{j=1 \\ i \neq j}}^m v^{t_{ij}} \prod_{\substack{i=1 \\ i \neq j}}^m \prod_{\substack{j=1 \\ i \neq j}}^m (\gamma_i \pi_i \pi_j)^{\frac{t_{ij}}{2}} \\
&\quad \prod_{\substack{i=1 \\ i \neq j}}^m \prod_{\substack{j=1 \\ i \neq j}}^m \frac{1}{(\gamma_i \pi_i + \pi_j + v \sqrt{\gamma_i \pi_i \pi_j})^{n_{ij}}} \\
&= \prod_{\substack{i=1 \\ i \neq j}}^m (\gamma_i \pi_i)^{\sum_{j=1, j \neq i}^m w_{ij}(1)} \prod_{\substack{i=1 \\ i \neq j}}^m \prod_{\substack{j=1 \\ i \neq j}}^m \pi_j^{w_{ij}(2)} v^{\sum_{i=1}^m \sum_{j=1, j \neq i}^m t_{ij}} \prod_{\substack{i=1 \\ i \neq j}}^m (\gamma_i \pi_i)^{\sum_{j=1, j \neq i}^m \frac{t_{ij}}{2}} \prod_{\substack{i=1 \\ i \neq j}}^m \prod_{\substack{j=1 \\ j \neq i}}^m \pi_j^{\frac{t_{ij}}{2}} \\
&\quad \prod_{\substack{i=1 \\ i \neq j}}^m \prod_{\substack{j=1 \\ i \neq j}}^m \frac{1}{(\gamma_i \pi_i + \pi_j + v \sqrt{\gamma_i \pi_i \pi_j})^{n_{ij}}} \\
&= v^{T_L} \prod_{\substack{i=1 \\ i \neq j}}^m (\gamma_i \pi_i)^{w_{iL}} \prod_{\substack{i=1 \\ i \neq j}}^m \prod_{\substack{j=1 \\ i \neq j}}^m \pi_j^{w_{ij}(2) + \frac{t_{ij}}{2}} \prod_{\substack{i=1 \\ i \neq j}}^m (\gamma_i \pi_i)^{\frac{t_{iL}}{2}} \prod_{\substack{i=1 \\ i \neq j}}^m \prod_{\substack{j=1 \\ i \neq j}}^m \frac{1}{(\gamma_i \pi_i + \pi_j + v \sqrt{\gamma_i \pi_i \pi_j})^{n_{ij}}} \\
&= \frac{v^{T_L} \prod_{\substack{i=1 \\ i \neq j}}^m (\gamma_i \pi_i)^{(w_{iL} + \frac{t_{iL}}{2})} \prod_{\substack{i=1 \\ i \neq j}}^m \prod_{\substack{j=1 \\ i \neq j}}^m \pi_j^{(w_{ij}(2) + \frac{t_{ij}}{2})}}{\prod_{\substack{i=1 \\ i \neq j}}^m \prod_{\substack{j=1 \\ i \neq j}}^m (\gamma_i \pi_i + \pi_j + v \sqrt{\gamma_i \pi_i \pi_j})^{n_{ij}}}
\end{aligned}$$

Para estimar por máxima verosimilitud, tomando logaritmo:

$$\begin{aligned}
 l(\pi_1, \dots, \pi_m; v; \gamma_1, \dots, \gamma_m) &= \ln(L(\pi_1, \dots, \pi_m; v; \gamma_1, \dots, \gamma_m)) \\
 &= \ln(v^{T_L} \prod_{\substack{i=1 \\ i \neq j}}^m (\gamma_i \pi_i)^{(w_{iL} + \frac{t_{iL}}{2})} \prod_{\substack{i=1 \\ i \neq j}}^m \prod_{\substack{j=1 \\ i \neq j}}^m \pi_j^{(w_{ij}(2) + \frac{t_{ij}}{2})} \prod_{\substack{j=1 \\ i \neq j}}^m (\gamma_i \pi_i + \pi_j + v\sqrt{\gamma_i \pi_i \pi_j})^{-n_{ij}}) \\
 &= \ln(v^{T_L}) + \sum_{\substack{i=1 \\ i \neq j}}^m \ln((\gamma_i \pi_i)^{(w_{iL} + \frac{t_{iL}}{2})}) + \sum_{\substack{i=1 \\ i \neq j}}^m \sum_{\substack{j=1 \\ i \neq j}}^m \ln(\pi_j^{(w_{ij}(2) + \frac{t_{ij}}{2})}) \\
 &\quad + \sum_{\substack{i=1 \\ i \neq j}}^m \sum_{\substack{j=1 \\ i \neq j}}^m \ln(\gamma_i \pi_i + \pi_j + v\sqrt{\gamma_i \pi_i \pi_j})^{-n_{ij}} \\
 &= T_L \ln(v) + \sum_{\substack{i=1 \\ i \neq j}}^m (w_{iL} + \frac{1}{2} t_{iL}) \ln(\gamma_i \pi_i) + \sum_{\substack{i=1 \\ i \neq j}}^m \sum_{\substack{j=1 \\ i \neq j}}^m (w_{ij}(2) + \frac{t_{ij}}{2}) \ln(\pi_j) \\
 &\quad - \sum_{\substack{i=1 \\ i \neq j}}^m \sum_{\substack{j=1 \\ i \neq j}}^m n_{ij} \ln(\gamma_i \pi_i + \pi_j + v\sqrt{\gamma_i \pi_i \pi_j}) \\
 &= T_L \ln(v) + \sum_{\substack{i=1 \\ i \neq j}}^m (w_{iL} + \frac{1}{2} t_{iL}) \ln(\pi_i) + \sum_{\substack{i=1 \\ i \neq j}}^m (w_{iL} + \frac{1}{2} t_{iL}) \ln(\gamma_i) \\
 &\quad + \sum_{\substack{i=1 \\ i \neq j}}^m \sum_{\substack{j=1 \\ i \neq j}}^m (w_{ij}(2) + \frac{t_{ij}}{2}) \ln(\pi_j) - \sum_{\substack{i=1 \\ i \neq j}}^m \sum_{\substack{j=1 \\ i \neq j}}^m n_{ij} \ln(\gamma_i \pi_i + \pi_j + v\sqrt{\gamma_i \pi_i \pi_j})
 \end{aligned}$$

Las ecuaciones de verosimilitud son entonces:

$$\begin{aligned}
 \frac{\partial}{\partial \pi_i} l(\pi_1, \dots, \pi_m; v; \gamma_1, \dots, \gamma_m) &= \frac{w_{iL} + \frac{1}{2} t_{iL}}{\pi_i} + \frac{\partial}{\partial \pi_i} \sum_{\substack{j=1 \\ i \neq j}}^m (w_{ij}(1) + \frac{1}{2} t_{ji}) \ln(\pi_i) \\
 &\quad - \frac{\partial}{\partial \pi_i} \sum_{\substack{j=1 \\ i \neq j}}^m (n_{ij} \ln(\gamma_i \pi_i + \pi_j + v\sqrt{\gamma_i \pi_i \pi_j}) + n_{ji} \ln(\gamma_i \pi_j + \pi_i + v\sqrt{\gamma_i \pi_i \pi_j})) \\
 &= \frac{w_{iL} + \frac{1}{2} t_{iL}}{\pi_i} + \frac{\partial}{\partial \pi_i} \sum_{\substack{j=1 \\ i \neq j}}^m (w_{ij}(1)) \ln(\pi_i) + \frac{\partial}{\partial \pi_i} \sum_{\substack{j=1 \\ i \neq j}}^m (\frac{1}{2} t_{ji}) \ln(\pi_i) \\
 &\quad - \frac{\partial}{\partial \pi_i} \sum_{\substack{j=1 \\ i \neq j}}^m (n_{ij} \ln(\gamma_i \pi_i + \pi_j + v\sqrt{\gamma_i \pi_i \pi_j})) - \frac{\partial}{\partial \pi_i} \sum_{\substack{j=1 \\ i \neq j}}^m (n_{ji} \ln(\gamma_i \pi_j + \pi_i + v\sqrt{\gamma_i \pi_i \pi_j})) \\
 &= \frac{w_{iL} + \frac{1}{2} t_{iL}}{\pi_i} + \frac{w_{iL} + \frac{1}{2} t_{iV}}{\pi_i} - \frac{1}{2} \left( \sum_{\substack{j=1 \\ i \neq j}}^m \frac{n_{ij} (2\gamma_i + v\sqrt{(\frac{\gamma_i \pi_j}{\pi_i})})}{\gamma_i \pi_i + \pi_j + v\sqrt{\gamma_i \pi_i \pi_j}} + \sum_{\substack{j=1 \\ i \neq j}}^m \frac{n_{ji} (2 + v\sqrt{(\frac{\gamma_i \pi_j}{\pi_i})})}{\gamma_i \pi_j + \pi_i + v\sqrt{\gamma_i \pi_i \pi_j}} \right) \\
 &= \frac{1}{2} \frac{4w_{iL} + t_{iL} + t_{iV}}{\pi_i} - \frac{1}{2} \left( \sum_{\substack{j=1 \\ i \neq j}}^m \frac{n_{ij} (2\gamma_i + v\sqrt{(\frac{\gamma_i \pi_j}{\pi_i})})}{\gamma_i \pi_i + \pi_j + v\sqrt{\gamma_i \pi_i \pi_j}} + \sum_{\substack{j=1 \\ i \neq j}}^m \frac{n_{ji} (2 + v\sqrt{(\frac{\gamma_i \pi_j}{\pi_i})})}{\gamma_i \pi_j + \pi_i + v\sqrt{\gamma_i \pi_i \pi_j}} \right) = 0
 \end{aligned}$$

$$\frac{\partial}{\partial v} l(\pi_1, \dots, \pi_m; v; \gamma_1, \dots, \gamma_m) = \frac{T_L}{v} - \sum_{\substack{i=1 \\ i \neq j}}^m \sum_{\substack{j=1 \\ i \neq j}}^m \frac{n_{ij} \sqrt{\gamma_i \pi_i \pi_j}}{\gamma_i \pi_i + \pi_j + v\sqrt{\gamma_i \pi_i \pi_j}} = 0$$



$$\begin{aligned} \frac{\partial}{\partial \gamma_i} l(\pi_1, \dots, \pi_m; v; \gamma_1, \dots, \gamma_m) &= \frac{w_{iL} + \frac{t_{iL}}{2}}{\gamma_i} - \sum_{\substack{j=1 \\ i \neq j}}^m \frac{n_{ij} \left( \pi_i + \frac{v \pi_i \pi_j}{2 \sqrt{\gamma_i \pi_i \pi_j}} \right)}{\gamma_i \pi_i + \pi_j + v \sqrt{\gamma_i \pi_i \pi_j}} \\ &= \frac{1}{2} \frac{(2w_{iL} + t_{iL})}{\gamma_i} - \frac{1}{2} \sum_{\substack{j=1 \\ i \neq j}}^m n_{ij} \frac{(2\pi_i + v \sqrt{\frac{\pi_i \pi_j}{\gamma_i}})}{\gamma_i \pi_i + \pi_j + v \sqrt{\gamma_i \pi_i \pi_j}} = 0 \end{aligned}$$

### 5.3. Modelo Bradley-Terry con ventaja individual de jugar de local como un modelo lineal generalizado

Un modelo lineal generalizado consta de 3 componentes: Una funcion de distribución, un predictor lineal y una funcion de enlace.

La ecuación definitiva del modelo lineal generalizado es:

$$g(\mu) = X\theta$$

Con  $\mu$  vector de medias, un vector de parámetros desconocidos  $\theta$ , la matriz del modelo  $X$ , y la función denominada función de enlace  $g$ .

Representaremos el modelo Bradley-Terry extendido por Davidson y Beaver incluyendo los empates y la ventaja individual de jugar como local como un modelo lineal generalizado, y estimaremos así los parámetros desconocidos  $\theta$  a través de métodos conocidos para los modelos lineales generalizados.

