Instituto de Matemática Interdisciplinar

IMI Publica

Boletín del IMI: https://ucm.es/imi/boletindelimi

UNIVERSIDAD COMPLUTENSE MADRID

Análisis crítico y propuestas de mejora del sorteo UEFA para la fase segunda de la Champions.

T. R. Basile Álvarez[a], J. A. Gallegos Salgado[a].

[a] Universidad Nacional Autónoma de México (UNAM).

Boletín del IMI Nº 70, (1 Dic. 2022), Sección "IMI-Publica".

https://doi.org/10.57037/b-imi.00070.imip

Tomás Ricardo Basile Álvarez

Jessica Andrea Gallegos Salgado

Abstract:

En este documento se muestra el trabajo presentado al V Concurso de Modelización Matemática del IMI (CMM IMI 2022; http://blogs.mat.ucm.es/cmm/edicion-2022/), en el que el problema propuesto llevaba por título "Análisis crítico y propuestas de mejora del sorteo UEFA para la fase segunda de la Champions".

En el trabajo se analiza el método de sorteo que sigue la UEFA para emparejar a los equipos en la fase de octavos de final de la Champions y se explica por qué no es un método equitativo. Para determinar qué tan inequitativo es el sorteo, se proponen varias medidas distintas que lo cuantifican. Posteriormente, se proponen métodos alternativos de sorteo, principalmente basados en minimizar alguna de las medidas de inequidad y se muestra un método de sorteo que consideramos el mejor, al ser el más equitativo a nuestro parecer. Finalmente, se ponen a prueba todos los métodos de sorteo para 5 temporadas de la Champions y se comparan los resultados de cada método.

El código Python que hace todo el análisis necesario para obtener los resultados que se presentan en este trabajo se encuentra disponible para ejecutar en Google Colab o para descargar en https://www.ucm.es/imi/file/codigo_cmm22.

Índice general

Capítulo 1

Introducción

Cada año, la Unión de Federaciones Europeas de Fútbol (UEFA) lleva a cabo la Champions League, una competencia en la que los mejores equipos de fútbol de europa se enfrentan entre sí. La competencia se lleva a cabo en los siguientes pasos [1]:

1. **Clasificatoria:** Antes de la competencia, 32 equipos de varios países europeos son clasificados a participar según sus resultados en sus ligas locales.

2. **Fase de Grupos:** Los 32 equipos son separados en 8 grupos de 4 equipos cada uno. Posteriormente, los equipos de cada grupo compiten entre sí (cada uno juega dos partidos con sus tres contrincantes) y queda un campeón y un subcampeón de cada grupo.

3. **Fase Eliminatoria:** Los 16 equipos que fueron campeones y subcampeones de grupo durante la fase de grupos pasan a la fase eliminatoria, la cual consiste en emparejarlos en 8 partidos (conocidos como octavos de final). Luego, los perdedores de estos partidos serán eliminados del torneo y los ganadores pasarán a la siguiente fase (cuartos de final), y así hasta llegar a la final y obtener al campeón del torneo.

Lo que nos concierne para este trabajo es el sorteo que se realiza para la fase eliminatoria, en el cual los 16 equipos tienen que ser agrupados en 8 parejas. Este agrupamiento se puede hacer de varias formas, pero para que el torneo sea lo más atractivo posible, la UEFA establece tres condiciones que se deben de cumplir.

Condiciones para la elección de la fase eliminatoria:

- **Condición 1:** Debe enfrentarse un campeón con un subcampeón.
- **Condición 2:** No deben enfrentarse equipos que ya hayan coincidido en la fase de grupos.
- **Condición 3:** No pueden enfrentarse equipos de la misma nacionalidad.

Podría pensarse que estas condiciones no complican demasiado el proceso de agrupación de los 16 equipos, sin embargo, esto no es así. Al considerar las tres condiciones, la cantidad de emparejamientos válidos para la fase de eliminación se reducen significativamente y esto complica la forma en que se puede llevar a cabo el sorteo.

Además, surge un problema probabilístico, ya que sin importar cómo se haga el sorteo bajo estas condiciones, habrán partidos que tengan más probabilidad de escogerse que otros. Esto es un gran problema, ya que se puede beneficiar a ciertos equipos y perjudicar a otros. Por ejemplo, pensemos en el caso de un equipo subcampeón "A" que tiene mucha probabilidad de jugar con el equipo campeón más fuerte y un segundo equipo subcampeón "B" que tiene mucha probabilidad de jugar con el equipo campeón más débil. Esto claramente marca una desventaja para el equipo "A" respecto al "B".

El problema es que las condiciones que plantea la UEFA son demasiado restrictivas y siempre habrá desigualdades como la mencionada anteriormente. Por lo tanto, lo que se busca encontrar en este trabajo son distintas formas de realizar el sorteo que se atenga a las condiciones, pero de tal forma que las probabilidades de los distintos partidos sean lo más equitativas posible.

En el capítulo 2 del trabajo se describirá el proceso de sorteo que sigue la UEFA y se presentarán algunos de los conceptos fundamentales para el resto del trabajo. Además, se verá de forma clara el problema que tiene el sorteo y se definirán distintos criterios para cuantificar qué tan equitativo es el sorteo.

En el capítulo 3 se enlistan algunos métodos de sorteo alternativos al de la UEFA y de cada uno se mide su equidad utilizando los criterios definidos en el capítulo 2. Al final del capítulo se comparan estos métodos y se propone el método que consideramos mejor de entre los propuestos.

Posteriormente, en el capítulo 4 se muestran los resultados de utilizar los distintos métodos de sorteo para varias ediciones de la competencia. Y finalmente, en el capítulo 5 se presentan las conclusiones del trabajo.

Capítulo 2

Establecer el Problema

Por ahora, para introducir el problema y ver posibles soluciones, vamos a hablar únicamente de la Champions del año 2021. Más adelante, en el capítulo 4, aplicaremos los métodos a otros años. Para empezar, en la figura 2.1 se observan los equipos campeones (segunda columna) y subcampeones (tercera columna) de cada uno de los 8 grupos de la fase de grupos.

Grupo	Campeones	Subcampeones
A	Manchester City (MCI)	Paris Saint-Germain (PSG)
B	Liverpoool (LIV)	Atlético de Madrid (ATM)
C	Ajax (AJX)	Sporting de Portugal (SPO)
D	Real Madrid (RMA)	Inter de Milán (INT)
E	Bayern Múnich (BAY)	Benfica (BEN)
F	Manchester United (MAU)	Villareal (VIL)
G	Lille (LIL)	Salzburgo (SAL)
H	Juventus (JUV)	Chelsea (CHE)

Figura 2.1: Equipos de la Fase de Eliminación, Champions 2021

Una vez que ya se tienen los equipos campeones y subcampeones, resta decidir quién jugará con quién y así determinar la fase eliminatoria, siguiendo las tres condiciones de la UEFA de las que hablamos en la introducción. Para poder visualizar mejor el problema al que nos enfrentamos y el efecto que tienen las tres condiciones, vamos a usar redes. En las redes que usaremos, los 16 equipos serán los nodos y va a haber una conexión entre dos de ellos si es posible que jueguen entre sí, siguiendo las condiciones.

Ninguna condición

Si no seguimos ninguna de las tres condiciones, todos los equipos pueden jugar con cualquier otro equipo, es decir, todos los nodos se conectan entre sí. Esto significa que hay 81,729,648,000 posibles fases de eliminación posibles (que es la cantidad de maneras en que 16 equipos se pueden dividir en 8 parejas). Se visualiza la red de la figura 2.2, en donde cada línea es un posible partido y las lineas rojas representan los 8 partidos de una fase de eliminación elegida al azar de entre las 81,729,648,000 fases de eliminación posibles.

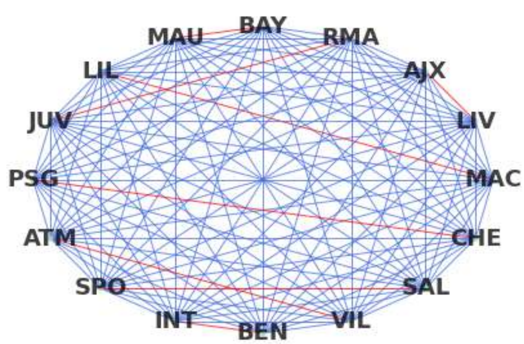

Figura 2.2: Red con todos los partidos posibles sin condiciones. En rojo se marca una posible fase de eliminación.

Primera condición

En este caso, los campeones tienen que jugar con subcampeones y no hay ninguna restricción adicional. Con ello, las posibles fases eliminatorias se reducen a la cifra de $8! = 40,320$.

Podemos observar la red que se genera con esta condición en la figura 2.3, en donde las líneas rojas nuevamente representan a los 8 partidos de una de las 40,320 posibles fases de eliminación. Notemos que, al no agregar aún la segunda condición, los campeones pueden jugar con los subcampeones de sus mismos grupos, lo cual justamente viene representado con las líneas rojas y que al agregar la segunda condición, se prohíbe.

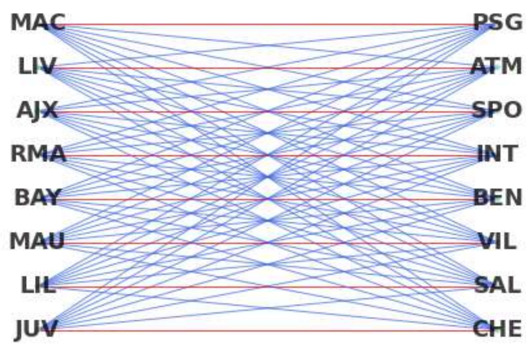

Figura 2.3: Red con todos los partidos posibles que cumplen con la primera condición. En rojo se marca una posible fase de eliminación.

Primera y segunda condición

En este caso únicamente puede jugar un campeón y un subcampeón que no hayan coincidido en la misma fase de grupos anterior. Por lo tanto, ahora las combinaciones se reducen a 14,833 posibles fases de eliminación y podemos ver los partidos válidos en la red de la figura 2.4, con las conexiones rojas representando una de estas posibles fases de eliminación.

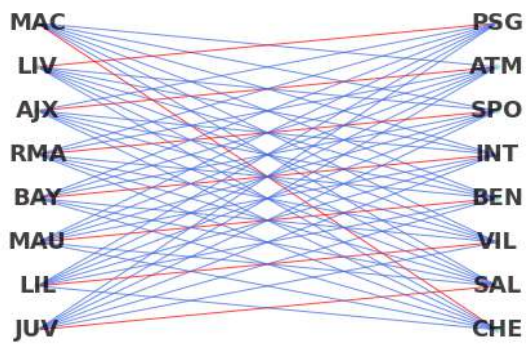

Figura 2.4: Red con todos los partidos posibles que cumplen con la primera y segunda condición. En rojo se marca una posible fase de eliminación

Primera, segunda y tercera condición

Al agregar la tercera condición y no permitir juegos entre equipos de la misma nacionalidad, las posibles fases de eliminación se reducen a la cifra de 4,781 y la red se vuelve más complicada, pues ahora las conexiones dependen de los países de origen de los equipos. En la figura 2.5 vemos la red correspondiente y nuevamente las líneas rojas marcan una fase de eliminación posible.

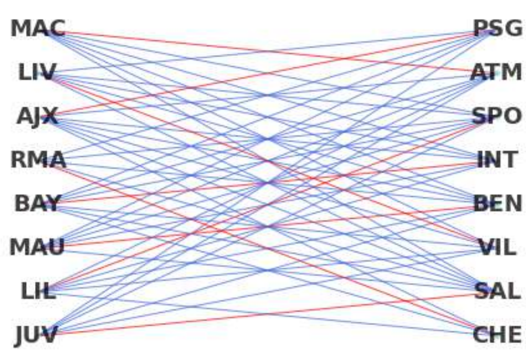

Figura 2.5: Red con todos los partidos posibles que cumplen con la primera, segunda y tercera condición. En rojo se marca una posible fase de eliminación

Aunque tratar el problema con ayuda de las redes hace más sencillo visualizar las posibles fases de eliminación, no será lo que usaremos a lo largo del trabajo, pues encontraremos una forma más útil para los objetivos del mismo.

Lo que haremos será representar a los equipos en una tabla como se muestra en la figura 2.6, donde cada fila corresponde a uno de los 8 equipos campeones (puestos en el orden de los grupos a los que pertenecen) y las columnas a los 8 equipos subcampeones (también en el orden de sus grupos). Luego, cada celda en esta tabla indica un partido entre un equipo campeón y un subcamepón, según la fila y columna en la que se encuentre la celda. Posicionar así a los equipos hace que los partidos posibles ya cumplan con la primera condición de emparejar campeones con subcampeones. Luego, se pintan de color negro las celdas correspondientes a partidos que no pueden jugarse. Los de la diagonal corresponden a la segunda condición, pues no pueden competir equipos de un mismo grupo. Y finalmente, los recuadros negros fuera de la diagonal corresponden a la tercera condición, ya que tampoco pueden competir equipos de la misma nacionalidad.(Por ejemplo, como MCI y CHE pertenecen a Inglaterra, la última celda de la primera fila donde coinciden ambos se encuentra bloqueada.).

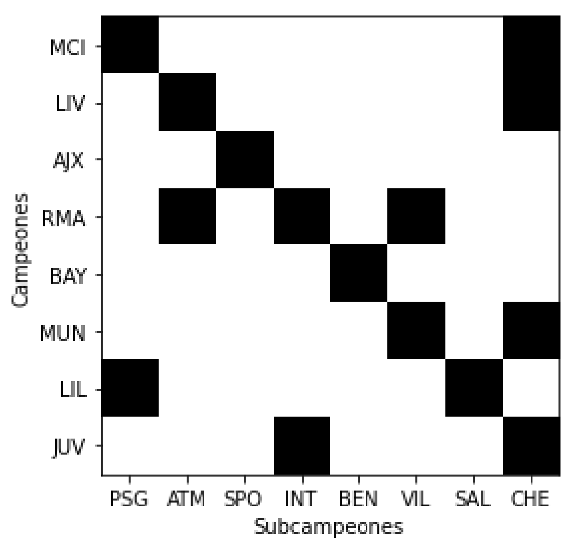

Figura 2.6: Tabla con todos los partidos posibles que cumplen con la primera, segunda y tercera condición.

2.1. Procedimiento de la UEFA

En esta sección describiremos a detalle el procedimiento que sigue la UEFA para hacer el sorteo de los emparejamientos de la fase eliminatoria y luego seguiremos los pasos para el caso particular del sorteo del 2021.

El procedimiento que sigue la UEFA para sortear la fase de eliminación es el siguiente:

1. La UEFA llena una urna con los equipos subcampeones y elige uno al azar.

2. Se usa la tabla 2.6 para revisar cuáles son los posibles campeones que pueden ser rivales del subcampeón elegido, se colocan en una urna y se elige uno al azar para emparejarlo con el subcampeón.

3. Después de hacer los pasos 1 y 2, y conseguir así un partido, se tiene que bloquear toda la columna y toda la fila del subcampeón y campeón en la tabla 2.6, ya que cada equipo se puede elegir una única vez.