Antes de identificar cómo será el vector de parámetros  $\theta$  y la matriz del modelo  $X$ , es necesario discutir cómo ajustaremos la distribución multinomial que tenemos a un modelo lineal generalizado, ya que la distribución multinomial no se modela como un modelo lineal generalizado directamente. De acuerdo al Th. 13.4-1 que desarrollaron Bishop, Fienberg y Holland (Bishop et al., 1975), una distribución multinomial puede modelarse como un vector de variables aleatorias de Poisson independientes, introduciendo restricciones adecuadas en el modelo (Critchlow and Flinger, 1991).

Podríamos ver que cada enfrentamiento entre el individuo  $i$  y el individuo  $j$  sigue una distribución multinomial, en este caso trinomial con probabilidades,

P(el individuo  $i$  le gane al individuo  $j$  jugando  $i$  como local)

$$p_{ij1} = \frac{\gamma_i \pi_i}{\gamma_i \pi_i + \pi_j + v \sqrt{\gamma_i \pi_i \pi_j}}$$

P(el individuo  $j$  le gane al individuo  $i$  jugando  $i$  como local)

$$p_{ij2} = \frac{\pi_j}{\gamma_i \pi_i + \pi_j + v \sqrt{\gamma_i \pi_i \pi_j}}$$

P(el individuo  $i$  empate con el individuo  $j$  jugando  $i$  como local)

$$p_{ij3} = \frac{v \sqrt{\gamma_i \pi_i \pi_j}}{\gamma_i \pi_i + \pi_j + v \sqrt{\gamma_i \pi_i \pi_j}}$$

$$Y_{ij} = [Y_{ij1}|Y_{ij2}|Y_{ij3}] \sim Mult_3(n_{ij}, p_{ij} = [p_{ij1}|p_{ij2}|p_{ij3}])$$

Multiplicando estas probabilidades por el número de comparaciones del par (i,j) siendo local i,  $n_{ij}$ , , obtendremos el valor esperado del número de veces que gana i a j siendo local i, el número de veces que gana j a i siendo local i y el número de veces que empatan siendo local i. Por lo que  $\mu = E[Y_{ijk}] = n_{ij}p_{ijk}$

La transformación Multinomial-Poisson (Baker, 1994) simplifica la estimación de los parámetros del modelo por máxima verosimilitud.

Tomaremos cada  $Y_{ijk}$  como el número de veces que al comparar el par (i,j) siendo local i ocurre k, con k=1 el número de veces que gana el individuo i siendo local i, con k=2 el número de veces que gana el individuo j siendo local i, y con k=3 el número de veces que empatan el individuo i con el individuo j siendo local i. Entonces  $Y_{ijk} \sim Poi(n_{ij}p_{ijk})$  con  $E[Y_{ijk}] = n_{ij}p_{ijk} = \mu_{ijk}$ .

Y sabemos que podemos expresar una distribución de Poisson como un modelo lineal generalizado tomando como función de enlace el logaritmo neperiano por lo que resulta:

$$\ln(\mu) = X\theta$$

con  $\mu = (\mu_{121}, \mu_{122}, \mu_{123}, \dots, \mu_{t(t-1)1}, \mu_{t(t-1)2}, \mu_{t(t-1)3})$ .

Tenemos que  $\mu_{ijk} = n_{ij}p_{ijk}$ , sustituyendo:

$$\begin{aligned} \ln(\mu_{ij1}) &= \ln(n_{ij}p_{ij1}) \\ &= \ln(n_{ij}) + \ln(p_{ij1}) \\ &= \ln(n_{ij}) + \ln\left(\frac{\gamma_i\pi_i}{\gamma_i\pi_i + \pi_j + v\sqrt{\gamma_i\pi_i\pi_j}}\right) \\ &= \ln(n_{ij}) + \ln(\gamma_i\pi_i) - \ln(\gamma_i\pi_i + \pi_j + v\sqrt{\gamma_i\pi_i\pi_j}) \\ &= \ln(\gamma_i) + \ln(\pi_i) + \ln\left(\frac{n_{ij}}{\gamma_i\pi_i + \pi_j + v\sqrt{\gamma_i\pi_i\pi_j}}\right) \end{aligned}$$

$$\begin{aligned} \ln(\mu_{ij2}) &= \ln(n_{ij}p_{ij2}) \\ &= \ln(n_{ij}) + \ln(p_{ij2}) \\ &= \ln(n_{ij}) + \ln\left(\frac{\pi_j}{\gamma_i\pi_i + \pi_j + v\sqrt{\gamma_i\pi_i\pi_j}}\right) \\ &= \ln(n_{ij}) + \ln(\pi_j) - \ln(\gamma_i\pi_i + \pi_j + v\sqrt{\gamma_i\pi_i\pi_j}) \\ &= \ln(\pi_j) + \ln\left(\frac{n_{ij}}{\gamma_i\pi_i + \pi_j + v\sqrt{\gamma_i\pi_i\pi_j}}\right) \end{aligned}$$

$$\begin{aligned}
 \ln(\mu_{ij3}) &= \ln(n_{ij}p_{ij3}) \\
 &= \ln(n_{ij}) + \ln(p_{ij3}) \\
 &= \ln(n_{ij}) + \ln\left(\frac{v\sqrt{\gamma_i\pi_i\pi_j}}{\gamma_i\pi_i + \pi_j + v\sqrt{\gamma_i\pi_i\pi_j}}\right) \\
 &= \ln(n_{ij}) + \ln(v\sqrt{\gamma_i\pi_i\pi_j}) - \ln(\gamma_i\pi_i + \pi_j + v\sqrt{\gamma_i\pi_i\pi_j}) \\
 &= \ln(v\sqrt{\gamma_i\pi_i\pi_j}) + \ln\left(\frac{n_{ij}}{\gamma_i\pi_i + \pi_j + v\sqrt{\gamma_i\pi_i\pi_j}}\right) \\
 &= \ln(v) + \ln(\sqrt{\gamma_i\pi_i\pi_j}) + \ln\left(\frac{n_{ij}}{\gamma_i\pi_i + \pi_j + v\sqrt{\gamma_i\pi_i\pi_j}}\right) \\
 &= \ln(v) + \frac{1}{2}\ln(\gamma_i\pi_i\pi_j) + \ln\left(\frac{n_{ij}}{\gamma_i\pi_i + \pi_j + v\sqrt{\gamma_i\pi_i\pi_j}}\right) \\
 &= \ln(v) + \frac{1}{2}(\ln(\gamma_i) + \ln(\pi_i) + \ln(\pi_j)) + \ln\left(\frac{n_{ij}}{\gamma_i\pi_i + \pi_j + v\sqrt{\gamma_i\pi_i\pi_j}}\right)
 \end{aligned}$$

donde  $\beta_i = \ln(\pi_i)$ ,  $\beta = \ln(v)$ ,  $\alpha_{ij} = \ln\left(\frac{n_{ij}}{\gamma_i\pi_i + \pi_j + v\sqrt{\gamma_i\pi_i\pi_j}}\right)$ ,  $\alpha_{ij} \neq \alpha_{ji}$  que será una constante de normalización y  $\psi_i = \ln(\gamma_i)$  Tendremos  $t(t-1)$   $\alpha_{ij}$  incluidas en el vector de parámetros  $\theta$ , lo que hará que la matriz del modelo aumente  $t(t-1)$  columnas. Estas columnas introducirán una variable de clasificación, clasificando nuestras observaciones de Poisson con la distribución multinomial a la que corresponden. Por lo que la columna asociada a la constante de normalización  $\alpha_{ij}$  tendrá un 1 en las filas correspondientes a  $Y_{ij1}, Y_{ij2}, Y_{ij3}$  y 0 en las demás filas.

Tenemos que

$$\ln(\mu_{ij1}) = \beta_i + \psi_i + \alpha_{ij} \quad (5.1)$$

$$\ln(\mu_{ij2}) = \beta_j + \alpha_{ij} \quad (5.2)$$

$$\ln(\mu_{ij3}) = \beta + \frac{1}{2}(\beta_i + \beta_j + \psi_i) + \alpha_{ij} \quad (5.3)$$

donde  $\beta_i = \ln(\pi_i)$ ,  $\beta = \ln(v)$ ,  $\alpha_{ij} = \ln\left(\frac{n_{ij}}{\gamma_i\pi_i + \pi_j + v\sqrt{\gamma_i\pi_i\pi_j}}\right)$  y  $\psi_i = \ln(\gamma_i)$  serán los parámetros desconocidos para mi modelo de regresión de Poisson.

Tenemos entonces un modelo de regresión de Poisson con:

- Componente aleatoria: Para  $i \neq j$ , siendo  $Y_{ijk}$  el número de veces que el elemento  $i$  gana ( $k=1$ ) al elemento  $j$  siendo local  $i$ , el número de veces que el elemento  $j$  gana ( $k=2$ ) al elemento  $i$  siendo local  $i$  o el número de veces que el elemento  $i$  empata ( $k=3$ ) con el elemento  $j$  siendo local  $i$  en las  $n_{ij}$  comparaciones entre  $(i,j)$  siendo local  $i$ . Todas las observaciones para las  $t$  componentes a comparar nos da el vector  $Y=(Y_{121}, Y_{122}, Y_{123}, \dots, Y_{t(t-1)1}, Y_{t(t-1)2}, Y_{t(t-1)3})$ , con un tamaño  $N = 3t(t-1)$ , cada  $Y_{ijk} \sim Poi(n_{ij}p_{ijk})$ , con  $n_{ij}$  el número de veces que se compara el elemento  $i$  con el  $j$  siendo  $i$  local y  $p_{ijk}$  la probabilidad de ganar ( $k=1$ ) el elemento  $i$  al elemento  $j$  siendo  $i$  local, la probabilidad de ganar ( $k=2$ ) el elemento  $j$  al elemento  $i$  siendo  $i$  local o la probabilidad de empatar ( $k=3$ ) el elemento  $i$  con el elemento  $j$  siendo  $i$  local. Tenemos que  $E[Y_{ijk}] = n_{ij}p_{ijk}$ . El vector de las probabilidades de  $Y$  será el vector  $p_{ijk} = (p_{121}, p_{122}, p_{123}, \dots, p_{t(t-1)1}, p_{t(t-1)2}, p_{t(t-1)3})$  con  $p_{ijk} = E\left[\frac{Y_{ijk}}{n_{ij}}\right]$ .
- Predictor lineal. El predictor lineal  $X\theta$  está formado por un vector de parámetros desconocidos  $\theta$  y la matriz del modelo  $X$ . El vector de parámetros desconocidos  $\theta$  será el formado por los  $\beta_i, \beta, \psi, \alpha_{ij}$ , con  $\theta = (\beta_1, \dots, \beta_t, \psi_1, \dots, \psi_t, \beta, \alpha_{12}, \dots, \alpha_{t(t-1)})$ . Y la matriz del modelo tendrá dimensión  $N = 3t(t-1)$  filas x  $(2t+1)+t(t-1)$  columnas.

Sea  $X$  la matriz del modelo, formada por  $X = [X_1|X_2]$ .

La matriz  $X_1$  tendrá  $N=3t(t-1)$  filas x  $(2t+1)$  columnas.

Como sabemos por (5.1), (5.2), (5.3) la  $m$ -ésima entrada en la  $(i,j,k)$ -ésima fila de la matriz  $X_1$  si  $k=1$  será 1 si  $m=i, m=t+i$  y 0 en caso contrario, si  $k=2$  será 1 si  $m=j$  y 0 en caso contrario, y si  $k=3$  será  $\frac{1}{2}$  si  $m=i, m=j$  o  $m=t+i$  y 1 si  $m=2t+1$  y 0 en caso contrario.