4. Se repiten los pasos 1, 2 y 3 pero considerando solamente los subcampeones y campeones que no hayan sido elegidos aún, hasta que todos queden emparejados. Sin embargo, al realizar el paso 2 en cada repetición hay que tener cuidado, ya que hay veces en las que si el subcampeón elegido queda emparejado con algún campeón en particular, esto bloqueará todas las opciones de partidos para otro equipo y hará que el sorteo no se pueda concluir siguiendo las 3 condiciones. Por lo tanto, solamente se tienen que poner en la urna de campeones a los equipos con los que esto no pueda suceder.

Para ver con más claridad lo anterior, sigamos como ejemplo el proceso del sorteo de la UEFA del año 2021.

1. **Primer emparejamiento (BEN-RMA):** Como mencionamos anteriormente, comenzamos llenando una urna con los subcampeones. Cuando se hace la primera elección al azar de subcampeones, sale BEN, el cual, según la figura 2.6, tiene como posibles contrincantes a MCI, LIV, AJX, RMA, MUN, LIL y JUV. Se realiza una

selección al azar entre estos equipos y el primer emparejamiento resulta ser BEN-RMA, el cual se puede ver coloreado de verde en el primer recuadro de la imagen 2.7. Esto provoca que se bloqueé toda la fila y toda la columna del recuadro BEN-RMA, pues ninguno de los dos equipos puede volverse a escoger para ser pareja de otro equipo.

2. **Segundo emparejamiento (VIL-MAC):** Se vuelve a elegir un subcampeón de forma aleatoria (entre los restantes), el cual resulta ser VIL. Este equipo tiene como posibles contrincantes a MAC, LIV, AJX, BAY, LIL y JUV. Al realizar la selección al azar entre estos equipos, el segundo emparejamiento resulta ser VIL-MAC, el cual se observa en verde en el segundo recuadro de la figura 2.7. Al igual que anteriormente, toda la fila y la columna de este recuadro se marcan en negro.

3. **Tercer emparejamiento (ATM-BAY):** Análogamente, se obtiene el subcampeón ATM, que tiene como opciones de contrincantes a AJX, BAY, MAU, LIY y JUV, de los cuales sale seleccionado el BAY. Esto se observa en el tercer recuadro de la figura 2.7. Igual que antes, se elimina como posibilidad su fila y su columna.

4. **Cuarto emparejamiento (SAL, LIV):** El siguiente subcampeón obtenido fue el SAL, que únicamente tiene como opción de contrincantes a LIV, AJX, MAU y JUV. Aleatoriamente sale seleccionado el cuarto emparejamiento SAL-LIV, el cual se observa en el cuarto recuadro de la figura 2.7. Igualmente se elimina su fila y columna.

5. **Quinto emparejamiento (INT, AJX):** Posteriormente se obtiene el subcampeón INT, el cual tiene como posibles contrincantes únicamente a AJX, MAU y LIL. Aleatoriamente se obtiene el quinto partido INT-AJX, cuya representación está en el quinto recuadro de la figura 2.7. Por lo que se bloquea su columna y fila.

6. **Sexto emparejamiento (SPO, JUV):** Para este emparejamiento en específico sucede algo particular. Al azar se obtiene como subcampeón al equipo SPO, el cual

tiene tres opciones de contrincantes: MAU, LIL y JUV. Sin embargo, ¿qué pasaría si como contrincante aleatorio es seleccionado LIL? Lo que sucedería es que CHE se quedaría sin equipo campeón con quien jugar, ya que los otros dos campeones sin pareja MAU y JUV no estaban disponibles para CHE (JUV jugó en su mismo grupo y MAU es de su misma nacionalidad). Por lo tanto, para que no se arruine el sorteo es necesario que LIL no forme parte de las opciones para SPO, por lo que se retira a LIL de la urna de campeones. Si se hace esto, las opciones de SPO entonces son MAU y JUV y aleatoriamente resulta seleccionado el partido SPO-JUV, por lo que los partidos restantes quedan emparejados como PSG-MAU y CHE-LIL.

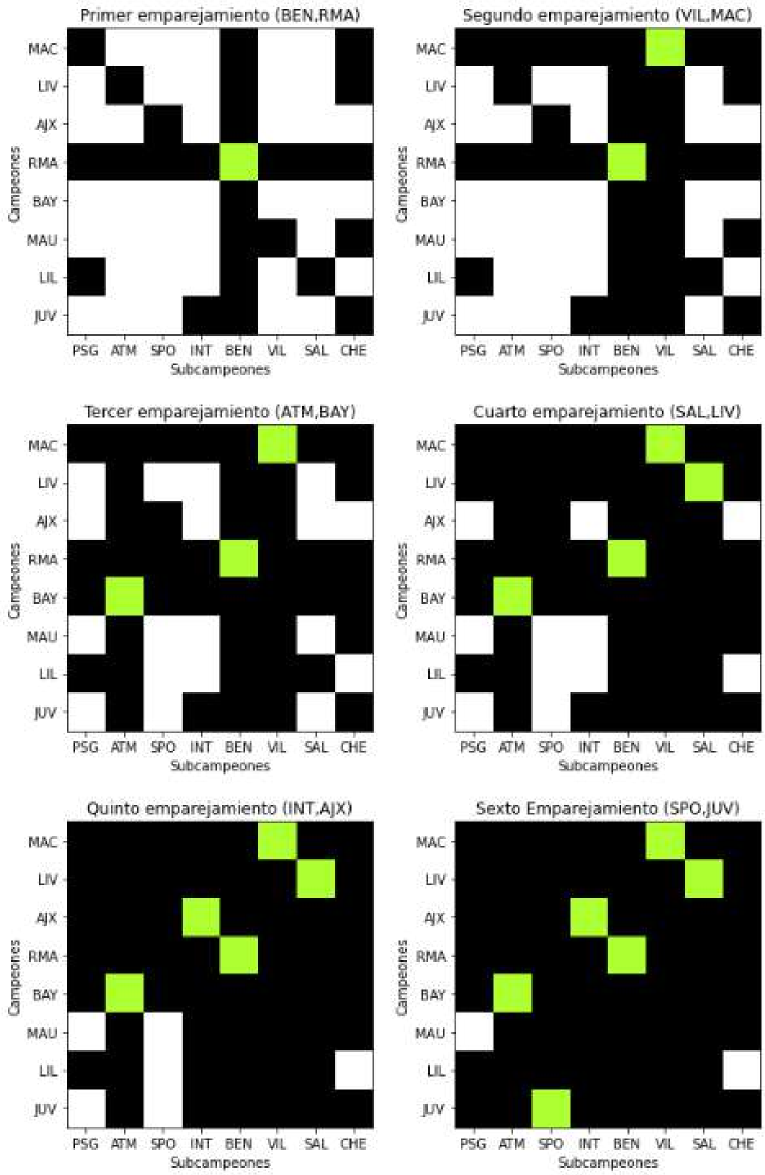

Figura 2.7: Evolución del sorteo de la UEFA, donde los cuadros blancos representan los partidos posibles, los negros representan los partidos que no pueden jugarse y los verdes representan las partidas seleccionadas.

Por lo tanto, al finalizar el procedimiento, el resultado del sorteo es el que se marca en la figura 2.8. Con esto queda claro el procedimiento de sorteo que hace la UEFA. Sin embargo, es importante mencionar que en este sorteo hubo algunos problemas debido a que no se colocaron correctamente los equipos campeones en la urna. Aunque por fortuna esto no cambió el resultado del sorteo, trajo atención al método que usa la UEFA y es la motivación para querer encontrar mejores métodos de sorteo en este trabajo [3]. Finalmente el sorteo fue cancelado debido a esos errores y se tuvo que repetir (usando el mismo procedimiento, pero esta vez sin equivocaciones).

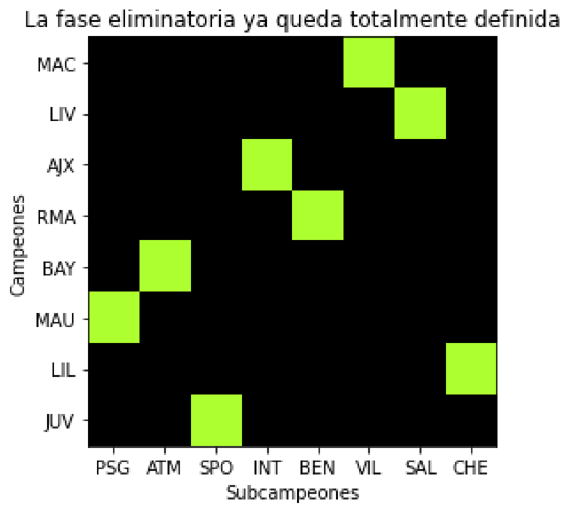

Figura 2.8: Resultado del sorteo de la UEFA 2021

Como se ve en la figura 2.8, la fase de eliminación que se obtiene empareja a los 8 subcampeones con los 8 campeones. Además, se puede observar que cualquier posible fase de eliminación (como la de la figura), corresponde con elegir celdas de la tabla de tal forma que se escoja una por renglón y una por columna. Esto debido a que cada campeón estará emparejado con un y sólo un subcampeón.

2.2. Tabla de Probabilidades

Ahora que sabemos cómo es el sorteo que realiza la UEFA para elegir la fase de eliminación, nos preguntamos qué tan equitativo es este procedimiento. Para ello, primero necesitamos definir aunque sea informalmente a qué nos referimos con que el método sea equitativo o no.

Definición (Método de Sorteo Equitativo): *Decimos que un método de sorteo es* *equitativo si todos los partidos válidos (es decir, en el caso de 2021, los que se muestran en blanco en la figura 2.6), son igual de probables de ser elegidos en el sorteo.*

Entonces, lo que haremos ahora es revisar si el método que realiza la UEFA para elegir la fase de eliminación cumple con esta definición o no. Para revisar esta cuestión, simulamos computacionalmente el sorteo de la UEFA 201,600 veces (lo que corresponde a repetirlo 5 veces para cada uno de las 40,320 órdenes distintos en los que se pueden elegir a los subcampeones). Al terminar las simulaciones, contamos la cantidad de veces que sucedía cada uno de los partidos. Luego, al dividir esta cantidad entre el número total de simulaciones, se obtiene una aproximación de la probabilidad que tiene de salir seleccionado cada partido en el sorteo. En la figura 2.9 se muestran estos resultados, utilizando lo que llamaremos tabla de probabilidades, que se define a continuación.

Definición (Tabla de Probabilidades): *Es una tabla como las que hemos utilizado para representar los partidos posibles, en la que como ya sabemos, cada casilla corresponde con un partido y se marcan en negro los partidos que no son válidos. Además, en los partidos que son válidos se escribe la probabilidad que tienen de ser elegidos cuando se realiza algún método de sorteo.*

Para el caso del método que sigue la UEFA, tras las 201,600 simulaciones del sorteo, se calcularon las probabilidades como se mencionó antes (para cada partido se divide la cantidad de veces que sale elegido entre la cantidad de simulaciones realizadas) y se colocó cada probabilidad en su celda correspondiente, como se ve en la figura 2.9.

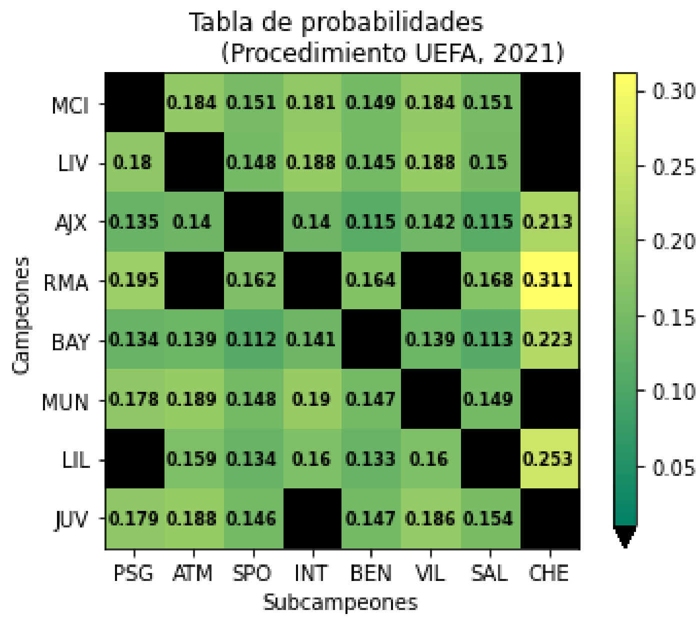

Figura 2.9: Tabla de probabilidades para el método de sorteo de la UEFA.

Por ejemplo, según la tabla de probabilidades de la figura 2.9, la probabilidad con la que se selecciona el partido de MCI contra ATM (primer renglón, segunda columna de la tabla) en el sorteo es de 0.184 (es decir, cada que se hace el sorteo, hay una probabilidad de 18.4 % de que este partido esté entre los seleccionados para los 8vos de final). Así, para cualquier celda en la tabla, el número en ella indica la probabilidad de que en un sorteo, ese partido sea uno de los que salen elegidos, a esto se debe su nombre de ”tabla de probabilidades”.

Propiedades de las tablas de probabilidades

Siguiendo esta interpretación, se puede notar que una tabla de probabilidades debe de cumplir con las siguientes propiedades:

- **Propiedad 1.** Para cada fila de la tabla, la suma de las probabilidades tiene que ser igual a 1. Esto se debe a que en un sorteo, cada campeón será emparejado necesariamente con un (y sólo un) subcampeón y entonces tenemos total certeza de que alguno de los recuadros de cada fila será elegido en el sorteo, por lo que la probabilidad de toda la fila tiene que ser 1.

- **Propiedad 2:** Similarmente, la suma de las probabilidades de cada columna de la tabla tiene que ser igual a 1. La razón es la misma que antes, pero aplicada a las columnas.

- **Propiedad 3:** Todos los números en la tabla de probabilidades se encuentran entre 0 y 1, ya que representan probabilidades. Además, los recuadros negros (que marcan partidos que no son válidos), se interpretarán como si tuvieran un 0 en la tabla, esto debido a que no tienen ninguna probabilidad de ser escogidos, pues son desechados para cumplir con las condiciones.

Con esta tabla se puede conocer qué tan equitativo es o no el método que usa la UEFA para elegir los partidos de la fase de eliminación. Por ejemplo, podemos ver que la celda correspondiente al partido entre RMA y CHE tiene una probabilidad de 0.311. Esto significa que en 31.1 % de los sorteos que se realizan con este método, ese partido sale entre los elegidos para la fase de eliminación. Por otro lado, la probabilidad de que se obtenga el partido entre RMA y SAL es de solamente 0.168, casi la mitad del 0.311 correspondiente a RMA y CHE. Por lo tanto, si CHE es un rival más fuerte que SAL, este método de sorteo indignaría a RMA, que tiene una probabilidad mucho más fuerte de enfrentarse a CHE que a SAL (inversamente, si CHE es un rival más débil que SAL, este método de sorteo beneficiaría mucho a RMA).

Por ello, se puede ver que el resultado del torneo no es perfectamente equitativo, ya que hay partidos con probabilidades de suceder más altas que otros. Debido a estas probabilidades, algunos equipos podrían salir beneficiados o perjudicados. Es decir, el método de sorteo de la UEFA no cumple con la definición de ser un Método de Sorteo Equitativo.

2.2.1. La imposibilidad de construir un Método de Sorteo Equitativo

Como vimos, el método de la UEFA no es un Método de Sorteo Equitativo, ya que no asigna las mismas probabilidades a todos los partidos posibles. Sin embargo, como veremos, para los equipos del 2021 es imposible construir un método de hacer el sorteo tal que sea equitativo.

Para empezar, notamos que si solamente se cumple con las condiciones 1 y 2 que impone la UEFA sobre los partidos (es decir, que los campeones se enfrenten a subcampeones y que no haya partidos entre equipos del mismo grupo, pero sin la condición sobre las nacionalidades), entonces sí se puede encontrar un Método de Sorteo Equitativo. En este caso, la tabla de partidos válidos se vería como la de la figura 2.10 (en la que cambiamos los nombres de los equipos por etiquetas genéricas).

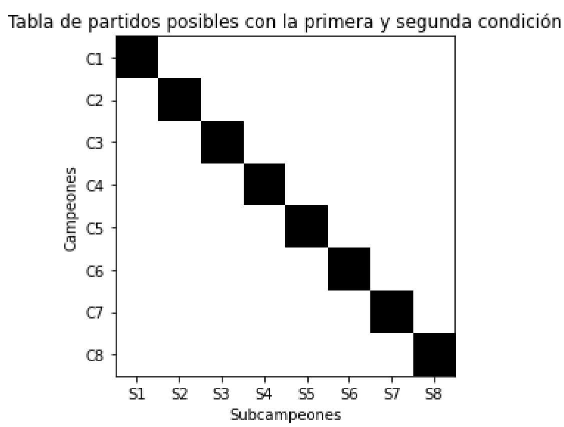

Figura 2.10: Partidos válidos con sólo la primera y segunda condición.

Luego, digamos que se sigue el método de Sorteo de la UEFA en esta situación. Primero se escogería un subcampeón al azar (digamos, sin pérdida de generalidad, que se escogió el S1) y tendría 7 posibles campeones rivales. De estos rivales, se escogería uno, dando así una probabilidad de 1/7 a cada uno de estos encuentros. Es decir, todos los recuadros blancos de la primera columna tendrían probabilidad de 1/7. Luego, como todos los subcampeones son equivalentes en esta tabla (todos tienen el mismo número total de rivales posibles), sus columnas deben de tener las mismas probabilidades que las de S1 (pues no tendría sentido que, por ejemplo, S2 tuviera probabilidades distintas a 1/7, ya que es equivalente a S1 en todos los aspectos excepto en el nombre). Entonces, la tabla de probabilidades que se obtiene con el procedimiento de la UEFA en este caso sería la de la figura 2.11.

Figura 2.11: Tabla de Probabilidades del método de la UEFA cuando sólo se siguen las condiciones 1 y 2.

Entonces, en este caso todos los partidos sí tienen la misma probabilidad, lo que quiere decir que el método es equitativo. Por lo tanto, concluimos que el problema de obtener una tabla de probabilidades que no sea equitativa se debe a la tercera condición.

De hecho, podemos concluir que en realidad para la mayoría de los conjuntos de campeones y subcampeones, si se incluyen las tres condiciones, es imposible encontrar un método de sorteo que sea equitativo. Esto se puede ver de manera simple en el caso del 2021, en el que la tabla de partidos válidos se muestra en la figura 2.6. Para este año, vemos que la cantidad de rivales válidos que tiene CHE (siguiendo las tres condiciones) es de 4 (que son AJX, RMA, BAY y LIL). Como la columna de probabilidades de CHE tiene que sumar 1, si se quiere tener un sorteo equitativo, estos cuatro encuentros tienen que tener cada uno una probabilidad de 1/4. Entonces, la tabla de probabilidades por el momento se vería como se muestra en la figura 2.12.

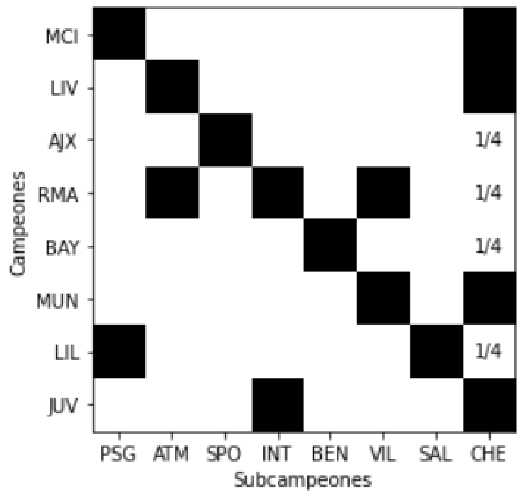

Figura 2.12: Tabla de Probabilidades con la columna de CHE equitativa.

Sin embargo, si nos fijamos ahora en la fila de AJX, notamos que tiene 7 contrincantes posibles (todos los subcampeones excepto SPO). Por lo tanto, como las probabilidades de esta fila (al igual que todas las filas y columnas) tienen que sumar 1, concluimos que las probabilidades de estos encuentros tienen que ser todas 1/7. El problema es que eso nos lleva a una contradicción, pues por un lado dijimos que el partido CHE-AJX tiene que tener probabilidad 1/4 para que la columna de CHE sea equitativa, pero por otro vimos que tiene que tener probabilidad 1/7 para que la fila de AJX sea equitativa. Por lo tanto, concluimos que es imposible asignar probabilidades en esta situación de manera que todos los partidos tengan la misma probabilidad.

Notamos que el problema que llevó a la imposibilidad de construir una tabla de probabilidades equitativa fue que hay equipos que tienen mayor número de rivales posibles que otros (en particular, usamos que CHE tiene 4 rivales posibles mientras que AJX tiene 7). Esta situación es de esperarse cuando se toma en cuenta la tercera condición, pues a menos que se tenga mucha suerte en los equipos que pasan a la fase eliminatoria, lo normal es que la cantidad de equipos de cada país sea muy variada. Por lo tanto, concluimos que en la (gran) mayoría de los casos, es imposible conseguir un Método de

Sorteo Equitativo.

2.3. Medir la equidad

Como se vio en la sección anterior, el método de sorteo que realiza la UEFA no es equitativo en las probabilidades de los partidos, pues algunos partidos son más probables que otros, como se puede ver claramente en la tabla de probabilidades de la figura 2.9. Vimos que en realidad esto es de esperarse, pues en la es imposible diseñar un sorteo que sea equitativo.

Sin embargo, aunque no haya métodos de sorteo equitativos, sin duda habrá métodos que sean más equitativos que otros. Diremos que un método de sorteo se acerca a ser equitativo si en general las probabilidades de todos los partidos válidos tienen valores similares. Esta definición es algo informal y nos gustaría encontrar una forma más clara de cuantificar qué tan equitativo es un método de sorteo. En este capítulo propondremos algunos criterios para determinar esto.

Buscaremos entonces criterios para cuantificar numéricamente qué tan equitativo es un método de sorteo a partir de revisar la tabla de probabilidades. En general, buscamos una cantidad que nos diga qué tan dispersos son los valores en la tabla de probabilidades, pues mientras menos dispersas sean estas probabilidades, mayor equidad entre las probabilidades de los partidos.

La dificultad aquí es que hay varias formas distintas de cuantificar esta dispersión en las probabilidades, y no está claro cual pueda ser más válida. Por lo tanto, a continuación enlistaremos varios posibles criterios distintos que se pueden utilizar para medir la equidad de una tabla de probabilidades. A lo largo del trabajo usaremos todos estos criterios para cuantificar la equidad.

2.3.1. Criterio de suma de varianzas

Antes de poder definir este criterio, es necesario primero definir el concepto de **varianza**. La varianza es una cantidad que se usa para cuantificar qué tanta dispersión hay en un conjunto de datos. Por ejemplo, digamos que tenemos un conjunto de datos numéricos que son $\{x_1, x_2, x_3, \cdots, x_N\}$ (estos datos pueden representar cualquier cosa, tanto medidas de una cantidad física como la estatura de n personas, como los números que se ponen en la tabla de probabilidades como la de la figura 2.9).

La varianza es una medida de qué tanta dispersión hay en estos datos. Para empezar, se calcula el promedio de los datos, que se denota por $\bar{x}$. Luego, para el dato x_i por ejemplo, se puede cuantificar qué tan disperso está al calcular la distancia a la que se encuentra de la media, lo que se haría como $x_i - \bar{x}$. Por ejemplo, si se trata de estaturas y el promedio es $\bar{x} = 1.70m$ y $x_i = 2.05m$, la distancia a la media sería de $0.35m$, lo cual en este caso es bastante grande e indica que x_i es un dato bastante disperso. Mientras que si un dato es $x_j = 1.68m$, la distancia a la media sería de $-0.02m$, que en este caso es bastante baja e indica que x_j no se encuentra tan alejado de la media.

Luego, para medir la dispersión total de los datos, se suman estas cantidades $x_i - \bar{x}$ para todos los datos registrados. Sin embargo, antes de hacer esto, estas cantidades se elevan al cuadrado; esto se hace para que estas cantidades sean números positivos, y así al sumar todas, no se anulen cantidades positivas con negativas. Entonces, una forma de medir la dispersión de los datos es calcular la suma de $(x_i - \bar{x})^2$ sobre todos los datos. Sin embargo, para definir la varianza es común dividir esta suma entre la cantidad total de datos N, para así estandarizarla y que no dependa de la cantidad de datos (en realidad, por razones en las que no es necesario entrar en detalle en este trabajo, la varianza de un conjunto de datos se suele definir dividiendo por $N - 1$, en vez de N). Entonces, la definición de varianza queda como [**?**]:

$$V(\{x_1, x_2, x_3, \cdots, x_N\}) := \frac{1}{N-1}\sum_{i=1}^{n}(x_i - \bar{x})^2. \tag{2.1}$$

Por todo lo que se discutió para llegar a esta definición, podemos ver que la varianza es una buena forma de medir la dispersión de un conjunto de datos. Esto es justo lo que buscábamos, ya que así podremos cuantificar qué tanta dispersión hay en una tabla de probabilidades y medir así qué tan equitativa es.

Entonces, ¿qué tiene que ver la varianza con poder medir la equidad? Sabiendo la interpretación de varianza, sabemos que ésta va a ser más pequeña mientras las probabilidades de la tabla sean más cercanas, puesto que no habrá tanta dispersión entre ellas. Por ejemplo, si tuviéramos una tabla de probabilidades totalmente equitativa (con cada celda con igual probabilidad de suceder), la varianza sería la mínima posible (que es 0), puesto que la distancia entre cada probabilidad y la media sería cero. Lo anterior nos da una clara forma de medir equidad.

En otras palabras, una forma de poder medir la equidad en una tabla de probabilidad es con la varianza, donde a menor varianza, mayor equidad. Por lo tanto, un valor pequeño de la varianza en una tabla de probabilidad indica un sorteo bastante equitativo, en el que los partidos tienen probabilidades similares de suceder.

Ya sabiendo la interpretación de varianza podemos hablar sobre el **criterio de suma de varianzas**. Este criterio consiste en obtener con la ecuación 2.1 las varianzas de cada renglón y cada columna de la tabla de probabilidades (considerando a las probabilidades del renglón o columna como los datos $x_1, x_2, \cdots, x_N$, omitiendo los partidos imposibles) y posteriormente sumarlas.

De esta forma, la varianza de una columna o fila nos dice qué tan equitativas son las probabilidades de los partidos que puede tener el subcampeón o campeón correspondiente.

Y luego, se suman todas para considerar a todos los subcampeones y campeones. Como se ha mencionado, mientras menor sea este valor, indica más equidad en la tabla.

2.3.2. Criterio de varianza total

Este criterio consiste en obtener, también con ayuda de la ecuación 2.1, la varianza de toda la tabla de probabilidad, no yendo de renglón en renglón y columna en columna, sino considerando a todos los posibles partidos válidos. Es decir, en este caso los datos $x_1, x_2, \cdots, x_N$ sobre los que se quiere medir la varianza son todos los datos de la tabla de probabilidades.
Al igual que el anterior, este criterio ayuda a medir la equidad gracias a la relación que hay entre ésta y la varianza. Una menor varianza total indica probabilidades cercanas entre sí y por tanto una mayor equidad.

2.3.3. Criterio de máximo

Como lo dice su nombre, este criterio se basa en encontrar el valor de probabilidad máxima de la tabla de probabilidades.

Hacer esto es una forma de medir qué tan equitativa es la tabla de probabilidades, ya que un valor alto indica que hay un partido con mucha probabilidad de suceder, lo cual es una señal de que la tabla no es muy equitativa y podría ser un partido que beneficia o perjudica mucho a uno de los equipos. Por otro lado, si el máximo de la tabla es pequeño, indica que no hay partidos demasiado probables, y como las probabilidades de las columnas y filas tienen que sumar 1, eso significa que tampoco pueden haber probabilidades demasiado bajas.

Por lo tanto, calcular el máximo de la tabla sirve como una buena forma de cuantificar qué tan equitativa es. Por ejemplo, una tabla totalmente equitativa, en la que todas las celdas tengan la misma probabilidad tendrá el máximo lo más pequeño posible. Esto porque no sería posible tener un máximo menor que el de una tabla totalmente equitativa, pues disminuir aunque sea una de las probabilidades hace que otras aumenten, ya que la suma de las columnas y filas tiene que seguir siendo 1.

2.3.4. Criterio de rango

Este criterio consiste en medir la diferencia que hay entre la probabilidad máxima y la probabilidad mínima de la tabla de probabilidades.

De esta forma, se puede cuantificar qué tan grande es la distancia entre el partido más probable y el menos probable, Si esta distancia fuera muy grande, indicaría una gran inequidad en las probabilidades; por otro lado, un rango pequeño indica que todas las probabilidades son parecidas, pues se encuentran dentro de ese rango. Por ejemplo, una tabla totalmente equitativa tendrá un rango de 0, pues todas las probabilidades son iguales y en particular el máximo y el mínimo .

2.3.5. Criterio de Gini

Con este criterio y los dos siguientes, pasamos a criterios más complejos y que se suelen usar en economía para medir la desigualdad económica entre individuos. A pesar de que se usen generalmente en economía, estos criterios se pueden utilizar para medir la inequidad de cualquier serie de datos. En particular, los podemos usar para medir la inequidad de las probabilidades en una tabla de probabilidades y así cuantificar qué tan equitativo es un método de sorteao de la fase eliminatoria de la Champions [**?**].

Al igual que como se hizo para definir la varianza, supongamos que tenemos una serie de datos $\{x_1, x_2, x_3, \cdots, x_N\}$ y veamos cómo se usa el criterio de Gini para medir su inequidad. Para empezar, si queremos medir la distancia entre una pareja de estos datos (digamos x_i y x_j), lo haríamos como $|x_i - x_j|$ (se pone el valor absoluto para asegurarnos de conseguir un valor positivo y que se pueda interpretar efectivamente como la distancia entre los datos).