La matriz  $X_2$  tendrá  $N=3t(t-1)$  filas x  $t(t-1)$  columnas. Cada columna se asocia a la constante de normalización  $\alpha_{ij}$ , por lo que tendrá un 1 en las filas correspondientes a  $Y_{ij1}, Y_{ij2}, Y_{ij3}$  y 0 en las demás filas.

$$X = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & 0 & 1 & 0 & \cdots & 0 & 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 1 & 0 & \cdots & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & \cdots & 0 & 0 & \frac{1}{2} & 0 & \cdots & 0 & 0 & 1 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 & 0 & 0 & \cdots & 0 & 1 & 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & \frac{1}{2} & \frac{1}{2} & 0 & 0 & \cdots & 0 & \frac{1}{2} & 1 & 0 & 0 & \cdots & 1 \end{pmatrix} \quad (5.4)$$

Tenemos

$$X\theta = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & 0 & 1 & 0 & \cdots & 0 & 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 1 & 0 & \cdots & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & \cdots & 0 & 0 & \frac{1}{2} & 0 & \cdots & 0 & 0 & 1 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 & 0 & 0 & \cdots & 0 & 1 & 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & \frac{1}{2} & \frac{1}{2} & 0 & 0 & \cdots & 0 & \frac{1}{2} & 1 & 0 & 0 & \cdots & 1 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \vdots \\ \beta_{t-1} \\ \beta_t \\ \psi_1 \\ \psi_2 \\ \vdots \\ \psi_{t-1} \\ \psi_t \\ \beta \\ \alpha_{12} \\ \alpha_{13} \\ \vdots \\ \alpha_{t(t-1)} \end{pmatrix} \quad (5.5)$$

- Función de enlace: La función de enlace para representar el modelo Bradley-Terry con empates y ventaja de local como un modelo de regresión de Poisson es la función ln.

Tenemos entonces :

$$\ln \begin{pmatrix} \mu_{121} \\ \mu_{122} \\ \mu_{123} \\ \vdots \\ \mu_{t(t-1)1} \\ \mu_{t(t-1)2} \\ \mu_{t(t-1)3} \end{pmatrix} = X\theta = \begin{pmatrix} \beta_1 + \psi_1 + \alpha_{12} \\ \beta_2 + \alpha_{12} \\ \beta + \frac{1}{2}(\beta_1 + \beta_2 + \psi_1) + \alpha_{12} \\ \vdots \\ \beta_t + \psi_t + \alpha_{t(t-1)} \\ \beta_{t-1} + \alpha_{t(t-1)} \\ \beta + \frac{1}{2}(\beta_{t-1} + \beta_t + \psi_t) + \alpha_{t(t-1)} \end{pmatrix} \quad (5.6)$$

(Critchlow and Fligner, 1991)

# Capítulo 6

## Aplicación del Modelo Bradley Terry a datos de fútbol

### 6.1. Clasificación mediante el modelo de Davidson

La Primera División de España o La Liga es la máxima categoría del fútbol profesional en España. Cada temporada suele empezar en Agosto y terminar en Mayo, la juegan 20 equipos que se enfrentan todos contra todos en 2 partidos con cada equipo, uno como local y otro como visitante por lo que cada equipo juega 38 partidos en la temporada, con un total de 380 partidos entre todos. Los equipos se ordenan en la clasificación recibiendo una puntuación de 3 puntos por cada partido ganado, de 1 punto por cada partido empatado y de 0 puntos por cada partido perdido. Durante los últimos años el Real Madrid, el Atlético de Madrid y el F.C. Barcelona suelen ser los aspirantes a ganar cada año La Liga.

En la temporada 2020/2021 de La Liga, el Atlético de Madrid fue campeón en la jornada 38, en la última jornada de una liga que siempre será recordada porque fue vivida sin público a causa de la Covid-19. Los cuatro primeros clasificados de La Liga jugarán la siguiente temporada Champions League, los clasificados en quinta y sexta posición jugarán Europa League y el séptimo clasificado la Conference League.

Primero vamos a analizar si el ranking final que nos ofrece el modelo de Davidson se parece al modelo de clasificación usado en La Liga y después vamos a analizar si el modelo sirve para predecir el ranking final de La Liga en diferentes momentos de la temporada. En el Apéndice A encontraremos el código utilizado para obtener los siguientes resultados.

Para ello definiremos la matriz de mi modelo y estimaremos los parámetros como un modelo lineal generalizado como hemos descrito anteriormente usando para ello la función glm de RStudio.

Empezaremos definiendo la matriz de mi modelo que cómo hemos detallado antes es de la forma

$$X = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & 0 & 0 & 1 & 0 & \cdots & 0 \\ \frac{1}{2} & \frac{1}{2} & \cdots & 0 & 0 & 1 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 & 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & \cdots & 0 & 1 & 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & \cdots & \frac{1}{2} & \frac{1}{2} & 1 & 0 & 0 & \cdots & 1 \end{pmatrix} \quad (6.1)$$

A continuación proceso los datos para obtener las observaciones ordenadas de forma que pueda aplicarle el modelo de Davidson. Una vez tenga los datos preparados y la matriz de mi modelo definida creamos el modelo mediante la función `glm`, como una distribución de Poisson con función de enlace log. Los coeficientes de mi modelo serán  $\beta_i = \ln(\pi_i)$ ,  $\beta = \ln(v)$ ,  $\alpha_{ij} = \ln\left(\frac{n_{ij}}{\pi_i + \pi_j + v\sqrt{\pi_i\pi_j}}\right)$  serán los parámetros desconocidos para mi modelo de regresión de Poisson. Para hallar los parámetros que necesito para clasificar a los equipos según su habilidad, los parámetros  $\pi_i$ , tengo que calcular  $e^{\beta_i}$  y obtengo así los parámetros que necesito por lo que ya puedo obtener la clasificación de los equipos de La Liga según el modelo de Davidson.

El valor estimado de parámetro común de empate  $v$  es de 0.9818693. El residual deviance calculado por `glm` es de 428.47 con 360 grados de libertad. En la Tabla 6.1 a la izquierda nos muestra los coeficientes  $\pi_i$  que nos miden la habilidad de cada equipo mediante el modelo de Davidson ordenado en orden decreciente, por lo que nos da el ranking ordenado de los equipos del más al menos fuerte según este modelo, y a la derecha podemos ver la clasificación final real de La Liga en esa temporada. Nuestro objetivo es que el modelo de Bradley-Terry contemplado los empates sea capaz de comparar a los equipos y clasificarlos en un ranking que sea lo más parecido posible al ranking real de La Liga. Contemplando ambas tablas podemos observar que los equipos se clasifican correctamente, con las únicas excepciones de Cádiz, Granada, Osasuna y Ath de Bilbao. El motivo de esta discrepancia es fácil de explicar, en el modelo Bradley-Terry tres de estos equipos obtuvieron el mismo parámetro de habilidad  $\pi_i$  (Cádiz, Granada, Osasuna)=(0.0183473,0.0183473,0.0183473) y el Ath de Bilbao (0.216661), pero el Granada ha ganado un número mayor de partidos, y en el modelo de clasificación de La Liga da más valor a los partidos ganados por ello el Granada obtuvo 46 al igual que el Ath de Bilbao que tuvo menos partidos perdidos, y el Osasuna y el Cádiz 44 puntos con igual número de partidos ganados, perdidos o empatados. Realmente solucionando esta diferencia entre Granada y Ath de Bilbao resultaría la clasificación final real de La Liga. Por lo general podemos concluir que el ranking dado mediante el Modelo de Davidson no presenta grandes diferencias con la clasificación real de La Liga.

Tabla 6.1: Clasificación de La Liga Temporada 2020-2021 mediante el modelo Bradley-Terry VS Clasificación Real

Pos	Equipos	Coeficiente	Equipos	V	E	D	Pts
1	Ath Madrid	0.2228735	Ath Madrid	26	8	4	86
2	Real Madrid	0.1977395	Real Madrid	25	9	4	84
3	Barcelona	0.1273987	Barcelona	24	7	7	79
4	Sevilla	0.1041239	Sevilla	24	5	9	77
5	Sociedad	0.0500579	Sociedad	17	11	10	62
6	Betis	0.0459344	Betis	17	10	11	61
7	Villarreal	0.0421839	Villarreal	15	13	10	58
8	Celta	0.0301734	Celta	14	11	13	53
9	Ath Bilbao	0.0216661	Granada	13	7	18	46
10	Cadiz	0.0183473	Ath Bilbao	11	13	14	46
11	Granada	0.0183473	Osasuna	11	11	16	44
12	Osasuna	0.0183473	Cadiz	11	11	16	44
13	Valencia	0.0183473	Valencia	10	13	15	43
14	Levante	0.0168745	Levante	9	14	15	41
15	Getafe	0.0130805	Getafe	9	11	18	38
16	Alaves	0.0130805	Alaves	9	11	18	38
17	Elche	0.0119960	Elche	8	12	18	36
18	Huesca	0.0109899	Huesca	7	13	18	34
19	Valladolid	0.0100557	Valladolid	5	16	17	31
20	Eibar	0.0083822	Eibar	6	12	20	30

Resulta interesante analizar si nuestro modelo predice bien el ranking final de La Liga en ciertos periodos de tiempo durante la temporada. Para ello hemos dividido los resultados de los partidos en 4 periodos de tiempo, un primer periodo con los resultados de los partidos desde el inicio de temporada hasta la Jornada 10, un segundo periodo con los resultados de los partidos desde el inicio de temporada hasta la Jornada 20, un tercer periodo con los resultados de los partidos desde el inicio de temporada hasta la Jornada 30 y un cuarto periodo con los resultados de toda la temporada al completo. Uno de los periodos más interesantes para analizar es el segundo periodo ya que en la Jornada 20 ya se han enfrentado todos los equipos entre sí.

En la Tabla 6.2. podemos observar como el ranking predicho en la Jornada 10 dista mucho del ranking real de La Liga, conforme vayamos añadiendo jornadas al modelo la predicción del ranking se acerca más al ranking real.

Para cuantificar la deficiencia del modelo de Davidson para predecir la clasificación final de La Liga en diferentes periodos de la temporada vamos a definir dos errores a analizar. Sea  $e_r$  el porcentaje de equipos que son mal clasificados en el periodo estudiado con respecto a la clasificación real de La Liga. Si tenemos  $t$  equipos,  $I_i=1$  si el equipo  $i$  es clasificado equivocadamente y 0 en caso contrario:

$$e_r = \frac{1}{t} \sum_{i=1}^t I_i$$

Vamos a cuantificar también con  $e_m$  el número de posiciones que varían de media entre la posición de la clasificación en el periodo estudiado con respecto a la posición en la clasificación real de La Liga. Sea  $r_i$  la posición del equipo  $i$  en la clasificación final de La Liga y  $r'_i$  la posición del equipo  $i$  en la clasificación en el periodo estudiado. Se define entonces  $e_m$  como:

$$e_m = \frac{1}{t} \sum_{i=1}^t |r_i - r'_i|$$

Una vez calculados los dos tipos de errores definidos anteriormente para nuestro modelo obtenemos la Tabla 6.3.

Tabla 6.3: Errores calculados en los distintos periodos de la temporada con respecto a la clasificación final

Periodo de la Temporada	er %	em
J10	95	4.0
J20	80	2.2
J30	55	1.1
J38	20	0.3

Hemos obtenido que la clasificación obtenida con los resultados en la jornada 10 J-10 con respecto a la clasificación final de La Liga clasifica mal a los equipos en el 95 % de las ocasiones, con una diferencia media entre la posición estimada y la real de 4 posiciones. Para la jornada 20 J-20 hemos obtenido que la clasificación obtenida en esta jornada con respecto a la clasificación final de La Liga clasifica mal a los equipos en el 80 % de las ocasiones, con una diferencia media entre la posición estimada y la real de 2.2 posiciones. Para la jornada 30 J-30 hemos obtenido que la clasificación obtenida en esta jornada con respecto a la clasificación final de La Liga clasifica mal a los equipos en el 55 % de las ocasiones, con una diferencia media entre la posición estimada y la real de 1.1 posiciones y finalmente la clasificación final obtenida en la jornada 38 mediante el modelo Bradley-Terry con respecto a la clasificación final de La Liga clasifica mal a los equipos en el 20 % de las ocasiones, con una diferencia media entre la posición estimada y la real de 0.3 posiciones.

Por lo que podemos concluir que el modelo de Davidson no predice bien la clasificación final en los distintos periodos, conforme van pasando las jornadas la predicción es más fiable. Por lo que la habilidad de cada equipo obtenida en un periodo determinado sólo determina la habilidad de ese equipo en ese periodo y no sería fiable tomarla como habilidad final del equipo.



Tabla 6.2: Clasificación de La Liga Temporada 2020-2021 mediante el modelo Bradley-Terry en diferentes jornadas

Pos-J10	Equipos	Coeficiente	Pos-J20	Equipos	Coeficiente
1	Ath Madrid	0.3519711	1	Ath Madrid	0.4796196
2	Sociedad	0.1677548	2	Real Madrid	0.1198512
3	Villarreal	0.1056620	3	Barcelona	0.0787276
4	Real Madrid	0.0994703	4	Sevilla	0.0667108
5	Barcelona	0.0492536	5	Villarreal	0.0529787
6	Granada	0.0432894	6	Sociedad	0.0329405
7	Cadiz	0.0354273	7	Granada	0.0193339
8	Sevilla	0.0248502	8	Betis	0.0175745
9	Getafe	0.0204198	9	Levante	0.0165467
10	Valencia	0.0142088	10	Cadiz	0.0154078
11	Betis	0.0135688	11	Celta	0.0140405
12	Elche	0.0108775	12	Getafe	0.0120092
13	Huesca	0.0104783	13	Ath Bilbao	0.0120092
14	Alaves	0.0092049	14	Valencia	0.0116287
15	Osasuna	0.0089326	15	Valladolid	0.0103442
16	Ath Bilbao	0.0080325	16	Eibar	0.0102546
17	Eibar	0.0075716	17	Osasuna	0.0088071
18	Valladolid	0.0072756	18	Alaves	0.0079632
19	Levante	0.0068703	19	Elche	0.0078626
20	Celta	0.0048805	20	Huesca	0.0053897

Pos-J30	Equipos	Coeficiente	Pos	Equipos	Coeficiente
1	Ath Madrid	0.2247314	1	Ath Madrid	0.2228735
2	Real Madrid	0.1839008	2	Real Madrid	0.1977395
3	Barcelona	0.1590013	3	Barcelona	0.1273987
4	Sevilla	0.1026927	4	Sevilla	0.1041239
5	Sociedad	0.0479182	5	Sociedad	0.0500579
6	Villarreal	0.0432193	6	Betis	0.0459344
7	Betis	0.0400053	7	Villarreal	0.0421839
8	Levante	0.0240066	8	Celta	0.0301734
9	Celta	0.0204981	9	Ath Bilbao	0.0216661
10	Granada	0.0204923	10	Cadiz	0.0183473
11	Ath Bilbao	0.0201418	11	Granada	0.0183473
12	Cadiz	0.0186384	12	Osasuna	0.0183473
13	Valencia	0.0181983	13	Valencia	0.0183473
14	Osasuna	0.0163138	14	Levante	0.0168745
15	Getafe	0.0121225	15	Getafe	0.0130805
16	Huesca	0.0112210	16	Alaves	0.0130805
17	Valladolid	0.0106698	17	Elche	0.0119960
18	Elche	0.0102752	18	Huesca	0.0109899
19	Alaves	0.0084422	19	Valladolid	0.0100557
20	Eibar	0.0075109	20	Eibar	0.0083822

## 6.2. Análisis de la ventaja de jugar de local común e individual para los distintos equipos de La Liga

Analizaremos la ventaja común e individual de cada equipo en Primera División a lo largo de las últimas 8 temporadas, desde la 2014-2015 a la actual Liga 2021-2022. En el Apéndice B encontraremos el código utilizado para obtener los siguientes resultados.

En la Tabla 6.4. se muestran los valores estimados de  $\gamma$  y  $\gamma_i$  los cuáles nos indican la ventaja común e individual de jugar en casa de los equipos que han jugado en Primera División en alguna de las 8 temporadas anteriores. Los valores de  $\gamma$  estimados por el Modelo de Bradley-Terry teniendo en cuenta los empates y una ventaja común de jugar en casa de todos los equipos vienen dados en la última fila de la Tabla 6.4. También se ha calculado la media de la ventaja individual de todos los equipos en cada temporada que vendrá dada por la penúltima fila de la Tabla 6.4. En la última columna de la Tabla 6.4. se ha calculado la media de ventaja individual de cada equipo durante los años que ha jugado en Primera en las 8 últimas temporadas. Los equipos serán ordenados en la tabla de forma decreciente en relación con esta columna.

Analicemos el significado de los valores más significativos de la Tabla 6.4. en la vida real. En la temporada 2014/15 el F.C.Barcelona obtuvo una constante de ventaja de jugar en casa de 0.83, esto se debe a que en casa consiguió 49 puntos y fuera de casa 45 por lo que prácticamente no se aprecia ventaja de jugar de local por el contrario el Valencia C.F. obtuvo una constante de ventaja de jugar en casa de 9.23 ya que en casa consiguió 48 puntos y jugando como visitante sólo 29, por lo que claramente esa temporada tuvo una buena ventaja de jugar en casa.

En la temporada 2015/16 destaca el Sevilla F.C. con una constante de ventaja de jugar en casa de 26.12 ya que consiguió 43 puntos en casa frente a los 9 que consiguió como visitante, pero no es el equipo que más ventaja individual ha tenido ya que Las Palmas en la temporada 2016/17 obtuvo una constante de ventaja de jugar en casa de 36.62 ya que consiguió 33 puntos jugando como local pero sólo 6 como visitante.

Un dato curioso y que refleja la importancia que tienen las aficiones de cada equipo es que en la Temporada 2020/21, la única temporada que se ha jugado sin público a lo largo de la historia de La Liga, la ventaja común de jugar en casa es la más pequeña de todas las temporadas, al igual que la media de las ventajas individuales de cada equipo, por lo que el apoyo del público juega un papel importante en el valor de este parámetro de ventaja local.

A lo largo de la temporada actual 2021/22 el Cádiz es el equipo con menor ventaja individual 0.4 ya que en casa ha conseguido 18 puntos y fuera de ella 21 por lo que jugar en casa no tiene un efecto de ventaja, el Español es el equipo que mayor ventaja de jugar en casa ha obtenido con 14.25 ya que jugando como local ha conseguido 33 puntos frente a los sólo 9 como visitante por lo que el efecto de jugar en casa es más que evidente.

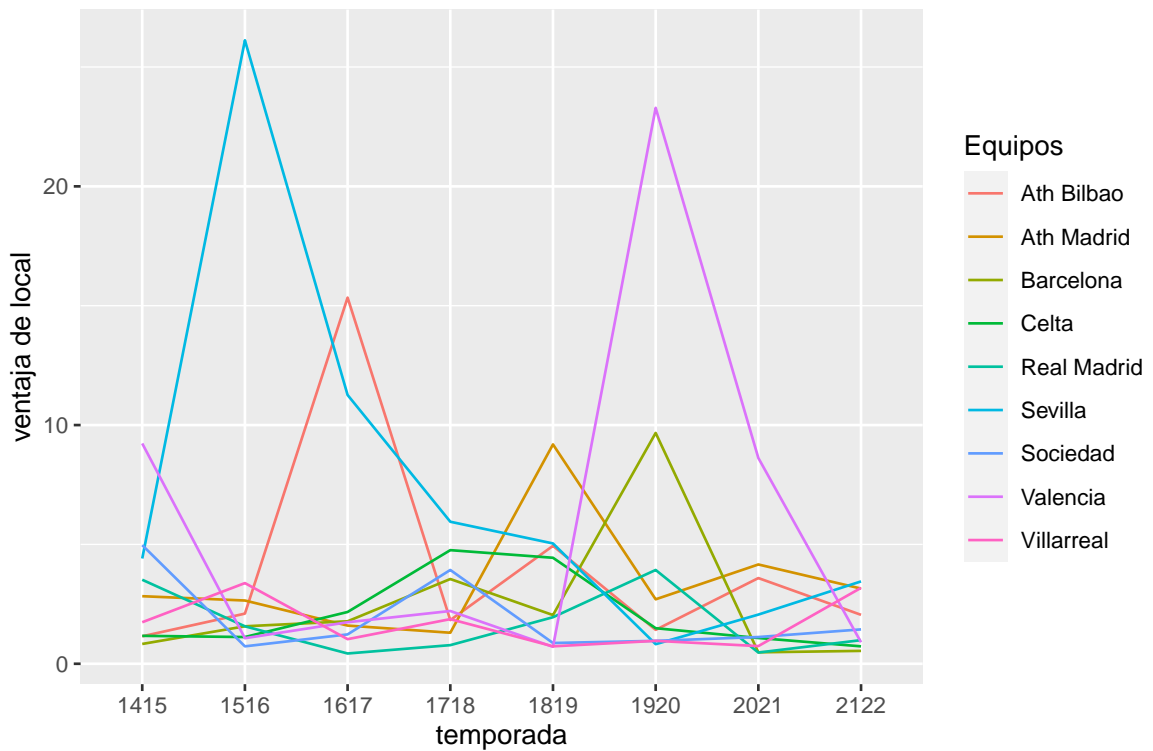
Tabla 6.4: Ventaja de jugar de local común e individual en las 8 últimas temporadas

Equipos	1415	1516	1617	1718	1819	1920	2021	2122	avg
Las Palmas	-	2.92	36.62	1.03	-	-	-	-	13.52
Mallorca	-	-	-	-	-	13.33	-	4.89	9.11
Sevilla	4.41	26.12	11.26	5.95	5.04	0.82	2.06	3.45	7.39
Valencia	9.23	1.07	1.74	2.21	0.71	23.29	8.62	0.9	5.97
Espanol	2.65	8.79	0.51	1.81	4.02	0.43	-	14.25	4.64
Ath Bilbao	1.15	2.1	15.34	1.83	4.94	1.42	3.59	2.05	4.05
Alaves	-	-	0.77	1.52	1.84	4.48	2.5	11.65	3.79
Ath Madrid	2.83	2.65	1.6	1.3	9.19	2.7	4.16	3.15	3.45
Levante	3.05	10.43	-	1.24	1.3	3.95	1.96	2.1	3.43
Vallecano	0.81	3.8	-	-	2.98	-	-	4.67	3.06
Malaga	2.65	1.92	3.31	3.98	-	-	-	-	2.96
Leganes	-	-	1.18	6.34	3.4	0.66	-	-	2.90
Granada	3.05	1.09	4.61	-	-	1.42	5.56	1.08	2.80
Barcelona	0.83	1.57	1.78	3.55	2.04	9.67	0.48	0.54	2.56
Getafe	2.53	4.05	-	1.04	1.54	0.96	2.5	4.67	2.47
Elche	0.56	-	-	-	-	-	3.17	3.03	2.25
Betis	-	0.6	2.83	1.53	1.53	6.59	2.07	0.35	2.21
Huesca	-	-	-	-	1.78	-	2.6	-	2.19
Celta	1.17	1.12	2.17	4.76	4.44	1.49	1.08	0.73	2.12
Sociedad	4.97	0.73	1.23	3.93	0.87	0.96	1.12	1.44	1.91
Eibar	0.58	2.92	1.76	0.86	4.97	1.8	0.37	-	1.89
Real Madrid	3.52	1.57	0.43	0.78	1.95	3.93	0.47	0.99	1.71
Villarreal	1.74	3.38	1.03	1.87	0.73	0.96	0.74	3.18	1.70
La Coruna	1.3	0.28	3.65	1.43	-	-	-	-	1.67
Sp Gijon	-	1.63	0.97	-	-	-	-	-	1.30
Osasuna	-	-	1.11	-	-	0.79	1.59	0.41	0.98
Valladolid	-	-	-	-	0.42	1.76	0.58	-	0.92
Almeria	0.91	-	-	-	-	-	-	-	0.91
Cordoba	0.46	-	-	-	-	-	-	-	0.46
Girona	-	-	-	0.6	0.23	-	-	-	0.42
Cadiz	-	-	-	-	-	-	0.33	0.4	0.36
avgi	2.42	3.94	4.69	2.38	2.7	4.07	2.28	3.2	-
gamma	1.71	2.07	2.09	1.82	1.81	2	1.55	1.77	-

Vamos a representar la ventaja de jugar en casa de los equipos que han estado en las 8 temporadas en La Liga.