Luego, una forma natural de medir la inequidad en toda la lista de datos es sumar esta diferencia para todas las parejas posibles. Así es básicamente como se define el índice de Gini, sólo que posteriormente se divide entre $2N^2\bar{x}$ (donde N es la cantidad de datos y $\bar{x}$ el promedio), esto se hace para estandarizar el valor del índice y asegurarse de que siempre se encuentre entre 0 y 1. Es decir [**?**],

$$G = \frac{1}{2N^2\bar{x}} \sum_{i=1}^{N} \sum_{j=1}^{N} |x_i - x_j|.$$

El índice de Gini da como resultado siempre un valor entre 0 y 1 y su interpretación es que mientras más cerca esté el índice a 1, mayor es la desigualdad de los datos y mientras más se acerque este índice a 0, es menor la desigualdad. El valor de 0 se consigue solamente cuando todos los datos son iguales, debido a que entonces todas las distancias entre parejas de datos son 0. Aunque normalmente se utiliza en términos económicos, el índice de Gini puede aplicarse para cualquier distribución a la que se le quiera cuantificar la inequidad, por ejemplo, para la inequidad de una tabla de probabilidades.

Por lo tanto, calcularemos el criterio de Gini para las probabilidades en la tabla de probabilidades y obtendremos así una medida de qué tan variadas son. Mientras mayor sea el índice, indicará una menor equidad probabilística y mientras menor sea el índice, indicará una mayor equidad.

2.3.6. Criterio de Theil

El índice de Theil consiste en otra métrica de desigualdad que se utiliza normalmente en economía. Sin embargo, al igual que para el índice de Gini, lo podemos aplicar a nuestro problema. Dado un conjunto de datos $\{x_1, x_2, \cdots, x_N\}$, este índice se calcula como [**?**]:

$$T = \frac{1}{N}\sum_{i=1}^{N}\frac{x_i}{\bar{x}}\ln\left(\frac{x_i}{\bar{x}}\right),$$

con $\bar{x}$ el promedio de los datos.

A diferencia de los otros criterios que hemos visto hasta ahora, éste no tiene una interpretación tan directa e intuitiva. Sin entrar en mucho detalle, pues se requerirían bastantes conceptos para hacerlo, el índice de Theil se relaciona con la entropía del conjunto de datos, que es una forma de medir la diversidad o falta de ella en los datos [**?**]. Un valor alto de este índice indica gran variedad en los datos, mientras que un valor cercano a 0 indica que los datos son muy similares entre sí. El 0 se obtiene cuando todos los datos son exactamente iguales.

Al igual que en el criterio de Gini, aplicaremos esta medida a todos los datos de la tabla de probabilidades y así cuantificaremos su dispersión. Un resultado bajo indicará que las probabilidades en la tabla son muy parecidas entre sí y se trata de un sorteo bastante equitativo, mientras que uno alto indirá desigualdades en el sorteo.

2.3.7. Criterio de Atkinson

Similarmente a los últimos dos criterios, el criterio de Atkinson se usa generalmente en economía para medir la desigualdad entre una serie de datos $\{x_1, x_2, \cdots, x_N\}$. Esta cantidad se define matemáticamente como [**?**]:

$$A = 1 - \frac{1}{\bar{x}}\left(\Pi_{i=1}^{N} x_i\right)^{1/N},$$

donde $\bar{x}$ indica el promedio de los datos y $\Pi_{i=1}^{N} x_i$ denota la multiplicación de los valores $x_1, x_2, \cdots, x_N$.

Aunque no tiene una interpretación tan clara y directa como otros de los criterios, es fácil ver que es una cantidad que se minimiza conforme más cerca estén entre sí los datos x_i.

Esto se puede ver fácilmente en un ejemplo. Consideramos el caso de solamente dos datos x_1, x_2 y digamos que los dos datos tienen un promedio de $\bar{x} = 5$ (o lo que es lo mismo, que $x_1 + x_2 = 10$). Entonces, se puede ver que el producto $\Pi_{i=1}^{2} x_i = x_1 x_2$ será lo más grande posible cuando $x_1 = x_2 = 5$, en cuyo caso vale 25. Sin embargo, si $x_1 = 9$ y $x_2 = 1$, el producto será 9. En general es fácil convencerse de que mientras más cercanos entre sí sean los dos números, mayor será su producto, mientras que si uno es grande y el otro chiquito, el producto no será muy grande.

Esto mismo se cumple cuando se considera el producto $\Pi_{i=1}^{N} x_i$ para más de 2 datos. Mientras más cercanos sean entre sí los datos, mayor será este producto y mientras más alejados, será menor. Entonces, como en la definición del criterio de Atkinson estamos tomando 1 menos este producto (con la raíz N-ésima y la división por $\bar{x}$, lo cual no cambia este argumento), maximizar al producto nos lleva a valores menores de A. Por lo tanto, concluimos que el criterio de Atkinson es más pequeño conforme más cercanos entre sí sean los datos.

Entonces, para medir la equidad de la tabla de probabilidades, podemos calcularle el criterio de Atkinson a los datos de dicha tabla. Valores altos indicarán una desiguldad en las probabilidades, mientras que valores bajos indicarán un sorteo que es equitativo.

Capítulo 3

Métodos de sorteo

En la sección anterior se presentó el método que utiliza la UEFA para sortear los partidos de la fase eliminatoria de la Champions y se observó que tiene un problema debido a la alta diferencia en probabilidades que tienen algunos partidos respecto a otros. Además, se presentaron varios criterios para cuantificar estas diferencias y determinar qué tan equitativa es una tabla de probabilidades.

En esta sección empezaremos presentando varios métodos para realizar los sorteos de los partidos de la fase eliminatoria. Iniciamos de nuevo con el método que utiliza la UEFA y lo analizamos a más detalle. Posteriormente, presentaremos métodos alternativos y los analizamos de igual manera. Además, al final de la sección compararemos la equidad de cada método según cada uno de los criterios de equidad definidos en el capítulo anterior y escogeremos como favorito al método que minimice a los criterios lo más posible.

3.1. UEFA

Este método para sortear la fase de eliminación ya fue presentado y explicado en la sección 2.1, por lo que no lo repetiremos aquí. Sin embargo, mencionaremos algunas características importantes del método y lo podremos comparar con otros métodos de sorteo propuestos más adelante.

- **Tabla de Probabilidades:** Como se menciona en la sección 2.2, la tabla de probabilidades que se obtiene con este método de sorteo es la siguiente:

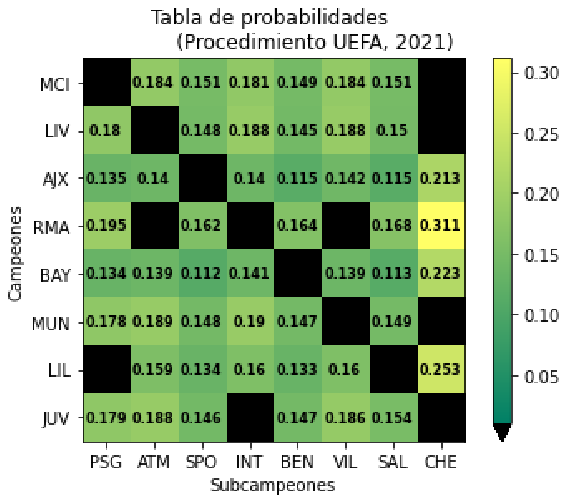

Figura 3.1: Tabla de probabilidades para el método de sorteo de la UEFA.

- **Evaluación de los criterios:** A esta tabla de probabilidades podemos calcularle los criterios que habíamos definido anteriormente para cuantificar la equidad. Hicimos esto computacionalmente (el proceso se encuentra en el código entregado) y obtuvimos los siguientes resultados para cada uno de los criterios evaluados en esta tabla:

Suma de varianzas	Varianza total	Rango	Máximo	Gini	Theil	Atkinson
0.015235	0.001288	0.199653	0.311384	0.110377	0.021733	0.010500

Figura 3.2: Resultado de los criterios aplicados a la tabla de probabilidades del método de la UEFA.

Por ahora estos resultados no nos dicen mucho, pues no tenemos con qué compararlos y no hay forma de saber si son muy grandes o pequeños. Sin embargo, para los métodos siguientes, usaremos los resultados de este método como punto de comparación para cuantificar qué tan superiores o inferiores son.

- **Requerimientos Adicionales:** Es importante notar que este método requiere del apoyo de una computadora que analice los posibles resultados a cada paso. Esto debido a que se necesita un programa que se asegure de que siga existiendo la posibilidad de completar el sorteo y que no se escoja un partido que haga imposible continuar con la selección, tal como se explica en la seccion 2.1.

- **Espectacularidad:** Una de las mayores ventajas de este método es la gran espectacularidad que ofrece durante el procedimiento del sorteo. Como se mencionó en la sección 2.1, en el sorteo se van seleccionando equipos de una urna. Esto crea gran expectativa en los aficionados y jugadores, ya que se tiene la emoción de ver como los organizadores sacan un nombre de la urna que puede decidir completamente el futuro del equipo en el torneo. Es por esta razón que este método se adapta muy bien a ser realizado ante una audiencia en vivo y televisado para todos los aficionados.

- **Sencillez:** La forma en que se realiza el sorteo puede ser un poco confusa para quienes no saben del tema debido a las condiciones que deben de cumplir los emparejamientos para ser válidos. Más aún, el sorteo se vuelve complicado debido a la constante necesidad de revisar que un emparejamiento escogido no bloqueé totalmente las posibilidades de partidos que pueda tener un equipo que no ha sido seleccionado aún. Como se vio en la sección 2.1, intentar evitar este problema puede dar lugar a errores como los que se tuvieron durante el sorteo del 2021.

3.2. Posibles Soluciones Alternativas

A partir de ahora, se presentan varios métodos alternativos para realizar el sorteo de los emparejamientos de la fase de eliminación. Se proponen varios métodos distintos y que se dividen en 5 tipos distintos:

- Un método similar al de la UEFA pero adaptando las probabilidades.
- Un método que ofrece las mismas probabilidades a todas las fases de eliminación posibles.
- Varios métodos que minimizan a los criterios de equidad.
- Métodos que minimizan un criterio de equidad que depende de un parámetro.
- Método que inicia con una tabla de probabilidades que minimiza el rango y luego minimiza la varianza.

Para cada uno de los métodos se obtendrá la tabla de probabilidades, se medirán los criterios de equidad y se propondrá la forma de realizar el sorteo. Al final compararemos todos los métodos entre sí.

3.2.1. Método de la UEFA Pesado

Primero se propone un método similar al de la UEFA pero adaptado para intentar que las probabilidades finales de la tabla sean más equilibradas. A grandes razgos, este método sigue los mismos pasos que el de la UEFA pero intenta compensar los partidos que en el método de la UEFA tienen probabilidades bajas al hacer más probable su elección.

Para que quede claro este procedimiento, empezamos presentando un ejemplo de cómo se podría realizar. Al igual que siempre, empezamos con la tabla de todos los partidos posibles según las tres condiciones, dicha tabla está dada en la figura 3.3.

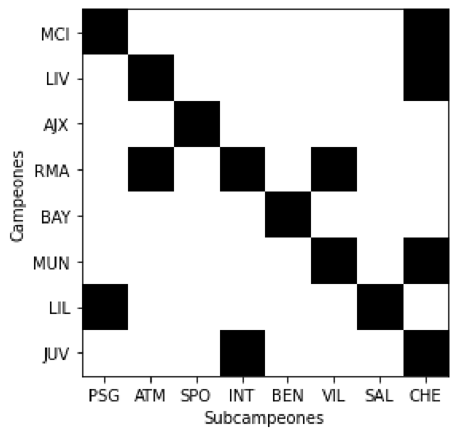

Figura 3.3: Tabla de partidos válidos.

Al igual que en el procedimiento de la UEFA, se empieza seleccionando uno de los subcampeones al azar. Digamos que queda seleccionado el CHE. Luego, se revisa a qué campeones puede enfrentarse (es decir, que sean de otro país y de otro grupo y que en caso de ser elegidos, no bloqueen el resto del sorteo). En este caso, los campeones posibles son AJX, RMA, BAY y LIL y se tiene que escoger alguno al azar. En el procedimiento que hace la UEFA, cada uno de estos equipos tendría la misma probabilidad de ser elegido, sin embargo, en este procedimiento modificaremos esta probabilidad de selección.

Lo que haremos primero es ver las probabilidades de cada uno de estos partidos según la tabla de probabilidades que resulta del procedimiento original de la UEFA, que están en la figura 3.1. Estas probabilidades son: 0.212 para (CHE-AJX), 0.312 para (CHE-RMA), 0.216 para (CHE-BAY) y 0.254 para (CHE-LIL). Para compensar a los partidos que tienen baja probabilidad en el método de la UEFA, ahora elegiremos a los campeones con una probabilidad inversamente proporcional a las de la UEFA. Es decir, elegiremos a AJX con una probabilidad proporcional a $1/0.212 = 4.717$, a RMA con probabilidad proporcional a $1/0.312 = 3.205$, a BAY con $1/0.216 = 4.63$ y a LIL con $1/0.254 = 3.937$. De esta forma, ahora AJX tendrá una mayor probabilidad de ser elegido y RMA menor, lo cual podría servir para compensar las probabilidades desiguales que tienen estos partidos en el método de UEFA.

En conclusión, el procedimiento es el siguiente:

1. Obtener la tabla de probabilidades que resulta de utilizar el método de la UEFA, para así poder ver qué partidos son más probables que otros. Esta tabla es la de la figura 3.1.

2. Escoger un equipo subcampeón al azar.

3. Revisar qué equipos campeones pueden jugar contra el subcampeón elegido, es decir, que cumpla con las condiciones de ser de un país distinto, de un grupo distinto y que de ser elegido, no bloquee el resto del sorteo.

4. Finalmente, se seleccionará al azar uno de estos campeones. Sin embargo, en vez de asignar una probabilidad equitativa a todos los campeones, se asegurará que los partidos que eran menos probables según el método de la UEFA tengan ahora mayor probabilidad de suceder. Para hacer esto, se asignará a cada campeón una probabilidad de ser elegido inversamente proporcional a la probabilidad que tiene su partido en la tabla de probabilidades de la UEFA.

5. Se reptien los pasos 2,3,4 hasta completar todo el emparejamiento.

Ahora que hemos definido este método, vamos a analizarlo, obtener su tabla de probabilidades y revisar sus características para ver si es superior al método de la UEFA.

- **Tabla de Probabilidades:** Para obtenerla, al igual que en el método de sorteo de la UEFA, se simula varias veces el procedimiento y se cuenta el número de veces que aparece cada partido en la fase eliminatoria. El método se repitió $201,600$ veces, realizando 5 veces cada forma de ordenar la elección de los subcampeones. Finalmente, dividiendo el número de veces que aparece cada partido en el resultado final entre la cantidad de simulaciones, se obtiene la tabla de probabilidades de la figura 3.4.