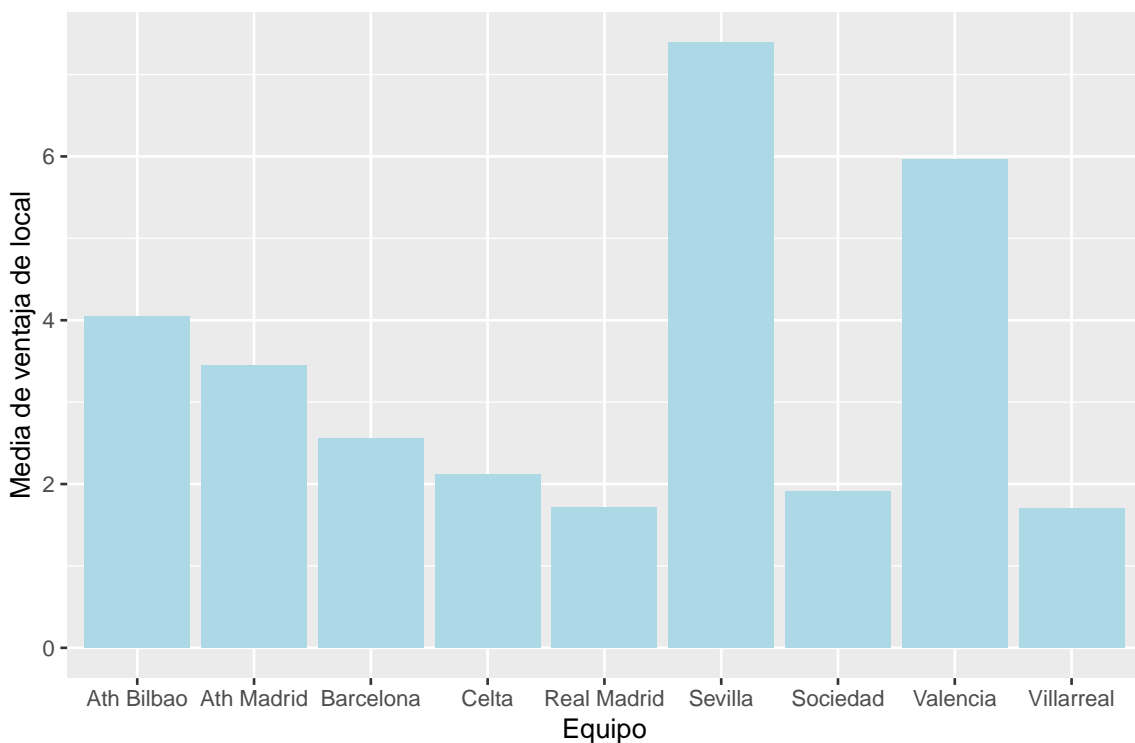
En el siguiente gráfico podemos destacar la ventaja de jugar de local que tuvo el Sevilla F.C. en la temporada 2015/16 de 26.12, la que tuvo el Athletic de Bilbao en la 2016/17 de 15.34 o la que tuvo el Valencia C.F. en la 2019/20 coincidiendo con su centenario de 23.29. Este parámetro no permanece constante a lo largo de los años ya que en él influyen muchos factores que varían en el tiempo como los resultados que va teniendo el equipo, la directiva, el estado de forma, el ambiente que haya entre la afición o la confianza que ésta tenga en el equipo.

### Ventaja individual de jugar de local en las últimas 8 temporadas

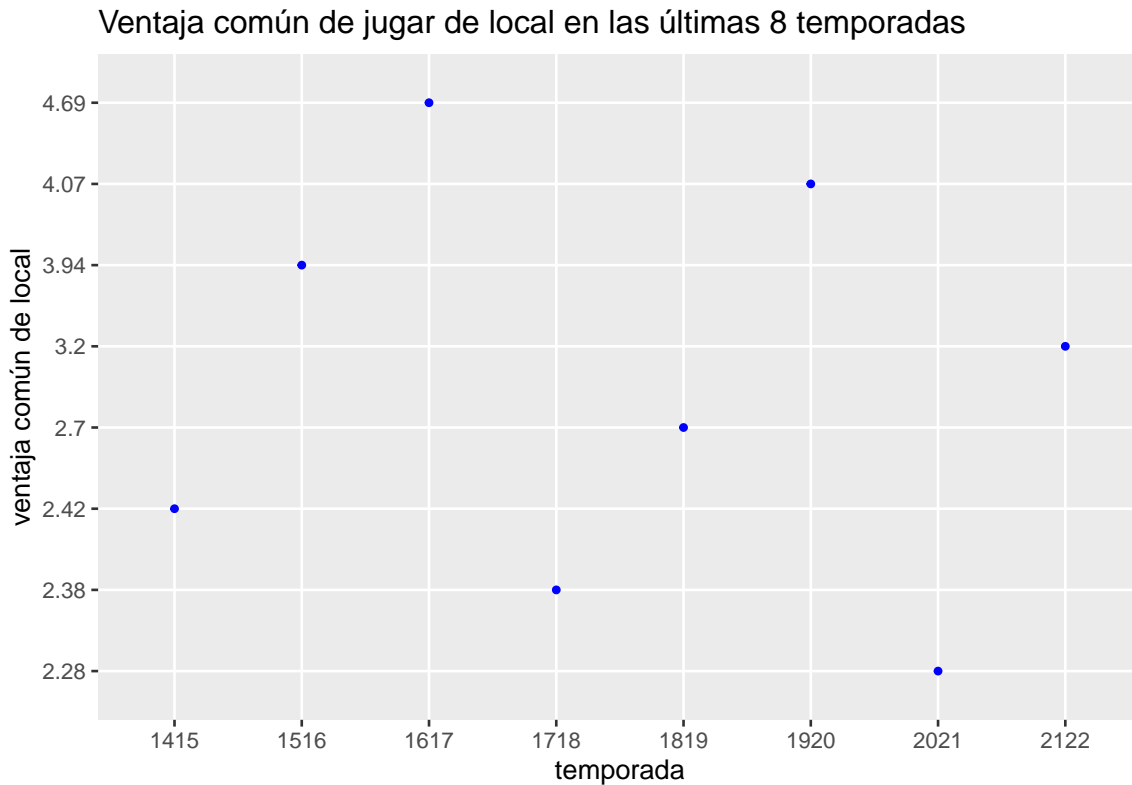


A la vista del siguiente gráfico podemos comprobar que el Sevilla FC es el equipo que más ventaja ha tenido jugando como local en las últimas 8 temporadas de La Liga con una media de 7.39, por el contrario el equipo que menos ventaja ha tenido jugando como local es el Villarreal con 1.70.

### Media de ventaja individual de local durante las últimas 8 temp.



En el siguiente gráfico podemos comprobar que la ventaja común de jugar en casa ha sido menos en la temporada 2020/21 coincidiendo con la temporada en la que los partidos se jugaron sin público a causa de la Covid-19.



Por último vamos a comparar el modelo Bradley-Terry usando la ventaja común de local de cada equipo con el modelo Bradley-Terry usando la ventaja individual de local de cada equipo. Para ello haremos un análisis de la bondad de ajuste de cada modelo para cada temporada. Si tomamos un nivel de significación de 5 %, los p-values de la Tabla 6.5 menores que 0.05 evidenciarán diferencias significativas entre los modelos, pero en los resultados obtenidos para las últimas 8 temporadas los p-values son mayores que 0.05, por lo que no existen evidencias significativas para rechazar la hipótesis de igualdad entre los modelos ya que la diferencia entre los residuos no son extremadamente significativas, pero el modelo usando la ventaja individual de cada equipo tiene durante las 8 temporadas menores residuos que el modelo usando la ventaja común por lo que podemos concluir que el modelo de ventaja individual tiene ligeramente mejor ajuste que el modelo de ventaja común.

Tabla 6.5: Comparación entre modelos de ventaja común e individual por temporada

Temporada	Residuos Comun	Residuos Individual	Dif res.(Chiq)	p-value
2014/15	674.5641	661.0011	13.56298	0.8085458
2015/16	697.9787	674.6862	23.29247	0.2247053
2016/17	659.7188	637.1171	22.60168	0.2553420
2017/18	688.9445	678.2964	10.64810	0.9351089
2018/19	751.4479	731.5644	19.88349	0.4016195
2019/20	720.7824	697.6948	23.08760	0.2335054
2020/21	726.5037	708.0639	18.43980	0.4932631
2021/22	741.0022	715.6956	25.30666	0.1506938



# Apéndice A

## Análisis de La Liga usando el modelo de Davidson en R

A continuación expondremos el código usado para clasificar a los equipos de La Liga en la temporada 2020-2021.

### A.1. Generar la matriz del modelo de Davidson

La siguiente función nos genera la matriz del modelo de Davidson para el número de equipos a comparar que le demos.

```
MatrizModeloEmpates <- function (t){
  x <- matrix(0, (3*t*(t-1))/2, (t+1)+(t*(t-1))/2)
  l=t*(t-1)/2  #total de permutaciones que puede haber ij
  vectori=c()
  for (i in 1:(t-1)){
    vectori=c(vectori,rep(i,(t-i)))
  }
  vectorj=c()
  for (j in 2:t){
    vectorj=c(vectorj,j:t)
  }
  for (k in 1:l) # le doy un numero a cada par ij
  {
    x[3*(k-1)+1,vectori[k]]<-1 #valor de beta_i que está
    #en el indice k
    x[3*(k-1)+2,vectorj[k]]<-1 #valor de beta_j que está en el indice k
    x[3*(k),vectori[k]]<-1/2
    x[3*(k),vectorj[k]]<-1/2
    x[3*(k-1)+1,t+1+k]<-1 #Gana i en el (i,j) serán las filas 1,4,7,10 aij1
    x[3*(k-1)+2,t+1+k]<-1 #Gana j en el (i,j) serán las filas 2,5,8,11 aij2
    x[3*(k),t+1+k]<-1 #Empatan en el (i,j) serán las filas 3,6,9,12
    #coeficiente de aij3
    x[3*(k),t+1]<-1 #Empatan en el (i,j) serán las filas 3,6,9,12
    #coeficiente de beta que será en la columna t+1
  }
}
```

```

return(x) }
MatrizModeloEmpates<-MatrizModeloEmpates(20)

```

## A.2. Analisis para el final de temporada

```

library(dplyr)
#Leo mis datos
epl2021empates <- read.table("liga2021EmpatesImportante.csv",
                           head=T, sep=",")
epl2021empates<-epl2021empates%>%
  select("EquipoA", "EquipoB", "w1jV", "wjiV", "tij")
colnames(epl2021empates)<-c("equipoi", "equipoj", "wij", "wji", "tij")
#Ordeno los datos por equipos alfabeticamente
epl2021empates<-epl2021empates%>%
  arrange(equipoj)%>%
  arrange(equipoi)

#Creo la variable respuesta con distribucion Poisson.
#Vector de  $w_{12}(1), w_{12}(2), t_{12}, \dots, w_{(t-1)t}, w_{(t-1)t}(2), w_{(t-1)t}$ 
epl2021empatesvec<-as.numeric(as.vector(t(epl2021empates[,3:5])))

#Creamos el modelo con poisson y link=log.
epl2021empatesmodelo <- glm(formula =
                           epl2021empatesvec ~ MatrizModeloEmpates+0 ,
                           family = poisson(link = log), maxit=1000000, trace=TRUE)

#Extraemos los coeficientes de mi modelo.
coefEmpate<-epl2021empatesmodelo$coefficients

#Veamos un resumen del modelo
summary(epl2021empatesmodelo)

##Calculemos ahora el valor de los distintos  $\pi_i$ , como
## $\pi_i = \ln(\beta_i) = \text{coeficiente del } 1 \text{ al } 20, \pi_i = \exp(\text{coeficientes})$ 
pisubiEmpate<-exp(coefEmpate[1:20])
epl2021.coefs <- exp(coefEmpate[1:21])
sumacoefEmpate<-sum(exp(coefEmpate[1:20]))
coefNormalizadoEmpate <- pisubiEmpate/sumacoefEmpate
coefNormalizadoEmpate[21]<-exp(coef(epl2021empatesmodelo))[21]
coefEmpate<-coefNormalizadoEmpate[21]
equiposTotales<-c(epl2021empates$equipoi, epl2021empates$equipoj)
equiposEmpates<-equiposTotales[!duplicated(equiposTotales)]

#Creamos una tabla con los distintos parámetros de mi modelo
names(coefNormalizadoEmpate) <- c(equiposEmpates, "tieparameter")
tabla2021<-data.frame(coefNormalizadoEmpate)