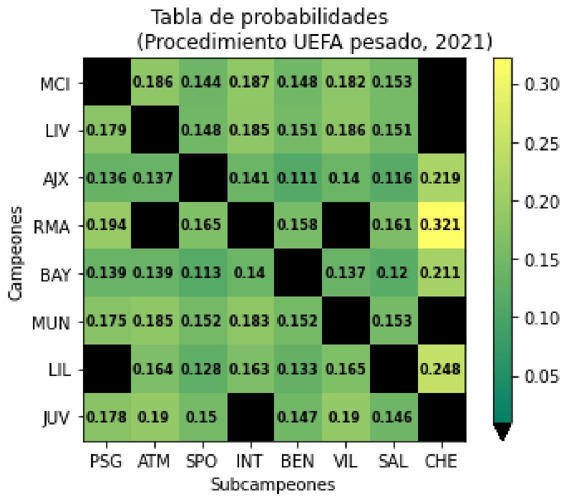

Figura 3.4: Tabla de probabilidades utilizando el método de UEFA pesado.

- **Criterios:** A simple vista, la tabla obtenida se parece mucho a la de la figura 3.1 del procedimiento de la UEFA. Sin embargo, para determinar si es verdaderamente un mejor resultado o no, le calculamos los criterios y obtenemos lo siguiente:

Suma de varianzas	Varianza total	Rango	Máximo	Gini	Theil	Atkinson
0.015995	0.001307	0.205531	0.318998	0.110014	0.021805	0.010482

Figura 3.5: Resultado de los criterios aplicados a la tabla de probabilidades del método de la UEFA pesado.

Al comparar con los criterios aplicados al método de la UEFA dados en la figura 3.2, podemos ver que los resultados son en general mayores, lo que indica que la tabla de probabilidades es menos equitativa. Por lo tanto, podemos ver que intentar compensar las probabilidades no ayuda y es más conveniente hacer directamente el método de la UEFA, en el cual cada campeón tiene la misma probabilidad. La comparación completa de estos resultados con los de los otros métodos y el de la UEFA se mostrarán más adelante en el trabajo en la sección 3.3.

- **Requerimientos Adicionales:** Al igual que en el sorteo de la UEFA, este método requiere del apoyo de una computadora, ya que en cada paso es necesario revisar qué campeones se pueden elegir para evitar que queden equipos sin contrincantes posibles. Además, se necesita una computadora para realizar la simulación del método de la UEFA antes de empezar (y poder así compensar las probabilidades) y también para seleccionar a los campeones con las probabilidades no uniformes.

- **Espectacularidad:** Similar al caso de la UEFA, para la elección del subcampeón se utiliza una urna, lo cual hace emocionante al proceso. Sin embargo, elegir al campeón es más complicado, ya que no se pueden poner simplemente en la urna, puesto que en el caso de la UEFA las probabilidades eran iguales para todos los campeones y para este método se requiere pesar las probabilidades de forma no uniforme. Por lo tanto, la elección del campeón se debe de hacer por computadora y ya no tendrá la misma emoción que en el sorteo de la UEFA.

- **Sencillez:** Este método es incluso más complicado que el de la UEFA, por la complejidad adicional de tener que asignar probabilidades no uniformes a los campeones.

En conclusión, aunque parezca una buena idea a primera vista, este método no es conveniente y es inferior al de la UEFA en todos los sentidos.

3.2.2. Método Uniforme

Para entender este método, es necesario recordar que dadas las condiciones que se tienen que cumplir para los emparejamientos, hay 4,781 posibles resultados para la fase eliminatoria. Es decir, al final del sorteo, sin importar el método que se utilice, obtendremos alguno de estos 4,781 torneos posibles.

El método del sorteo uniforme consiste en enlistar esas 4,781 posibilidades y escoger una de ellas al azar (donde todas tienen la misma posibilidad de ser escogidas). Una vez que se hace esta elección, ya quedan decididos los 8 partidos que cumplen con las condiciones.

De esta forma, nos aseguramos que todas las posibles fases de eliminatoria que son válidas tienen la misma probabilidad de suceder. Sin embargo, eso no significa que todos los partidos tienen la misma probabilidad, ya que hay partidos que se presentan en un mayor porcentaje de fases de eliminatoria posibles que otros.

- **Tabla de Probabilidades:** Para empezar, obtenemos la tabla de probabilidades de los partidos seleccionados por medio de este método de sorteo. Para hacerlo, como en este método todos los torneos posibles tienen la misma probabilidad de suceder, simplemente contamos cuantas veces aparece cada partido en los $4,781$ torneos posibles y dividimos esta cantidad por 4781. El resultado de hacer esto se muestra en la tabla de probabilidades de la figura 3.6.

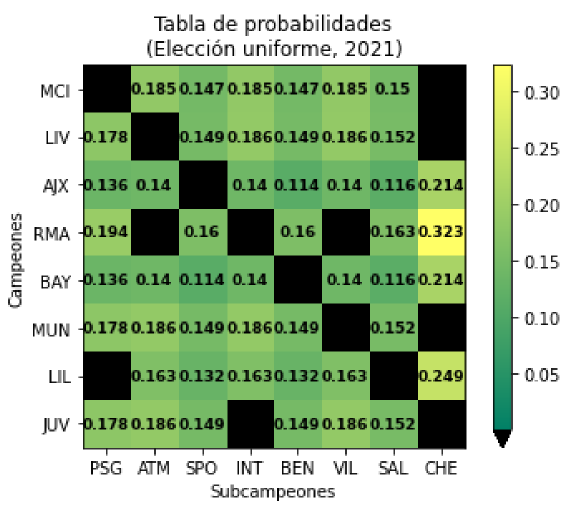

Figura 3.6: Tabla de probabilidades utilizando el método uniforme.

- **Criterios de Equidad:** Le medimos los criterios de equidad a esta tabla de probabilidades, al igual que se ha hecho con los otros métodos.

Suma de varianzas	Varianza total	Rango	Máximo	Gini	Theil	Atkinson
0.016060	0.001304	0.209580	0.323154	0.108255	0.021686	0.010418

Figura 3.7: Resultado de los criterios utilizando el método uniforme.

Al igual que antes, dejaremos la comparación con otros métodos para más adelante.

- **Espectacularidad:** El sorteo consiste en elegir al azar uno de los $4,781$ torneos posibles, el cual se podría elegir de una urna (aunque no sería muy práctico por la cantidad de bolas que se tendrían que poner en ella) o computacionalmente eligiendo al azar un número entre 1 y $4,781$. La revelación de los resultados no es tan emocionante como en el sorteo de la UEFA, pues tras la elección, se conocen todos los partidos a la vez. Sin embargo, una forma de darle más espectacularidad a este método es si se van leyendo los partidos del torneo uno por uno, dando tiempo para la emoción y simulando la sorpresa de los resultados.

- **Requerimientos Adicionales:** Para este proceso se necesita antes usar una computadora para calcular todos los torneos posibles y para luego elegir uno al azar el día del sorteo.

- **Sencillez:** El método es muy sencillo, es fácil de explicar y tiene la ventaja de que a todas las posibles fases de eliminación se les asigna la misma probabilidad.

3.2.3. Métodos que minimizan un criterio

Hasta ahora hemos definido métodos que se basan en intentar hacer un sorteo de la forma más equitativa posible. Una vez definido el método, obtenemos su tabla de probabilidades y le medimos los distintos criterios, esperando que los resultados sean lo más bajos posibles para que indiquen una tabla equitativa.

Sin embargo, en esta sección tendremos un enfoque inverso al anterior. Empezaremos escogiendo uno de los criterios, buscaremos entre todas las posibles tablas de probabilidades alguna que minimice dicho criterio y finalmente construiremos un sorteo que dé resultados que sigan dicha tabla de probabilidades.

Presentaremos varios métodos dependiendo de qué criterio sea el que buscamos minimizar. Sin embargo, para todos los criterios se seguirá el mismo procedimiento, el cual se detalla a continuación.

3.2.3.1. Procedimiento para encontrar la tabla que minimiza a un criterio

1. Consideraremos a cada valor en la tabla de probabilidades como una variable libre (excepto los correspondientes a partidos que no son válidos, es decir, asignamos una variable para cada una de las celdas blancas en la figura 2.6). Para el caso del torneo del 2021, estas variables se ven como en la figura 3.8.

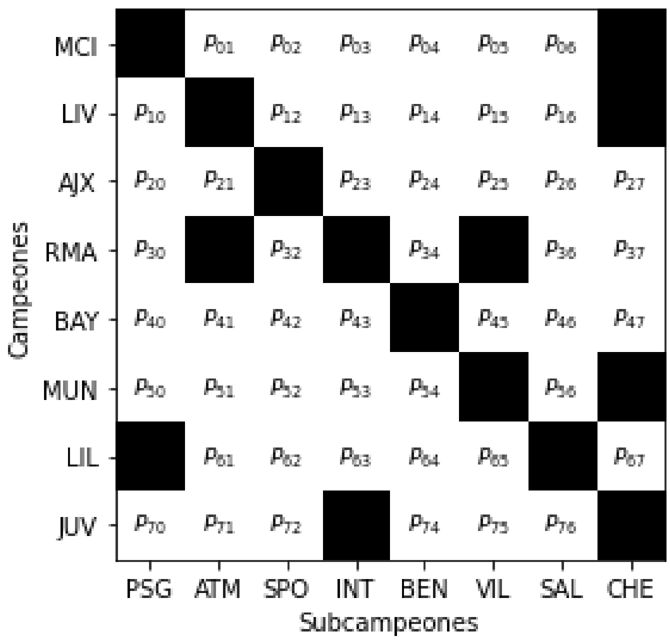

Figura 3.8: Tabla con las variables para cada uno de los partidos posibles.

2. Planteamos las condiciones que deben de cumplir estas variables para que formen una tabla de probabilidades válida. Estas condiciones se mencionan en la sección 2.2, que son:

   - Cada valor tiene que ser un número entre 0 y 1, ya que tiene que representar una probabilidad. Es decir, $0 \leq p_{ij} \leq 1$.
   - La suma de los valores en cada renglón tiene que ser igual a 1. Es decir, para todo i se debe de cumplir que $\sum_j p_{ij} = 1$.

- La suma de los valores en cada columna tiene que ser igual a 1. Es decir, para todo j se debe de cumplir que $\sum_i p_{ij} = 1$.

3. Luego, se escribe matemáticamente el criterio que se quiere minimizar en términos de estas variables. Como todos los criterios son a final de cuentas funciones de los elementos de la tabla, podemos escribir el criterio como función de las p_{ij}.

4. Finalmente, ya teniendo la expresión que se busca minimizar, las variables y las condiciones que deben de cumplir, se utiliza algún método de optimización para encontrar los valores de las variables que minimicen al criterio y cumplan las condiciones. En este caso, se utilizó el método SLSQP que incluye la librería scipy de Python (como se puede ver en el código). Éste es uno de muchos métodos que se pueden utilizar para minimizar una función bajo ciertas condiciones que deban de cumplir las variables.

De esta forma, se encuentran computacionalmente los valores que deben de tener las probabilidades p_{ij} de los partidos de manera que minimicen el criterio que nos interesa y que cumplan con las condiciones necesarias para formar una tabla de probabilidades. Los resultados encontrados pueden variar según el método de minimización utilizado y los criterios que se utilicen para detener la búsqueda.

3.2.3.2. Crear un sorteo a partir de la tabla de probabilidades

Una vez que consigamos la tabla de probabilidades que minimice a cierto criterio, nos preguntamos de qué manera podemos realizar un sorteo cuya tabla de probabilidades sea la obtenida. Es decir, en general nos interesa encontrar la forma de definir un sorteo a partir de conocer la tabla de probabilidades que queremos que tenga.

Para hacerlo, primero recordamos de la sección 2.2 que una tabla de probabilidades debe de cumplir que todos sus valores sean números entre 0 y 1, que los renglones sumen 1 y que las columnas también sumen 1. En general, a las matrices que cumplen con estas condiciones se les conoce como **matrices bi-estocásticas** [2]. Es decir, las tablas de probabilidades con las que hemos estado trabajando son matrices bi-estocásticas (con la condición adicional de que las probabilidades de los partidos que no son válidos tienen que valer 0).

Una propiedad muy importante de las matrices bi-estocásticas y que nos será útil para construir el sorteo es el teorema de Birkhoff-von Neumann, que se enuncia a continuación.

Teorema de Birkhoff-von Neumann: Dada una matriz bi-estocástica B, existen números $b_1, b_2, \cdots, b_k > 0$ que cumplen $\sum_{i=1}^{k} b_i = 1$ y **matrices de permutación** $P_1, P_2, \cdots, P_k$ tales que:

$$B = b_1 P_1 + b_2 P_2 + \cdots + b_k P_k.$$

Este teorema y su demostración se pueden encontrar en [2]. Para entender el teorema, hay que definir antes qué es una matriz de permutación, lo cual se hace a continuación:

Definición (Matriz de Permutación): Es una matriz cuadrada cuyas entradas son todas 0 ó 1. Además, cada columna y cada fila tiene únicamente un 1. Por ejemplo, una matriz de permutación de dimensiones 8×8 podría ser como la siguiente:

$$P = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix}.$$

Algo interesante aquí es que las matrices de permutación de 8×8 se corresponden exactamente con posibles fases eliminatorias. Esto porque si a cada fila le asociamos un campeón y a cada columna un subcampeón, así como lo hemos estado haciendo, entonces podemos interpretar las entradas en las que se tiene un 1 como que ese partido fue elegido para la fase eliminatoria del torneo. Por las propiedades de la matriz de permutación, cada campeón estará emparejado con un y sólo un subcampeón y se tendrá así una posible fase de eliminatoria.

Por ejemplo, si en la matriz P mencionada anteriormente le asociamos cada fila a un campeón de grupo (que son MCI, LIV, AJX, RMA, BAY, MUN, LIL, JUV) y cada columna a un subcamepeón (PSG, ATM, SPO, INT, BEN, VIL, SAL, CHE). Entonces, la matriz P simboliza un torneo en el que los partidos son: MCI-ATM, LIC-PSG, AJX-INT, RMA-SPO, BAY-VIL, MUN-SAL, LIL-CHE, JUV-BEN.

Es decir, en general, **a cada elección de una fase de eliminación le corresponde una matriz de permutación**.

Pero además, vimos que una tabla de probabilidades es una matriz estocástica y por el teorema de Birkhoff-von Neumann, estas matrices se pueden escribir como una suma convexa de matrices de permutación. Por lo tanto, según este teorema, dada una tabla de probabilidades que llamaremos B, podemos encontrar matrices de permutación (que corresponden con posibles fases de eliminación) $P_1, P_2, P_3, \cdots, P_k$ y números $0 < b_i$ con $\sum_i b_i = 1$, tales que:

$$B = b_1P_1 + b_2P_2 + \cdots + b_kP_k.$$

Esto nos dice directamente la manera para encontrar cómo realizar un sorteo dada la tabla de probabilidades que queremos que lo describa:

- Dada la tabla de probabilidades B, usar el teorema de Birkhoff-von Neumann para escribirla como:

$$B = b_1 P_1 + \cdots + b_k P_k,$$

donde P_i son posibles emparejamientos para la fase de eliminación (representados como matrices de permutación en las que se coloca un 1 en los partidos que se van a jugar) y $0 < b_i$ con $\sum_i b_i = 1$.