```



```

#Creamos una tabla con los distintos parámetros de mi modelo, sin el
#parámetro de empate.
coefNormalizadoEmpate1 <- pisubiEmpate/sumacoefEmpate
names(coefNormalizadoEmpate1) <- c(equiposEmpates)
tabla2021clasi<-data.frame(equiposEmpates,coefNormalizadoEmpate1)
tabla2021clasi<-tabla2021clasi%>%
  arrange(coefNormalizadoEmpate1)

#Los ordenamos en orden descendente
ordenado2021<-arrange(tabla2021clasi,
                      desc(tabla2021clasi$coefNormalizadoEmpate1))

#Preparamos los datos para visualizar el ranking obtenido.
clasificacion=data.frame("Posicion"=c(1:20),ordenado2021)
colnames(clasificacion)<-c("Pos","Equipo","Coeficiente")
#print(clasificacion,row.names = FALSE)
clasificacionReal<-read.csv("clasificacionReal.csv")
knitr::kable(list(clasificacion,clasificacionReal),booktabs=TRUE,
                 caption="\\label{tabla01}Clasificación de La Liga Temporada
                 2020-2021 mediante el modelo Bradley-Terry
                 VS Clasificación Real",
                 row.names = FALSE)

```

### A.3. Análisis para los diferentes periodos de la temporada

Comenzaremos exponiendo el código usado para crear las diferentes clasificaciones a lo largo de los 4 periodos que estudiaremos.

```

#Voy a cargar los resultados ya
#preparados como los necesitamos,
#pero sólo los de las 10 primeras jornadas.
ep12021empatesJ10 <- read.table("jornada10LaLiga2021.csv",
                               head=T, sep=",")
ep12021empatesJ10<-ep12021empatesJ10%>%
  select("EquipoA","EquipoB","w1jV","w1jV","t1j")
colnames(ep12021empatesJ10)<-c("equipoA","equipoB","w1j","w1j","t1j")
#Ordeno los datos por equipos alfabeticamente
ep12021empatesJ10<-ep12021empatesJ10%>%
  arrange(equipoB)%>%
  arrange(equipoA)

#Vector de  $w_{12}(1), w_{12}(2), t_{12}, \dots, w_{(t-1)t}, w_{(t-1)t}(2), w_{(t-1)t}$ 
ep12021empatesJ10vec<-as.numeric(as.vector(t(ep12021empatesJ10[,3:5])))

ep12021empatesJ10modelo <- glm(formula = ep12021empatesJ10vec ~
                               MatrizModeloEmpates+0 ,
                               family = poisson(link = log), maxit=100000, trace=TRUE)

```

```

coefEmpateJ10<-ep12021empatesJ10modelo$coefficients
summary(ep12021empatesJ10modelo)
##Calculemos ahora el valor de los distintos pi_i, como
##pi_i=ln(beta_i)=coeficiente del 1 al 20, pi_i=exp(coeficientes)
pisubiEmpateJ10<-exp(coefEmpateJ10[1:20])
sumacoefEmpateJ10<-sum(exp(coefEmpateJ10[1:20]))
coefNormalizadoEmpateJ10 <- pisubiEmpateJ10/sumacoefEmpateJ10

equiposTotalesJ10<-c(ep12021empatesJ10$equipoi,
                    ep12021empatesJ10$equipoj)
equiposEmpatesJ10<-equiposTotalesJ10[!duplicated(equiposTotalesJ10)]

names(coefNormalizadoEmpateJ10) <- c(equiposEmpatesJ10)
clasificacionJ10<-data.frame(coefNormalizadoEmpateJ10)

clasificacionJ10<-clasificacionJ10%>%
  mutate("Equipos"=equiposEmpatesJ10)
#Los ordenamos en orden descendente
#Obtenemos la clasificación hasta la jornada 10.
clasificacion2021J10<-arrange(clasificacionJ10,
  desc(clasificacionJ10$coefNormalizadoEmpateJ10))%>%
  mutate("Pos"=c(1:20))%>%
  select(Pos,Equipos,coefNormalizadoEmpateJ10)
colnames(clasificacion2021J10)<-c("Pos-J10","Equipos","Coeficiente")

#A continuacion vamos a hacerlo para la jornada 20.
#Voy a cargar los resultados ya preparados como
#los necesitamos, pero sólo los de las 20 primeras jornadas

ep12021empatesJ20 <- read.table("jornada20laliga2021.csv",
                              head=T, sep=",")
ep12021empatesJ20<-ep12021empatesJ20%>%
  select("EquipoA","EquipoB","wijV","wjiV","tij")
colnames(ep12021empatesJ20)<-c("equipoi","equipoj",
                              "wij","wji","tij")
#Ordeno los datos por equipos alfabeticamente
ep12021empatesJ20<-ep12021empatesJ20%>%
  arrange(equipoj)%>%
  arrange(equipoi)

#Vector de w_12(1),w_12(2),t_12,...,w_(t-1)t,w_(t-1)t(2),w_(t-1)t

ep12021empatesJ20vec<-as.numeric(as.vector(t(ep12021empatesJ20[,3:5])))

ep12021empatesJ20modelo <- glm(formula = ep12021empatesJ20vec ~
                              MatrizModeloEmpates+0 ,
family = poisson(link = log), maxit=1000000, trace=TRUE)
coefEmpateJ20<-ep12021empatesJ20modelo$coefficients
summary(ep12021empatesJ20modelo)
##Calculemos ahora el valor de los distintos pi_i, como

```

```

##pi_i=ln(beta_i)=coeficiente del 1 al 20, pi_i=exp(coeficientes)
pisubiEmpateJ20<-exp(coefEmpateJ20[1:20])
sumacoefEmpateJ20<-sum(exp(coefEmpateJ20[1:20]))
coefNormalizadoEmpateJ20 <- pisubiEmpateJ20/sumacoefEmpateJ20

equiposTotalesJ20<-c(ep12021empatesJ20$equipoi,
                    ep12021empatesJ20$equipoj)
equiposEmpatesJ20<-equiposTotalesJ20[!duplicated(equiposTotalesJ20)]

names(coefNormalizadoEmpateJ20) <- c(equiposEmpatesJ20)
clasificacionJ20<-data.frame(coefNormalizadoEmpateJ20)

clasificacionJ20<-clasificacionJ20%>%
  mutate("Equipos"=equiposEmpatesJ20)
##Los ordenamos en orden descendente
clasificacion2021J20<-arrange(clasificacionJ20,
                             desc(clasificacionJ20$coefNormalizadoEmpateJ20))%>%
  mutate("Pos"=c(1:20))%>%
  select(Pos,Equipos,coefNormalizadoEmpateJ20)
colnames(clasificacion2021J20)<-c("Pos-J20", "Equipos", "Coeficiente")

##A continuacion vamos a hacerlo para la jornada 30.
ep12021empatesJ30 <- read.table("jornada30laliga2021.csv",
                              head=T, sep=",")
ep12021empatesJ30<-ep12021empatesJ30%>%
  select("EquipoA", "EquipoB", "wijV", "wjiV", "tij")
colnames(ep12021empatesJ30)<-c("equipoi", "equipoj",
                              "wij", "wji", "tij")

##Ordeno los datos por equipos alfabeticamente
ep12021empatesJ30<-ep12021empatesJ30%>%
  arrange(equipoj)%>%
  arrange(equipoi)

##Vector de w_12(1),w_12(2),t_12,...,w_(t-1)t,w_(t-1)t(2),w_(t-1)t
ep12021empatesJ30vec<-as.numeric(as.vector(t(ep12021empatesJ30[,3:5])))

ep12021empatesJ30modelo <- glm(formula = ep12021empatesJ30vec ~
                              MatrizModeloEmpates+0 ,
                              family = poisson(link = log), maxit=1000000, trace=TRUE)
coefEmpateJ30<-ep12021empatesJ30modelo$coefficients
summary(ep12021empatesJ30modelo)
##Calculamos ahora el valor de los distintos pi_i, como
##pi_i=ln(beta_i)=coeficiente del 1 al 20, pi_i=exp(coeficientes)
pisubiEmpateJ30<-exp(coefEmpateJ30[1:20])
sumacoefEmpateJ30<-sum(exp(coefEmpateJ30[1:20]))
coefNormalizadoEmpateJ30 <- pisubiEmpateJ30/sumacoefEmpateJ30

equiposTotalesJ30<-c(ep12021empatesJ30$equipoi,
                    ep12021empatesJ30$equipoj)

```

```

equiposEmpatesJ30<-equiposTotalesJ30[!duplicated(equiposTotalesJ30)]

names(coefNormalizadoEmpateJ30) <- c(equiposEmpatesJ30)
clasificacionJ30<-data.frame(coefNormalizadoEmpateJ30)

clasificacionJ30<-clasificacionJ30%>%
  mutate("Equipos"=equiposEmpatesJ30)
#Los ordenamos en orden descendente
clasificacion2021J30<-arrange(clasificacionJ30,
  desc(clasificacionJ30$coefNormalizadoEmpateJ30))%>%
  mutate("Pos"=c(1:20))%>%
  select(Pos,Equipos,coefNormalizadoEmpateJ30)
colnames(clasificacion2021J30)<-c("Pos-J30","Equipos","Coeficiente")

#A continuación imprimiremos las clasificaciones para los distintos
#periodos
knitr::kable(list(clasificacion2021J10,
clasificacion2021J20,clasificacion2021J30,clasificacion),
booktabs=TRUE,
caption="\label{tabla02}Clasificación de La Liga
Temporada 2020-2021 mediante el modelo Bradley-Terry en diferentes
jornadas",row.names = FALSE)

```

A continuación exponemos el código usado para calcular los dos tipos de errores que hemos descrito en la sección correspondiente.

```

#Vamos a empezar calculando el er de la J-10
clasificacionReal<-clasificacionReal%>%
  mutate(Pos=c(1:20))
Completay10<-clasificacion2021J10%>%
  full_join(clasificacionReal,clasificacion2021J10,by="Equipos")%>%
  group_by("Equipos")%>%
  mutate("rest"=Pos-`Pos-J10`)%>%
  filter(rest==0)
erJ10<-(20-nrow(Completay10))/20*100

#Vamos a empezar calculando el er de la J-20
Completay20<-clasificacion2021J20%>%
  full_join(clasificacionReal,clasificacion2021J20,by="Equipos")%>%
  group_by("Equipos")%>%
  mutate("rest"=Pos-`Pos-J20`)%>%
  filter(rest==0)
erJ20<-(20-nrow(Completay20))/20*100

#Vamos a empezar calculando el er de la J-30
Completay30<-clasificacion2021J30%>%
  full_join(clasificacionReal,clasificacion2021J30,by="Equipos")%>%
  group_by("Equipos")%>%
  mutate("rest"=Pos-`Pos-J30`)%>%
  filter(rest==0)
erJ30<-(20-nrow(Completay30))/20*100

```

```

#Vamos a empezar calculando el er de la J-30
Completay38<-clasificacion%>%
  full_join(clasificacionReal,clasificacion,by="Equipos")%>%
  group_by("Equipos")%>%
  mutate("rest"=`Pos.x`-`Pos.y`)%>%
  filter(rest==0)
erJ38<-(20-nrow(Completay38))/20*100

#Vamos a empezar calculando el em de la J-10
Completay10m<-clasificacion2021J10%>%
  full_join(clasificacionReal,clasificacion2021J10,by="Equipos")%>%
  group_by("Equipos")%>%
  mutate("rest"=abs(Pos-`Pos-J10`))
cambios10m<-sum(Completay10m$rest)
emJ10<-cambios10m/20

#Vamos a empezar calculando el em de la J-20
Completay20m<-clasificacion2021J20%>%
  full_join(clasificacionReal,clasificacion2021J20,by="Equipos")%>%
  group_by("Equipos")%>%
  mutate("rest"=abs(Pos-`Pos-J20`))
cambios20m<-sum(Completay20m$rest)
emJ20<-cambios20m/20

#Vamos a empezar calculando el em de la J-30
Completay30m<-clasificacion2021J30%>%
  full_join(clasificacionReal,clasificacion2021J30,by="Equipos")%>%
  group_by("Equipos")%>%
  mutate("rest"=abs(Pos-`Pos-J30`))
cambios30m<-sum(Completay30m$rest)
emJ30<-cambios30m/20

#Vamos a empezar calculando el em de la J-38
Completay38m<-clasificacion%>%
  full_join(clasificacionReal,clasificacion,by="Equipos")%>%
  group_by("Equipos")%>%
  mutate("rest"=abs(`Pos.x`-`Pos.y`))
cambios38m<-sum(Completay38m$rest)
emJ38<-cambios38m/20

#Creamos la tabla final con todos los errores.
Errores<-data.frame(c("J10","J20","J30","J38"),
c(erJ10,erJ20,erJ30,erJ38),c(emJ10,emJ20,emJ30,emJ38))
names(Errores)<-c("Periodo de la Temporada","er %","em")
knitr::kable(Errores,caption="\label{tabla03} Errores
calculados en los distintos periodos de la
temporada con respecto a la clasificación final",
row.names = FALSE)

```



# Apéndice B

## Análisis de la ventaja de jugar de local común e individual para los distintos equipos de La Liga

A continuación expondremos el código usado para calcular la ventaja común e individual de los equipos a los largos de las últimas 8 temporadas.