Además, podemos estar seguros de que ninguna de las matrices P_i incluye un partido que no sea válido según las condiciones de la UEFA. Esto debido a que la tabla B vale 0 en los partidos no válidos y como los coeficientes $b_1, b_2, \cdots, b_k$ son positivos y las entradas de P_i son todas 1 o 0, la única forma de que B_{mn} sea cero es que todas las matrices P_i tengan un cero en la posición (m, n) (pues con que por lo menos una tenga un 1 en dicha posición, la suma $b_1 P_1 + \cdots + b_k P_k$ tendrá un número positivo en esa posición). Entonces, ninguna de las matrices P_i tiene un 1 en un partido no válido.

- Luego, consideramos el sorteo en el que se elige un emparejamiento de la fase de eliminación al azar de entre las opciones $P_1, P_2, \cdots, P_k$, cada uno con probabilidad $b_1, b_2, \cdots, b_k$. Como hay una probabilidad b_1 de elegir al torneo P_1, b_2 de elegir a P_2, etc., la tabla de probabilidad de este sorteo es $b_1 P_1 + \cdots + b_k P_k = B$.

Entonces, para todos los métodos que siguen, en los que primero se obtiene la tabla de probabilidades antes de definir la forma en que se hará el sorteo, el sorteo será como se acaba de mencionar. Es decir, una vez escrita la tabla de probabilidades como indica el teorema de Birkhoff, se hace un sorteo en el que se asigna una probabilidad b_i a cada emparejamiento de fase eliminatoria P_i. En el sorteo se elige entonces al azar una de las P_i, cada una con un peso b_i de ser elegida. De esta forma, nos aseguramos que la tabla de probabilidad correspondiente a este sorteo es la que queremos y el sorteo replica las probabilidades que queremos para cada partido.

Hay algunos comentarios sobre este proceso que valen la pena destacar:

- La descomposición de una matriz bi-estocástica en suma de matrices de permutación según describe el teorema de Birkhoff no es única. Esto significa que se pueden encontrar varios posibles sorteos para replicar una misma tabla de probabilidades.

- Existen algoritmos para encontrar esta descomposición de la matriz bi-estocástica. Uno de ellos se puede descargar de una librería en Python y se utiliza para encontrar las matrices P_i y las probabilidades b_i para replicar cada una de las tablas de probabilidad. Los resultados se muestran en el código y no se escriben aquí por ser demasiado largos para detallar.

- La cantidad de matrices P_i necesarias suele ser de alrededor de 30 y 40 para las tablas de probabilidad usadas (las matrices en particular que se utilizan dependen de cuál es la tabla de probabilidad).

- **Espectacularidad:** Para realizar el sorteo en estos casos, como se dijo antes, se selecciona un torneo P_i al azar con probabilidades b_i. Esto se tiene que hacer computacionalmente y se tiene la desventaja de que una vez que se selecciona una opción P_i, ya quedan determinados todos los partidos. Sin embargo, como se mencionó para otros métodos, se puede aparentar la espectacularidad al ir leyendo poco a poco los partidos del emparejamiento seleccionado. De esta forma, se puede simular la expectativa y sorpresa de decir a qué equipo le toca con qué otro equipo.

Para todos los métodos mencionados de aquí en adelante, lo descrito anteriormente será el proceso que debe de realizarse para hacer el sorteo, dada la tabla de probabilidades que resulta cada método.

3.2.3.3. Optimizar la suma de varianzas

Para este método usaremos el procedimiento mencionado anteriormente para encontrar la tabla de probabilidades que minimiza a un criterio y lo haremos en particular para el criterio de la suma de varianzas.

- **Tabla de Probabilidades:** Tras hacer la optimización computacionalmente, resulta que la tabla de probabilidades que minimiza al criterio de suma de varianzas es la siguiente (figura 3.9):

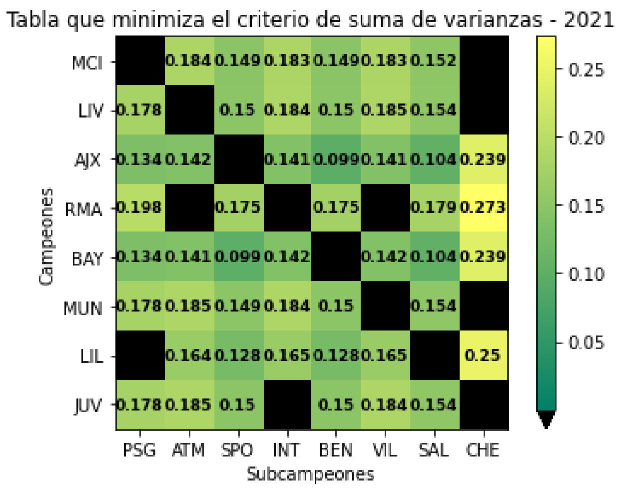

Figura 3.9: Tabla de probabilidades para el método de optimizar la suma de varianzas.

- **Criterios de Equidad:** A esta tabla le medimos los criterios de equidad y nos queda lo siguiente (figura 3.10:

Suma de varianzas	Varianza total	Rango	Máximo	Gini	Theil	Atkinson
0.013395	0.001280	0.173781	0.272892	0.114899	0.022806	0.011310

Figura 3.10: Resultado de los criterios aplicados al método que minimiza la suma de las varianzas.

Vemos que esta tabla efectivamente parece tener filas y columnas bastante uniformes, en las que la varianza es baja.

- **Requerimientos adicionales:** Se requiere del algoritmo de optimización para encontrar la tabla antes del torneo y luego para realizar el proceso del sorteo que se mencionó en la sección 3.2.3.2.

- **Espectacularidad:** Se aplica lo que se menciona en la sección 3.2.3.2.

3.2.3.4. Optimizar varianza total

Para este método usaremos el procedimiento mencionado anteriormente para encontrar la tabla de probabilidades que minimiza a un criterio y lo haremos en particular para el criterio de la varianza total.

- **Tabla de Probabilidades:** Tras hacer la optimización computacionalmente, resulta que la tabla de probabilidades que minimiza al criterio de varianza total es la siguiente (figura 3.11):

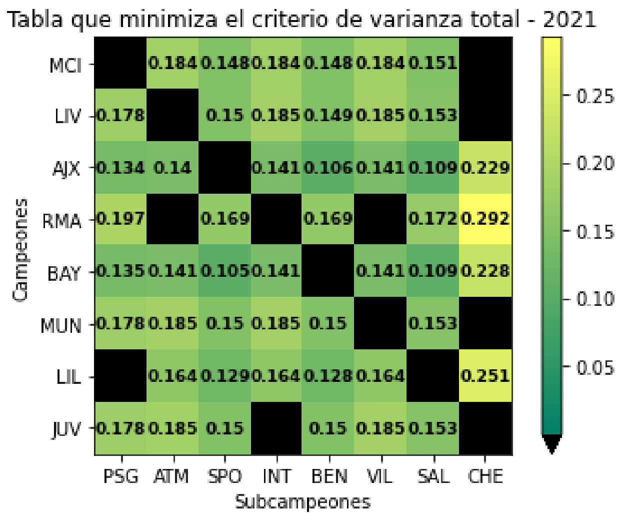

Figura 3.11: Tabla de probabilidades para el método de optimizar la varianza total.

Vemos que esta tabla es en general bastante uniforme y efectivamente parece que la varianza total es baja, excepto por algunos pocos valores demasiado altos o demasiado bajos.

- **Criterios de Equidad:** A esta tabla le medimos los criterios de equidad y nos queda la figura 3.12:

Suma de varianzas	Varianza total	Rango	Máximo	Gini	Theil	Atkinson
0.013809	0.001263	0.186556	0.291968	0.112425	0.021969	0.010768

Figura 3.12: Resultado de los criterios aplicados al método que minimiza la varianza total.

- **Requerimientos adicionales:** Se requiere del algoritmo de optimización para encontrar la tabla antes del torneo y luego para realizar el proceso del sorteo que se mencionó en la sección 3.2.3.2.

- **Espectacularidad:** Se aplica lo que se menciona en la sección 3.2.3.2.

3.2.3.5. Optimizar el máximo

Para este método usaremos el procedimiento mencionado anteriormente para encontrar la tabla de probabilidades que minimiza a un criterio y lo haremos en particular para el criterio del máximo.

- **Tabla de Probabilidades:** Tras hacer la optimización computacionalmente, resulta que la tabla de probabilidades que minimiza al criterio del máximo es la dada por la figura 3.13:

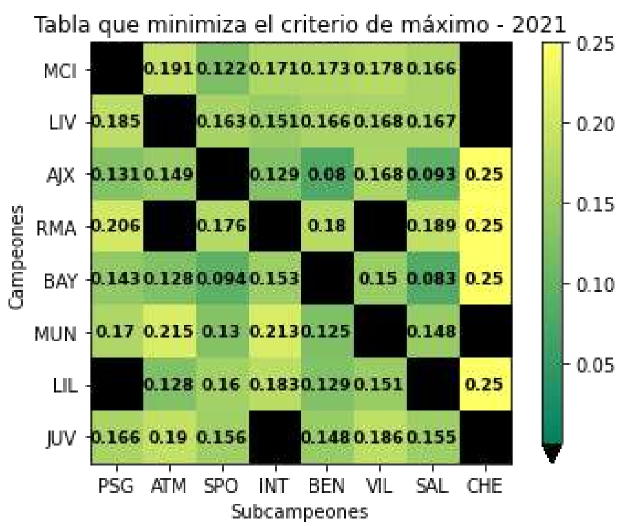

Figura 3.13: Tabla de probabilidades para el método de optimizar el máximo.

Vemos que la matriz tiene máximos de 0.25 en la columna de CHE, lo cual efectivamente es lo más bajo que pueden tomar los valores de esta columna para que siga sumando 1. Sin embargo, vemos que para compensar esto, hay algunas entradas muy bajas, como la de 0.08.

- **Criterios de Equidad:** A esta tabla le medimos los criterios de equidad y nos queda la figura 3.14:

Suma de varianzas	Varianza total	Rango	Máximo	Gini	Theil	Atkinson
0.018468	0.001596	0.169634	0.250000	0.132712	0.029616	0.015019

Figura 3.14: Resultado de los criterios aplicados al método que minimiza el máximo.

- **Requerimientos adicionales:** Se requiere del algoritmo de optimización para encontrar la tabla antes del torneo y luego para realizar el proceso del sorteo que se mencionó en la sección 3.2.3.2.

- **Espectacularidad:** Se aplica lo que se menciona en la sección 3.2.3.2.

- **Unicidad:** En este caso las soluciones no son únicas, pues hay muchas otras tablas de probabilidad que tienen como máximo a 0.25. Una de ellas se presenta en el planteamiento del concurso [3]. Más aún, en uno de los métodos posteriores que se definen en este trabajo, se deja fijo el máximo de la tabla y se varían los otros valores, con lo que se encuentra que hay muchas tablas con el mismo máximo y entonces existe la libertad de minimizar otros criterios.

3.2.3.6. Optimizar el rango

Para este método usaremos el procedimiento mencionado anteriormente para encontrar la tabla de probabilidades que minimiza a un criterio y lo haremos en particular para el criterio del rango.

- **Tabla de Probabilidades:** Tras hacer la optimización computacionalmente, resulta que la tabla de probabilidades que minimiza al criterio del rango es la siguiente (figura 3.15):

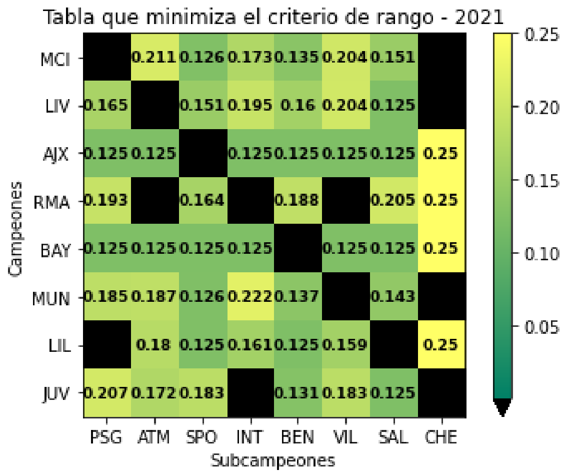

Figura 3.15: Tabla de probabilidades para el método de optimizar el rango.

Vemos que al igual que en la optimización del máximo, los valores más altos son los 0.25 de la columna de CHE. Sin embargo, a diferencia de la optimización del máximo, al requerir aquí que se optimice el rango, los mínimos no son tan bajos, teniendo un valor de 0.125. Esto es una mejoría respecto a la tabla que optimiza al máximo y es hasta el momento el mejor resultado.

- **Criterios de Equidad:** A esta tabla le medimos los criterios de equidad y nos queda lo siguiente (figura 3.16):

Suma de varianzas	Varianza total	Rango	Máximo	Gini	Theil	Atkinson
0.019285	0.001603	0.125037	0.250000	0.132604	0.028288	0.013902

Figura 3.16: Resultado de los criterios aplicados al método que minimiza el rango.

- **Requerimientos adicionales:** Se requiere del algoritmo de optimización para encontrar la tabla antes del torneo y luego para realizar el proceso del sorteo que se mencionó en la sección 3.2.3.2.

- **Espectacularidad:** Se aplica lo que se menciona en la sección 3.2.3.2.

- **Unicidad:** En este caso las soluciones no son únicas, pues hay muchas otras tablas de probabilidad que tienen como máximo a 0.25. Una de ellas se presenta en el planteamiento del concurso [3]. Más aún, en uno de los métodos posteriores que se definen en este trabajo, se deja fijo el rango de la tabla y se varían los otros valores, con lo que se encuentra que hay muchas tablas con el mismo máximo y entonces existe la libertad de minimizar otros criterios.

3.2.3.7. Optimizar el criterio de Gini

Para este método usaremos el procedimiento mencionado anteriormente para encontrar la tabla de probabilidades que minimiza a un criterio y lo haremos en particular para el criterio de Gini.

- **Tabla de Probabilidades:** Tras hacer la optimización computacionalmente, resulta que la tabla de probabilidades que minimiza al criterio del rango es la siguiente (figura 3.17):

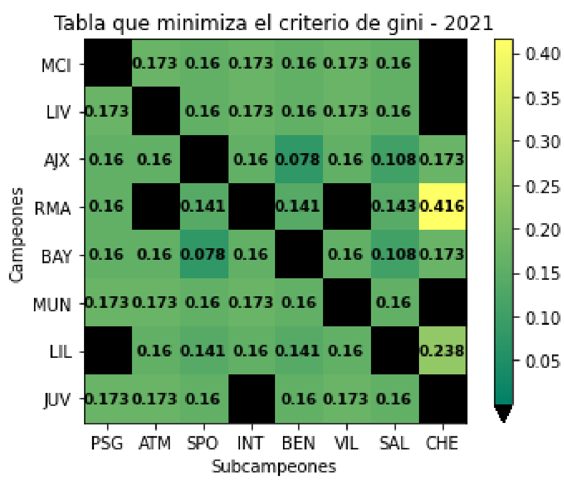

Figura 3.17: Tabla de probabilidades para el método de optimizar el criterio de Gini.