### B.1. Generar la matriz del modelo con ventaja común

La siguiente función nos genera la matriz del modelo de Bradley-Terry con ventaja común.

```
MatrizComún <- function (t){
  x <- matrix(0, (3*t*(t-1)), (t+2)+(t*(t-1)))
  l=t*(t-1)  #total de permutaciones que puede haber ij
  vectori=c()
  vectorj=c()
  for (i in 1:t){
    vectori=c(vectori,rep(i,(t-1)))
    if(i==1){
      vectorj=c(vectorj,2:t)
    }

    else{
      if(i==t){
        vectorj=c(vectorj,1:(t-1))
      }
      else
      {vectorj=c(vectorj,c(1:(i-1),(i+1):t))}
    }
  }

  for (k in 1:l) # le doy un numero a cada par ij
```

```

{
  x[3*(k-1)+1,vectori[k]]<-1 #valor de beta_i que está en el indice
                                #k del vector i
  x[3*(k-1)+2,vectorj[k]]<-1 #valor de beta_j que está en el indice
                                #k del vector j
  x[3*(k),vectori[k]]<-1/2
  x[3*(k),vectorj[k]]<-1/2

  x[3*(k),t+2]<-1 #Empatan en el (i,j) serán las filas 3,6,9,12
                    #coeficiente de beta que será en la columna t+2
  x[3*(k-1)+1,t+1]<-1 #Comun en el (i,j) serán las filas 1,4,7,10
                    #coeficiente de \gamma que será en la columna t+1
  x[3*(k),t+1]<-1/2 #Comun en el (i,j) serán las filas 3,6,9,12
                    #coeficiente de \gamma que será en la columna t+1

  x[3*(k-1)+1,t+1+(k+1)]<-1 #Gana i en el (i,j) serán las filas
                                #1,4,7,10 aij1
  x[3*(k-1)+2,t+1+(k+1)]<-1 #Gana j en el (i,j) serán las filas
                                #2,5,8,11 aij2
  x[3*(k),t+1+(k+1)]<-1 #Empatan en el (i,j) serán las filas 3,6,9,12
                                #coeficiente de aij3
}
return(x) }

```

## B.2. Generar la matriz del modelo con ventaja individual

La siguiente función nos genera la matriz del modelo de Bradley-Terry con ventaja individual.

```

MatrizComún1 <- function (t){
  x <- matrix(0, (3*t*(t-1)), ((2*t)+1)+(t*(t-1)))
  l=t*(t-1) #total de permutaciones que puede haber ij
  vectori=c()
  vectorj=c()
  for (i in 1:t){
    vectori=c(vectori,rep(i,(t-1)))
    if(i==1){
      vectorj=c(vectorj,2:t)
    }

    else{
      if(i==t){
        vectorj=c(vectorj,1:(t-1))
      }
      else
        {vectorj=c(vectorj,c(1:(i-1),(i+1):t))}
    }
  }
}

```



```

for (k in 1:l) # le doy un numero a cada par ij
{
  x[3*(k-1)+1,vectori[k]]<-1 #valor de beta_i que
  #está en el indice k del vectori
  x[3*(k-1)+2,vectorj[k]]<-1 #valor de beta_j que
  #está en el indice k del vector
  x[3*(k),vectori[k]]<-1/2
  x[3*(k),vectorj[k]]<-1/2

  x[3*(k-1)+1,t+(vectori[k])]<-1 #Comun en el (i,j) serán
  #las filas 1,4,7,10
  #coeficiente de \gamma
  #que será en la columna t+1
  x[3*(k),t+(vectori[k])]<-1/2 #Comun en el (i,j) serán
  #las filas 3,6,9,12
  #coeficiente de \gamma
  #que será en la columna t+1

  x[3*(k),2*t+1]<-1 #Empatan en el (i,j) serán
  #las filas 3,6,9,12
  #coeficiente de beta
  #que será en la columna t+2

  x[3*(k-1)+1,(2*t)+(k+1)]<-1 #Gana i en el (i,j)
  #serán las filas 1,4,7,10 aij1
  x[3*(k-1)+2,(2*t)+(k+1)]<-1 #Gana j en el (i,j)
  #serán las filas 2,5,8,11 aij2
  x[3*(k),(2*t)+(k+1)]<-1 #Empatan en el (i,j)
  #serán las filas 3,6,9,12 #coeficiente de aij3
}
return(x) }

```

### B.3. Creamos la tabla final con las ventajas individuales y comunes de todos los equipos a lo largos de las últimas 8 temporadas

```

library(dplyr)
#Leo mis datos
ep12122LaLiga1 <- read.table("Liga2122.csv",
                             head=T, sep=",")
ep12122LaLiga1<-ep12122LaLiga1%>%select("HomeTeam","AwayTeam","FTR")

#Proceso los datos para quedarme sólo con lo que me interesa
functionMas<-function(i){
  ep12122LaLiga1%>%
  filter(HomeTeam %in% c(ep12122LaLiga1[i,1]) &

```

```

        AwayTeam %in% c(ep12122LaLiga1[i,2]))%>%
mutate("wij(1)"=ifelse(FTR=="H",1,0),
       "wij(2)"=ifelse(FTR=="A",1,0),
       "tij"=ifelse(FTR=="D",1,0),
       "Equipo i"=HomeTeam,
       "Equipo j"=AwayTeam)%>%
select("Equipo i","Equipo j","wij(1)","wij(2)","tij")%>%
arrange("Equipo i")}
functionMas(1)

l=nrow(ep12122LaLiga1)
equipoi=c()
equipoj=c()
wij1=c()
wij2=c()
tij=c()
for (i in 1:l){
  equipoi=c(equipoi,functionMas(i)[1,1])
  equipoj=c(equipoj,functionMas(i)[1,2])
  wij1=c(wij1,functionMas(i)[1,3])
  wij2=c(wij2,functionMas(i)[1,4])
  tij=c(tij,functionMas(i)[1,5])
}

##Vamos a poner los datos como los necesitamos

ep12122LaLigaha<-data.frame(equipoi,equipoj,wij1,wij2,tij)
#Ordeno los datos por equipos locales alfabeticamente
ep12122LaLigaha<-ep12122LaLigaha%>%
  arrange(equipoj)%>%
  arrange(equipoi)

#Vector de w_12(1),w_12(2),t_12,...,w_t(t-1)(1),w_t(t-1)(2),t_t(t-1)

ep12122LaLigavec<-as.numeric(as.vector(t(ep12122LaLigaha[,3:5])))

# Creamos el modelo

eplmodelo2122LaLiga <- glm(formula = epl2122LaLigavec ~
                          MatrizModelo+0 ,
                          family = poisson(link = log), maxit=1000000,
                          trace=TRUE)
eplmodelo2122LaLiga$coefficients

##Calculemos ahora el valor de los distintos pi_i, como
##pi_i=ln(beta_i)=coeficiente del 1 al 20, pi_i=exp(coeficientes)
pisubi2122LaLiga<-exp(eplmodelo2122LaLiga$coefficients[1:20])
sumacoef2122LaLiga<-sum(exp(coef(eplmodelo2122LaLiga))[1:20])
coefNormalizado2122LaLiga <- pisubi2122LaLiga/sumacoef2122LaLiga
coefNormalizado2122LaLiga[21]<-exp(coef(eplmodelo2122LaLiga))[21]
coefNormalizado2122LaLiga[22]<-exp(coef(eplmodelo2122LaLiga))[22]

```

```

equipos2122LaLiga<-ep12122LaLigaha$equipoi
[!duplicated(ep12122LaLigaha$equipoi)]
names(coefNormalizado2122LaLiga) <- c(equipos2122LaLiga,
"homeadv2122", "tieparameter")
tabla2122LaLiga<-data.frame(coefNormalizado2122LaLiga)

#En este caso estamos estudiando el parámetro de ventaja
#común de jugar de local
#por lo que sólo me interesaría el coefNormalizado2122LaLiga[21]
ventajacomun2122LaLiga<-coefNormalizado2122LaLiga[21]
empatecomun2122LaLiga<-coefNormalizado2122LaLiga[22]

#Repito el proceso anterior para la temporada 2021,1920,1819,1718,1617,
#1516 y 1415
ventajas8ultimastemporadasr<-c(ventajacomun1415LaLiga,
ventajacomun1516LaLiga,ventajacomun1617LaLiga,
ventajacomun1718LaLiga,ventajacomun1819LaLiga,
ventajacomun1920LaLiga,ventajacomun2021LaLiga,
ventajacomun2122LaLiga)
ventajas8ultimastemporadas<-round(ventajas8ultimastemporadasr,2)

#A continuacion calculo las ventajas individuales para todas
#las temporadas.
library(tidyverse)
#Leo mis datos
ep12122LaLiga2 <- read.table("Liga2122.csv",
                             head=T, sep=",")
ep12122LaLiga2<-ep12122LaLiga2%>%select("HomeTeam","AwayTeam","FTR")

#Proceso los datos para quedarme sólo con lo que me interesa
functionMas1<-function(i){
  ep12122LaLiga2%>%
    filter(HomeTeam %in% c(ep12122LaLiga2[i,1])
           & AwayTeam %in% c(ep12122LaLiga2[i,2]))%>%
    mutate("wij(1)"=ifelse(FTR=="H",1,0),
           "wij(2)"=ifelse(FTR=="A",1,0),
           "tij"=ifelse(FTR=="D",1,0),
           "Equipo i"=HomeTeam,
           "Equipo j"=AwayTeam)%>%
    select("Equipo i","Equipo j",
           "wij(1)","wij(2)","tij")%>%
    arrange("Equipo i")}
functionMas1(1)

l=nrow(ep12122LaLiga2)
equipoi=c()
equipoj=c()
wij1=c()
wij2=c()
tij=c()
for (i in 1:l){