Cabe destacar aquí que la tabla resultante es muy distinta a las demás, pues tiene un valor muy alto (0.416) pero todos los demás valores están bastante equilibrados. Esto seguramente no sería conveniente para el sorteo, ya que hay un partido demasiado probable. Sin embargo, este partido muy probablemente se compensa con la casi homogeneidad de las probabilidades de los demás partidos. Podría ser una selección interesante para realizar los sorteos, ya que muchos equipos estarían conformes con la homogeneidad de sus posibles resultados (pero un par de equipos podrían estar enojados por la casi certeza de conocer a sus rivales).

- **Criterios de Equidad:** A esta tabla le medimos los criterios de equidad y nos queda la siguiente figura 3.18:

Suma de varianzas	Varianza total	Rango	Máximo	Gini	Theil	Atkinson
0.034369	0.001958	0.337615	0.415587	0.094684	0.030037	0.014187

Figura 3.18: Resultado de los criterios aplicados al método que minimiza el criterio de Gini.

Al igual que para los otros métodos, compararemos estos resultados con los demás más adelante.

- **Requerimientos adicionales:** Se requiere del algoritmo de optimización para encontrar la tabla antes del torneo y luego para realizar el proceso del sorteo que se mencionó en la sección 3.2.3.2.

- **Espectacularidad:** Se aplica lo que se menciona en la sección 3.2.3.2.

- **Unicidad:** El resultado parece ser único, pues al repetir la optimización con distintos parámetros y puntos iniciales, siempre se llegó a esta misma tabla óptima.

3.2.3.8. Optimizar el criterio de Theil

Para este método usaremos el procedimiento mencionado anteriormente para encontrar la tabla de probabilidades que minimiza a un criterio y lo haremos en particular para el criterio de Theil.

- **Tabla de Probabilidades:** Tras hacer la optimización computacionalmente, resulta que la tabla de probabilidades que minimiza al criterio de Theil es la siguiente dada por la figura 3.19:

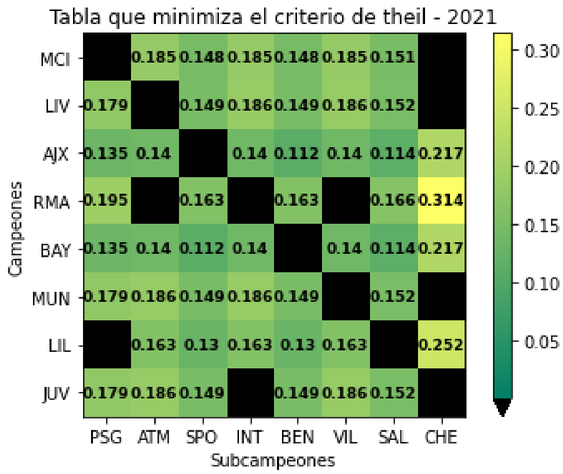

Figura 3.19: Tabla de probabilidades para el método de optimizar el criterio de Theil.

- **Criterios de Equidad:** A esta tabla le medimos los criterios de equidad y nos queda la figura 3.20:

Suma de varianzas	Varianza total	Rango	Máximo	Gini	Theil	Atkinson
0.015276	0.001285	0.201706	0.314140	0.109446	0.021646	0.010454

Figura 3.20: Resultado de los criterios aplicados al método que minimiza el criterio de Theil.

Se obtiene un resultado similar al de Gini pero menos extremo, en el que uno de los partidos es bastante más probable que los demás pero los demás tienen una distribución muy uniforme.

- **Requerimientos adicionales:** Se requiere del algoritmo de optimización para encontrar la tabla antes del torneo y luego para realizar el proceso del sorteo que se mencionó en la sección 3.2.3.2.

- **Espectacularidad:** Se aplica lo que se menciona en la sección 3.2.3.2.

- **Unicidad:** El resultado parece ser único, pues al repetir la optimización con distintos parámetros y puntos iniciales, siempre se llegó a esta misma tabla óptima.

3.2.3.9. Optimizar el criterio de Atkinson

Para este método usaremos el procedimiento mencionado anteriormente para encontrar la tabla de probabilidades que minimiza a un criterio y lo haremos en particular para el criterio de Atkinson.

- **Tabla de Probabilidades:** Tras hacer la optimización computacionalmente, resulta que la tabla de probabilidades que minimiza al criterio del rango está dada por (figura 3.21:

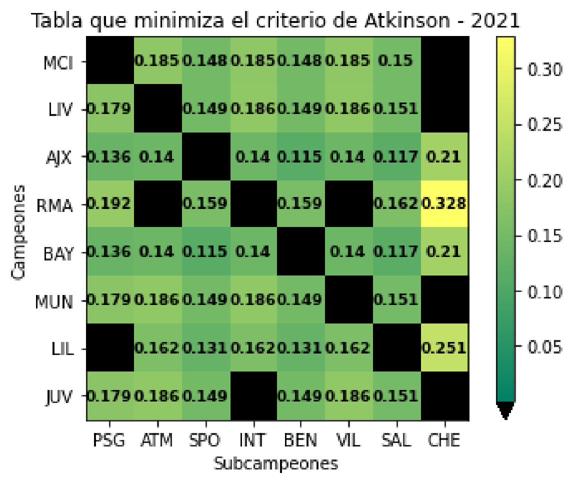

Figura 3.21: Tabla de probabilidades para el método de optimizar el criterio de Atkinson.

- **Criterios de Equidad:** A esta tabla le medimos los criterios de equidad y nos queda lo siguiente (figura 3.22):

Suma de varianzas	Varianza total	Rango	Máximo	Gini	Theil	Atkinson
0.016668	0.001319	0.212732	0.328196	0.107533	0.021751	0.010407

Figura 3.22: Resultado de los criterios aplicados al método que minimiza el criterio de Atkinson.

Curiosamente, al igual que en los otros dos criterios basados en medidas económicas, se obtienen matrices con un valor muy alto y luego valores bastante uniformes. Parece que estos métodos no castigan tan fuertemente a este tipo de resultados y la homogeneidad de la mayoría de los valores compensa la existencia de uno tan alto.

- **Requerimientos adicionales:** Se requiere del algoritmo de optimización para encontrar la tabla antes del torneo y luego para realizar el proceso del sorteo que se mencionó en la sección 3.2.3.2.
- **Espectacularidad:** Se aplica lo que se menciona en la sección 3.2.3.2.
- **Unicidad:** El resultado parece ser único, pues al repetir la optimización con distintos parámetros y puntos iniciales, siempre se llegó a esta misma tabla óptima.

3.2.4. Métodos con un parámetro libre

En esta sección del documento analizamos métodos cuya definición depende de un parámetro libre que se puede ir variando para encontrar el valor en el que alcanza un minimiza algo que se quiera minimizar.

3.2.4.1. Método de generalizar la varianza

Como vimos en la sección de los criterios, la varianza de un conjunto de N datos x_i se define como:

$$V(\{x_i\}) = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2,$$

donde $\bar{x}$ es el promedio de $\{x_i\}$. Esta definición puede generalizarse de forma sencilla cambiando el exponente 2 por otros valores. En general, si definimos un exponente $\alpha \in (0, \infty)$, podemos modificar la definición de la varianza como

$$V_\alpha(\{x_i\}) = \frac{1}{N-1}\sum_{i}|x_i - \bar{x}|^\alpha,$$

que ahora depende del parámetro libre α. Es necesario aplicar el valor absoluto a la diferencia $x_i - \bar{x}$, para asegurarnos que la suma de $|x_i - \bar{x}|^\alpha$ no se anula ni se hace negativa.

A partir de esta definición, se propone un nuevo método de sorteo que consiste en lo siguiente:

1. Definir un valor para el exponente α.
2. Encontrar la tabla de probabilidades que minimiza a V_α calculada para todos los datos de la tabla. Para encontrar la tabla, se utiliza el código computacional para la optimización y se sigue el procedimiento explicado anteriormente.
3. Calcular el criterio del rango para esta tabla.
4. Repetir para distintos valores de α y encontrar el valor de α en el que se minimiza el rango.
5. La tabla de probabilidades correspondiente a este valor de α será nuestro resultado.

La idea de este método es que para cada α se minimiza V_α, que es una cantidad similar a la varianza y con una interpretación similar. Pero además, entre éstas, se busca la α que minimiza luego al rango, que ofrece una forma distinta a la varianza de medir la equidad. De esta forma, idealmente se podrá tener lo mejor de ambas medidas (varianza y rango) y obtener una tabla bastante equilibrada.

Se realizó este procedimiento para los equipos del 2021. Para α se usaron valores desde 0.25 hasta 7.5, tomados en pasos de 0.25. Como se menciona en el procedimiento, para cada uno de estos valores de α, se obtiene la tabla de probabilidades que minimiza V_α y se le calcula el rango (se puede ver en el código anexado a la solución). En la figura 3.23 se muestran estos rangos para cada uno de los valore de α.

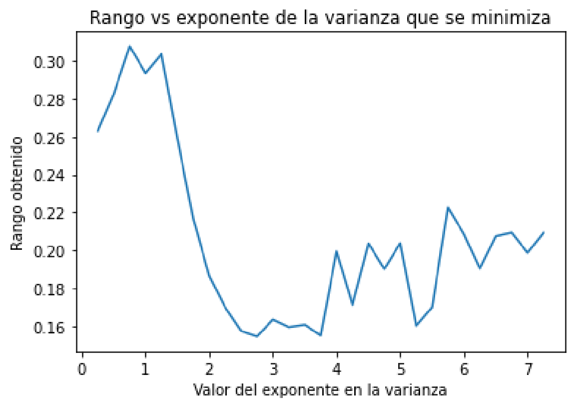

Figura 3.23: Rango obtenido para la tabla que minimiza V_α para varios valores del exponente α

Se puede observar que el mínimo del rango se encuentra cuando α es de 2.75. Es decir, de entre todas las tablas que minimizan V_α para distintos valores de α, la que minimiza el rango es la que tiene $\alpha = 2.75$. En la figura 3.24 se muestra esta tabla.

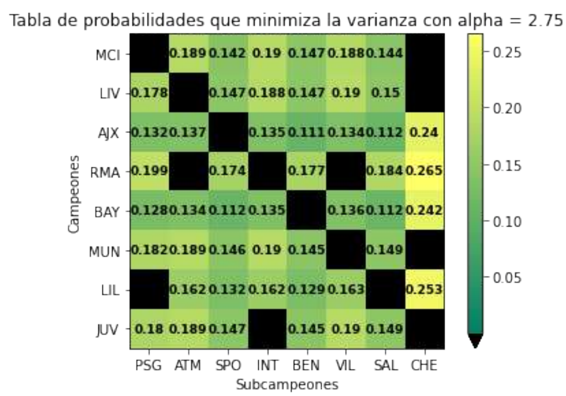

Figura 3.24: Tabla de Probabilidades que minimiza V_α con $\alpha = 2.75$

Medimos ahora los criterios para esta tabla y se obtiene la figura 3.25:

Suma de varianzas	Varianza total	Rango	Máximo	Gini	Theil	Atkinson
0.013930	0.001306	0.154569	0.265384	0.118432	0.022946	0.011272

Figura 3.25: Resultado de los criterios para la tabla que minimiza V_α con $\alpha = 2.75$

Requerimientos adicionales: Se requiere del algoritmo de optimización para encontrar la tabla antes del torneo y luego para realizar el proceso del sorteo que se mencionó en la sección 3.2.3.2.

Espectacularidad: Se aplica lo que se menciona en la sección 3.2.3.2.

3.2.4.2. Diferencia entre promedio de máximos y mínimos

Para este método, se generaliza el criterio del rango de una forma natural. Sabemos que el rango mide la diferencia entre el valor máximo y el mínimo de la tabla, por lo que una forma natural de generalizarlo es medir la diferencia entre el promedio de los n valores máximos y el promedio de los n valores mínimos de la tabla. De esta forma, esta medida no sólo toma en cuenta los valores extremos de la tabla, sino también algunos cercanos a estos. Es decir, para una muestra de datos $\{x_i\}$ (los datos de la tabla), definimos lo siguiente:

$$Ran_n(\{x_i\}) := \overline{\{x_i\}^n} - \overline{\{x_i\}_n},$$

donde $\{x_i\}^n$ simboliza el conjunto de los n valores más grandes de los datos, y $\{x_i\}_n$ el de los n menores y la barra denota el promedio. En particular, para $n = 1$ recuperamos el rango común y corriente.

Dado este criterio, podemos definir el siguiente procedimiento para encontrar una tabla de probabilidades que podría resultar equilibrada:

1. Elegir un valor valor entero para n.
2. Encontrar la tabla de probabilidades que minimiza a Ran_n, utilizando el mismo método de optimización que se usó para los otros problemas.
3. Para esta tabla de probabilidades, calcular la varianza total.
4. Repetir para distintos valores de n y encontrar el valor de n en el que la varianza de la tabla es mínima.
5. La tabla de probabilidades que corresponde a este valor de n será el resultado.

La idea de este método es que para cada valor de n se minimiza una cantidad similar al rango, y dentro de todas estas tablas, se encuentra el valor de n que minimiza la varianza. Se realizó este procedimiento para los equipos del 2021. Para n se usaron valores desde 1 hasta 15, tomados en pasos de uno en uno. Como se menciona en el procedimiento, para cada uno de estos valores de n, se obtiene la tabla de probabilidades que minimiza Ran_n y se le calcula la varianza. En la figura 3.26 se muestran estas varianzas para cada uno de los valores de n.

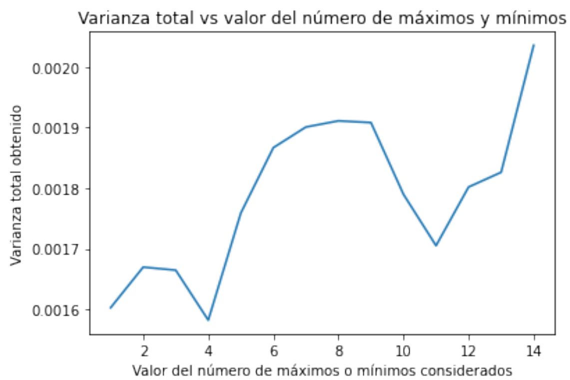

Figura 3.26: Varianza total obtenida para la tabla que minimiza Ran_n para varios valores de n

Se puede observar que el mínimo de la varianza total se encuentra cuando $n = 4$. Es decir, de entre todas las tablas que minimizan Ran_n para distintos valores de n, la que minimiza la varianza total es la que tiene $n = 4$. En la figura 3.27 se muestra esta tabla.