```

```
equipoi=c(equipoi,functionMas1(i)[1,1])
equipoj=c(equipoj,functionMas1(i)[1,2])
wij1=c(wij1,functionMas1(i)[1,3])
wij2=c(wij2,functionMas1(i)[1,4])
tij=c(tij,functionMas1(i)[1,5])
}

##Vamos a poner los datos como los necesitamos

ep12122LaLigaha1<-data.frame(equipoi,equipoj,wij1,wij2,tij)

ep12122LaLigaha1<-ep12122LaLigaha1%>%
  arrange(equipoj)%>%
  arrange(equipoi)

#Vector de w_12(1),w_12(2),t_12,...,w_t(t-1)(1),w_t(t-1)(2),t_t(t-1)

ep12122LaLigavec<-as.numeric(as.vector(t(ep12122LaLigaha1[,3:5])))

# Creamos el modelo

ep12122LaLigamodelo1 <- glm(formula = ep12122LaLigavec ~ MatrizModelo1+0 ,
  family = poisson(link = log), maxit=1000000,
  trace=TRUE)
ep12122LaLigamodelo1$coefficients

##Calculemos ahora el valor de los distintos \gamma_{i}

equipos2122LaLiga1<-ep12122LaLigaha1$equipoi
[!duplicated(ep12122LaLigaha1$equipoi)]

#Vamos a calcular la ventaja de jugar en casa para todos los años.
ventaja2122Liga<-round(exp(coef(ep12122LaLigamodelo1))[21:40],2)
tabla2122Liga1<-data.frame(equipos2122LaLiga1,ventaja2122Liga)
colnames(tabla2122Liga1)<-c("Equipos","2122")
print(tabla2122Liga1,row.names = FALSE)

#Hago lo mismo para Temporada 2021,1920,1819,1718,1617,1516 y 1415

#Creamos la tabla con todos los datos juntos cada uno con
#su respectivo equipo.
tablafinal1<-full_join(tabla2021LaLiga1,
  tabla2122Liga1,by = "Equipos")
tablafinal2<-full_join(tabla1920LaLiga1,
  tablafinal1,by = "Equipos")
tablafinal3<-full_join(tabla1819LaLiga1,
  tablafinal2,by = "Equipos")
tablafinal4<-full_join(tabla1718LaLiga1,
  tablafinal3,by = "Equipos")
tablafinal5<-full_join(tabla1617LaLiga1,
```

```

        tablafinal4,by = "Equipos")
tablafinal6<-full_join(tabla1516LaLiga1,
        tablafinal5,by = "Equipos")
tablafinal<-full_join(tabla1415LaLiga1,
        tablafinal6,by = "Equipos")

#Eliminamos los valores que tengan valores NA, para calcular
#las dos ultimas filas y la última columna.
tablaFinalSinNa<-tablafinal
tablaFinalSinNa[is.na(tablafinal)]<-0
x<-cbind(as.numeric(tablaFinalSinNa[,2]),as.numeric(tablaFinalSinNa[,3]),
        as.numeric(tablaFinalSinNa[,4]),as.numeric(tablaFinalSinNa[,5]),
        as.numeric(tablaFinalSinNa[,6]),as.numeric(tablaFinalSinNa[,7]),
        as.numeric(tablaFinalSinNa[,8]),as.numeric(tablaFinalSinNa[,9]))

#Nos da el numero de temporadas que cada equipo ha estado en Primera
#de las 8 posibles.
##Ese vector que me da el numero de temporadas en primera de cada equipo
#dividido entre la suma total de los parametros
#de ventaja de jugar de local nos da la media individual en todas
#las temporadas en Primera.
functionTemporadas<-function(i){
  8-sum(is.na(tablafinal[i,2:9]))
}
functionTemporadas(1)
n=nrow(tablafinal)
temporadasi=c()
for (i in 1:n){
  temporadasi=c(temporadasi,functionTemporadas(i))
}
avgr<-c(rowSums(x)/temporadasi,NA,NA)
avg<-round(avgr,2)
#Nos falta hacer una media por temporada de
#los parametros individuales de ventaja
avgir<-c(colSums(x)/20)
avgi<-round(avgir,2)

tablafinal[32,]<-c("avgi",avgi)
tablafinal[33,]<-c("gamma",ventajas8ultimastemporadas)
tablafinal[,10]<-c(avg)
tablafinal<-tablafinal%>%arrange(desc(avg))
colnames(tablafinal)<-c("Equipos","1415","1516",
"1617","1718","1819","1920","2021","2122","avg")

```

## B.4. Creamos los gráficos usados

```

#En el primer grafico representamos la Ventaja
#individual de jugar de local en las últimas 8 temporadas.

```

```

tablafinalCompleta <- drop_na(tablafinal)[,1:9]
ventaja_plot <- tablafinalCompleta %>%
  gather(key = time, value = read_counts, -Equipos)
# View(tr_plot)
grafico1<-ggplot(data = ventaja_plot, aes(x = time,
                                          y = as.numeric(read_counts),
                                          group = Equipos)) +
  geom_line(aes(color = Equipos))
grafico1+labs(x="temporada",y="ventaja de local",
              title="Ventaja individual de jugar de local
                    en las últimas 8 temporadas")

#En el segundo grafico representamos la Media de ventaja
#individual de local durante las últimas 8 temp.
tablafinalG<-drop_na(tablafinal)[,c(1,10)]
grafico2<-ggplot(tablafinalG)+geom_bar(mapping=aes(x=Equipos,y=avg),
                                          stat="identity",fill="lightblue")+
  labs(x="Equipo",y="Media de ventaja de local",
       title="Media de ventaja individual
              de local durante las últimas 8 temp.")

#En el tercer grafico representamos la Ventaja común de jugar de local
#en las últimas 8 temporadas.
tablafinalG1<-tablafinal[32,c(1:9)]
ventaja_plotG1 <- tablafinalG1 %>%
  gather(key = time, value = avg1 , -Equipos)
# View(tr_plot)
grafico3<-ggplot(data = ventaja_plotG1, aes(x = time, y = avg1)) +
  geom_point(size=1,colour="blue")
grafico3+labs(x="temporada",y="ventaja común de local",
              title="Ventaja común de jugar de local
                    en las últimas 8 temporadas")

```

## B.5. Generamos la tabla para comparar los ajustes de los modelos

```

#Realizamos un anova entre los modelos con ventaja individual
#y común en cada temporada.
anova1<-anova(eplmodelo1415LaLiga,
              epl1415LaLigamodelo1,test="Chisq")
anova2<-anova(eplmodelo1516LaLiga,
              epl1516LaLigamodelo1,test="Chisq")
anova3<-anova(eplmodelo1617LaLiga,
              epl1617LaLigamodelo1,test="Chisq")
anova4<-anova(eplmodelo1718LaLiga,
              epl1718LaLigamodelo1,test="Chisq")
anova5<-anova(eplmodelo1819LaLiga,
              epl1819LaLigamodelo1,test="Chisq")

```

```

anova6<-anova(eplmodelo1920LaLiga,
               epl1920LaLigamodelo1,test="Chisq")
anova7<-anova(eplmodelo2021LaLiga,
               epl2021LaLigamodelo1,test="Chisq")
anova8<-anova(eplmodelo2122LaLiga,
               epl2122LaLigamodelo1,test="Chisq")

Temporada<-c("2014/15", "2015/16", "2016/17", "2017/18",
             "2018/19", "2019/20", "2020/21", "2021/22")
Comun<-c(anova1$`Resid. Dev` [1], anova2$`Resid. Dev` [1],
         anova3$`Resid. Dev` [1], anova4$`Resid. Dev` [1],
         anova5$`Resid. Dev` [1], anova6$`Resid. Dev` [1],
         anova7$`Resid. Dev` [1], anova8$`Resid. Dev` [1])
Individual<-c(anova1$`Resid. Dev` [2], anova2$`Resid. Dev` [2],
             anova3$`Resid. Dev` [2], anova4$`Resid. Dev` [2],
             anova5$`Resid. Dev` [2], anova6$`Resid. Dev` [2],
             anova7$`Resid. Dev` [2], anova8$`Resid. Dev` [2])
Chiq<-c(anova1$Deviance [2], anova2$Deviance [2],
        anova3$Deviance [2], anova4$Deviance [2],
        anova5$Deviance [2], anova6$Deviance [2],
        anova7$Deviance [2], anova8$Deviance [2])
pvalue<-c(anova1$`Pr(>Chi)` [2], anova2$`Pr(>Chi)` [2],
          anova3$`Pr(>Chi)` [2], anova4$`Pr(>Chi)` [2],
          anova5$`Pr(>Chi)` [2], anova6$`Pr(>Chi)` [2],
          anova7$`Pr(>Chi)` [2], anova8$`Pr(>Chi)` [2])
comparacion<-data.frame(Temporada,Comun,Individual,Chiq,pvalue)
knitr::kable(comparacion,
             caption="\\label{tabla05} Comparación entre
             modelos de ventaja común e individual por temporada",
             row.names = FALSE)

```





# Bibliografía

- Agresti, A. (2002). *Categorical data analysis*. Wiley-Interscience, 2nd ed.
- Allaire, J., Xie, Y., McPherson, J., Luraschi, J., Ushey, K., Atkins, A., Wickham, H., Cheng, J., Chang, W., and Iannone, R. (2022). *rmarkdown: Dynamic Documents for R*. R package version 2.14.
- Atkinson, D. R., Wampold, B. E., Lowe, S. M., Matthews, L., and nie Ahn, H. (1998). Asian american preferences for counselor characteristics: Application of the bradley-terry-luce model to comparison data. *Counseling Psychologist*, 26:101–123.
- Baker, S. G. (1994). The multinomial-poisson transformation. *Journal of the Royal Statistical Society. Series D (The Statistician)*., 43(4):495–504.
- Bishop, Y. M. M., Holland, P., and Fienberg, S. E. (1975). Discrete multivariate analysis: Theory and practice.
- Bradley, R. (1965). Another interpretation of a model for paired comparisons. *Psychometrika*, 30(3):315–318.
- Bradley, R. and Terry, M. (1952). Rank analysis of incomplete block designs: The method of paired comparisons. *Biometrika*, 39(3):324–345.
- Caron, F. and Doucet, A. (2012). Efficient bayesian inference for generalized bradley-terry models. *Journal of Computational and Graphical Statistics.*, 21(1):174–196.
- Cattelan, M. (2012). Models for paired comparison data: A review with emphasis on dependent data. *Statistical Science*, 27(3):414–433.
- Critchlow, D. E. and Fligner, M. A. (1991). Paired comparison, triple comparison, and ranking experiments as generalized linear models, and their implementation on glim. *Psychometrika.*, 56(3):517–533.
- Davidson, R. R. (1970). On extending the bradley-terry model to accommodate ties in paired comparison experiments. *Journal of the American Statistical Association*, 65(329):317–328.
- Davidson, R. R. and Beaver, R. J. (1977). On extending the bradley-terry model to incorporate within-pair order effects. *Biometrics.*, 33(4):693–702.
- Glickman, M. E. (2011). Introductory note to 1928.
- Guttman, L. (1946). An approach for quantifying paired comparisons and rank order. *The Annals of Mathematical Statistics*, 17(2):144–163.
- Heather Turner, D. F. and Robertson, G. (2013). Generalized bradley-terry modelling of football results. *University of Warwick, UK. Independent statistical.*

- Hollander, M. (2000). A conversation with ralph a. bradley. *Statistical Science*, 16(1):75–100.
- Hunter, D. (2004). Mm algorithms generalized bradley-terry models. *The Annals of Statistics*, 32(1):384–406.
- Ibarrola, R. V. and Perez, A. G. (2002). *Principios de Inferencia Estadística*. Madrid, UNED.
- Jamil, H. (2010). Analysis of paired comparison data using bradley-terry models with applications to football data. *University of Warwick, UK. Statistics Department*.
- Kendall, M. G. and Smith, B. B. (1940). On the method of paired comparisons. *Biometrika*, 31(3/4):324–345.
- Luce, R. D. (2008). Luce’s choice axiom. *Scholarpedia*, 3(12):8077.
- Luque-Calvo, P. L. (2017). *Escribir un Trabajo Fin de Estudios con R Markdown*.
- Luque-Calvo, P. L. (2019). *Cómo crear Tablas de información en R Markdown*.
- Matthews, J. N. S. and Morris, K. P. (1995). An application of bradley-terry-type models to the measurement of pain. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 44(2):243–255.
- Rao, P. V. and Kupper, L. L. (1967). Ties in paired-comparison experiments: A generalization of the bradley-terry model. *Journal of the American Statistical Association*, 62(317):194–204.
- Sham, P. and Curtis, D. (1995). An extended transmission/disequilibrium test (tdt) for multi-allele marker loci. *Annals of Human Genetics.*, 59(3):323–336.
- Thurstone, L. (1927). A law of comparative judgment. *Psychological review*, 34(4):273–286.
- Wickham, H., Chang, W., Henry, L., Pedersen, T. L., Takahashi, K., Wilke, C., Woo, K., Yutani, H., and Dunnington, D. (2022a). *ggplot2: Create Elegant Data Visualisations Using the Grammar of Graphics*. R package version 3.3.6.
- Wickham, H., François, R., Henry, L., and Müller, K. (2022b). *dplyr: A Grammar of Data Manipulation*. R package version 1.0.9.
- Xie, Y. (2022). *knitr: A General-Purpose Package for Dynamic Report Generation in R*. R package version 1.39.
- Zermelo, E. (1929). Die berechnung der turnier-ergebnisse als ein maximumproblem der wahr-scheinlichkeitsrechnung. *Mathematische Zeitschrift*, 29(1):436–460.