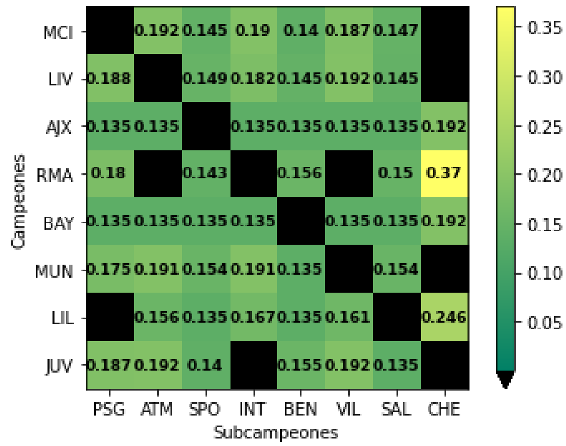

Figura 3.27: Tabla de Probabilidades que minimiza Ran_n con $n = 4$

El resultado no es muy bueno. Aunque gran parte de la tabla parece estar equilibrada, hay un valor muy alto en una de las celdas. Esto se puede deber a que al interesarnos en el promedio de los n valores más altos en vez de sólo el valor más alto, puede haber un valor muy alto pero de tal forma que sus valores cercanos sean considerablemente menores y bajen el promedio.

Medimos ahora los criterios para esta tabla y se obtiene la figura 3.28:

Suma de varianzas	Varianza total	Rango	Máximo	Gini	Theil	Atkinson
0.024441	0.001582	0.235617	0.370351	0.106636	0.024516	0.011416

Figura 3.28: Criterios evaluados en la tabla que minimiza Ran_n

Requerimientos adicionales: Se requiere del algoritmo de optimización para encontrar la tabla antes del torneo y luego para realizar el proceso del sorteo que se mencionó en la sección 3.2.3.2.

Espectacularidad: Se aplica lo que se menciona en la sección 3.2.3.2.

3.2.5. Método favorito

A continuación se presenta nuestro método favorito para resolver el problema y que se basa en lo aprendido a partir de los otros métodos. El procedimiento para realizarlo es el siguiente:

1. Encontrar la tabla de probabilidades que minimiza al rango, al igual que se hizo en el método de minimizar al rango.

2. Encontrar el valor máximo y mínimo de esa tabla. Con ello, sabemos que cualquier tabla de probabilidades que tenga el rango lo más pequeño posible, tiene que tener todas sus entradas entre ese mínimo y máximo.

3. Aplicar de nuevo el algoritmo de optimización, pero ahora para la suma de varianzas y partiendo desde la tabla del paso 1. Además, ahora la optimización con la restricción adicional de que todos los valores de la tabla tengan que estar confinados entre el mínimo y máximo.

   De esta forma, la tabla que se obtenga tendrá el mismo rango que la del paso 1, por lo que podemos estar seguros de que minimiza al rango (y probablemente al máximo también). Pero además, de entre todas las matrices que minimizan al rango, buscamos la que minimice luego a la suma de varianzas. De esta manera, obtenemos una matriz más equilibrada según el criterio de varianzas pero que aún minimice el rango.

Se realizó este procedimiento para los datos del 2021 y la tabla obtenida es la de la figura 3.29.

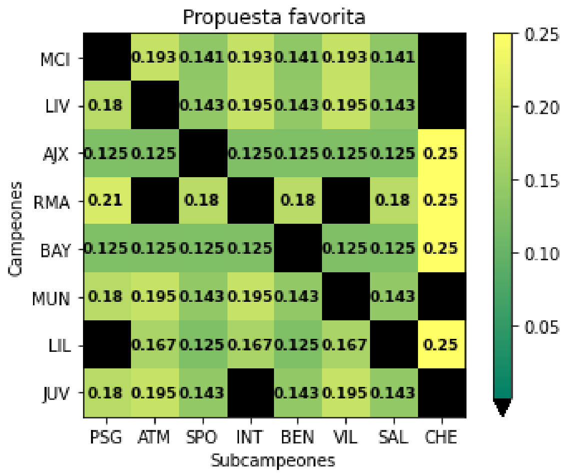

Figura 3.29: Tabla de Probabilidades con el método favorito.

Luego, se le calcula a esta matriz los criterios y se obtiene la figura 3.30:

Suma de varianzas	Varianza total	Rango	Máximo	Gini	Theil	Atkinson
0.016055	0.001425	0.125037	0.250000	0.123908	0.025068	0.012301

Figura 3.30: Resultados para la propuesta favorita.

Requerimientos adicionales: Se requiere del algoritmo de optimización para encontrar la tabla antes del torneo y luego para realizar el proceso del sorteo que se mencionó en la sección 3.2.3.2.

Espectacularidad: Se aplica lo que se menciona en la sección 3.2.3.2.

Unicidad: Aunque es difícil decir si la tabla obtenida por este método es única, el método prueba que hay muchas matrices con el rango y el máximo óptimo, lo que prueba que esas matrices no son únicas.

3.3. Comparativa

En esta sección vamos a comparar los resultados de los métodos que hemos definido hasta ahora en la tabla 3.31. En las filas se presentan los métodos utilizados y en las columnas se presentan los criterios medidos para cada método. El problema para realizar la comparación es que cada criterio tiene un rango distinto de valores posibles; por ejemplo, la varianza da resultados del orden de 0.0013, mientras que el máximo da resultados del orden de 0.25.

Por lo tanto, lo que hicimos fue dividir cada criterio entre el valor del criterio evaluado a la tabla de probabilidades con el método de la UEFA. Así, por ejemplo, la varianza total del método de la UEFA vale 0.001288, mientras que con el método uniforme es de 0.001304, por lo que en la tabla 3.31, pondremos un valor de $0.001304/0.001288 = 1.0124$ para el criterio de varianza total medido en el método uniforme. La ventaja de esto es que un valor mayor que 1 indica que el criterio evaluado en el método es peor que si se hiciera el sorteo de la UEFA (porque se obtuvo un valor más alto en el criterio), mientras que un resultado menor que 1 indica que el método obtuvo un mejor valor al medirle el criterio.

Procedimiento Realizado	Valor Suma de varianzas	Valor Varianza total	Valor Rango	Valor Máximo	Valor Gini	Valor Theil	Valor Atkinson
UEFA	1	1	1	1	1	1	1
UEFA Pesado	1.053	1.016	1.024	1.023	1.0	1.004	0.998
Uniforme	1.058	1.014	1.044	1.036	0.984	0.999	0.992
Minimizar Suma de varianzas	0.882	0.996	0.866	0.875	1.045	1.05	1.077
Minimizar Varianza total	0.909	0.982	0.929	0.936	1.022	1.012	1.025
Minimizar Rango	1.27	1.247	0.623	0.801	1.206	1.303	1.324
Minimizar Máximo	1.216	1.242	0.845	0.801	1.207	1.364	1.43
Minimizar Gini	2.264	1.523	1.682	1.332	0.861	1.383	1.351
Minimizar Theil	1.006	1.0	1.005	1.007	0.995	0.997	0.996
Minimizar Atkinson	1.098	1.026	1.06	1.052	0.978	1.002	0.991
Generalizar la Varianza	0.917	1.016	0.77	0.851	1.077	1.057	1.073
Diferencia de promedios	1.61	1.231	1.174	1.187	0.97	1.129	1.087
Favorito	1.057	1.109	0.623	0.801	1.127	1.154	1.171

Figura 3.31: Comparación de resultados de cada criterio para los métodos utilizados.

Analizando la tabla, tenemos las siguientes observaciones:

- Debido a que los resultados se dividen entre el valor que tiene el sorteo de la UEFA, todos los resultados en el renglón de la UEFA son 1.
- Vemos que ningún renglón tiene puros valores menores que 1, lo que significa que el método renglón es mejor que el método de la UEFA para todos los criterios usados. Esto implica que decir si un resultado es mejor que otro es algo muy relativo y depende del criterio que se utilice.
- Se puede ver que el procedimiento de minimizar la suma de varianzas tiene el valor mínimo en la columna de suma de varianzas (como debería, pues este método busca minimizar a este criterio). Similarmente, el procedimiento de minimizar la varianza

total tiene el valor mínimo en la columna de varianza total y lo mismo para los otros métodos de minimización.

- Podemos ver que los métodos de minimizar Gini, Theil y Atkinson no son muy buenos según los criterios que miden varianza, pues dan resultados altos en estas columnas. Esto se puede deber a que habíamos notado que estos métodos suelen tener un valor muy alto (que contribuye mucho a la varianza) y que al parecer no se compensa con la homogeneidad del resto de sus valores. Como sus resultados no son muy buenos, podemos concluir que no hay que darle demasiada importancia a las columnas de criterios de Gini, Theil y Atkinson, pues ya vemos que incluso al optimizarlas no se obtienen muy buenos resultados.

- Los métodos de minimizar la suma de varianzas y la varianza total son considerablemente buenos, pues no sólo tienen valores bajos para las varianzas, sino que sus valores para el rango y máximo están bastante bien y aunque no se acercan al óptimo en estos criterios, son considerablemente mejores que otros métodos.

- Los métodos de optimización de rango y máximo tienen, como es de esperar, los mejores valores en el rango y máximo. Sin embargo, en cuanto a varianzas no son tan buenos, pues ahí tienen valores más de 20 % peores a los de la UEFA. Entre ellos, el de rango es mejor, pues no sólo tiene el mismo máximo que el método de máximo, sino que su criterio de rango es claramente mucho mejor.

- El método que generaliza la varianza tiene valores de varianza total y suma de varianzas bajos (aunque no tanto como los métodos que minimizan estas cantidades). Sin embargo, esta pérdida se compensa con una gran mejoría en el rango y máximo respecto a los métodos de optimización de varianzas.

- El método de diferencia de promedios no es bueno, pues se ve superado por muchos otros métodos en las primeras 4 columnas.

- El método propuesto como favorito lo consideramos así ya que tiene las mismas ventajas que el de optimización del rango (que minimiza al rango y al máximo), pero además tiene mejores valores que éste en la suma de varianzas y la varianza total.

Capítulo 4

Comparación de varios años

Vamos a seguir el mismo análisis pero de una manera más resumida para los 5 años anteriores al 2021; año con el que estuvimos trabajando a lo largo de la investigación.

4.1. 2020

Comenzaremos obteniendo la tabla de las posibles partidas que se tienen para este año, como se ve en la figura 4.1. Se puede calcular que para estos equipos se tienen 3,305 posibles fases de eliminación.

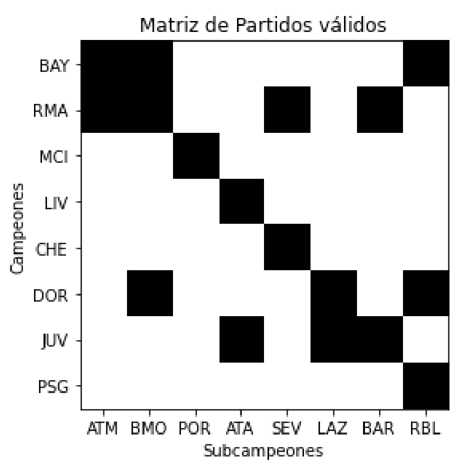

Figura 4.1: Tabla con todos los partidos posibles en 2020. Los recuadros negros representan los partidos que no pueden jugarse porque rompen alguna de las tres condiciones.

Posteriormente, aplicaremos todos los métodos que vimos a lo largo del trabajo a estos equipos, incluyendo el que usa la UEFA. Estos métodos son aplicados con ayuda del programa de Python que se envía junto a este trabajo y son los que se muestran en las figuras 4.2 y 4.3.

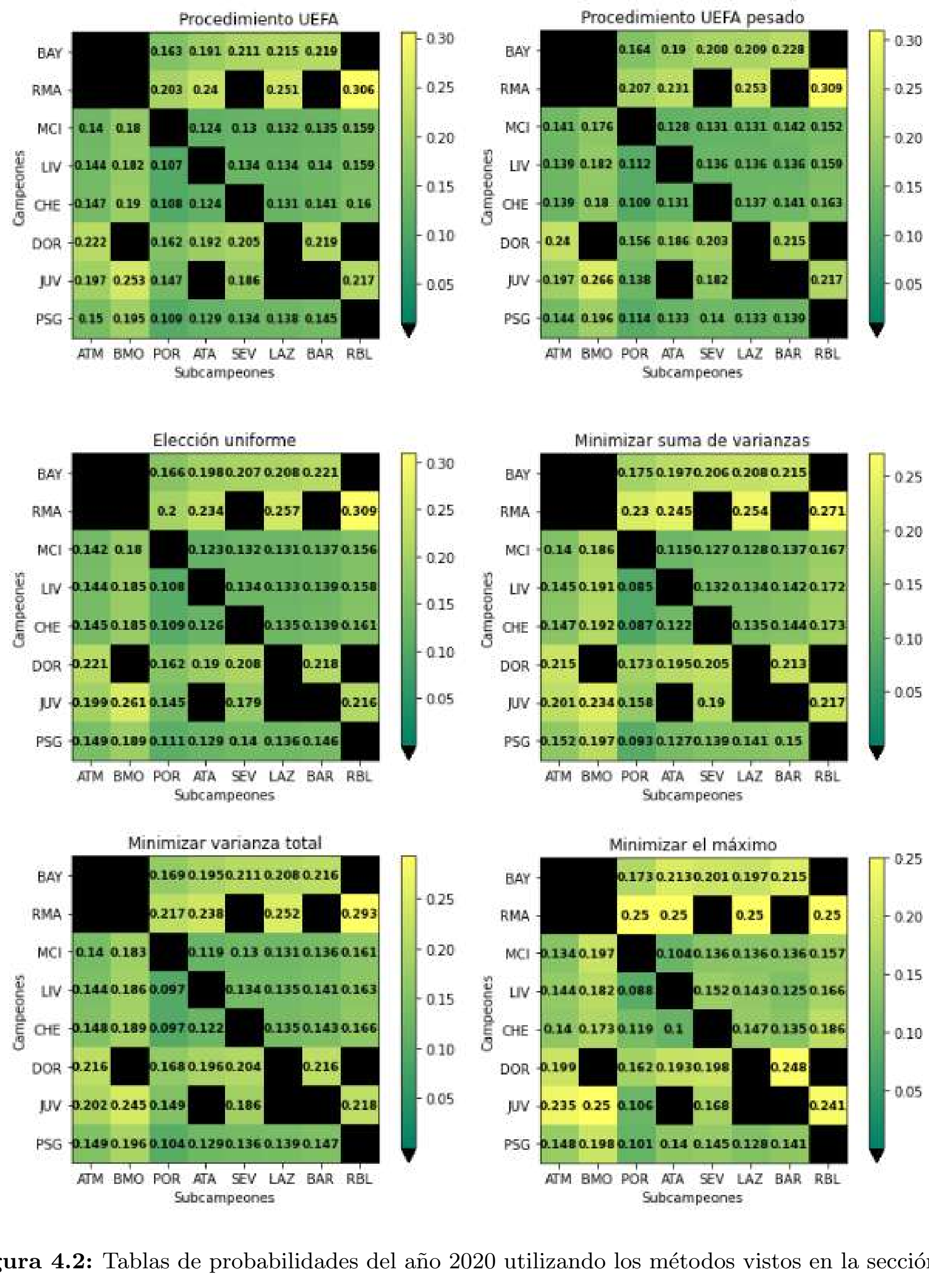

Figura 4.2: Tablas de probabilidades del año 2020 utilizando los métodos vistos en la sección 3.

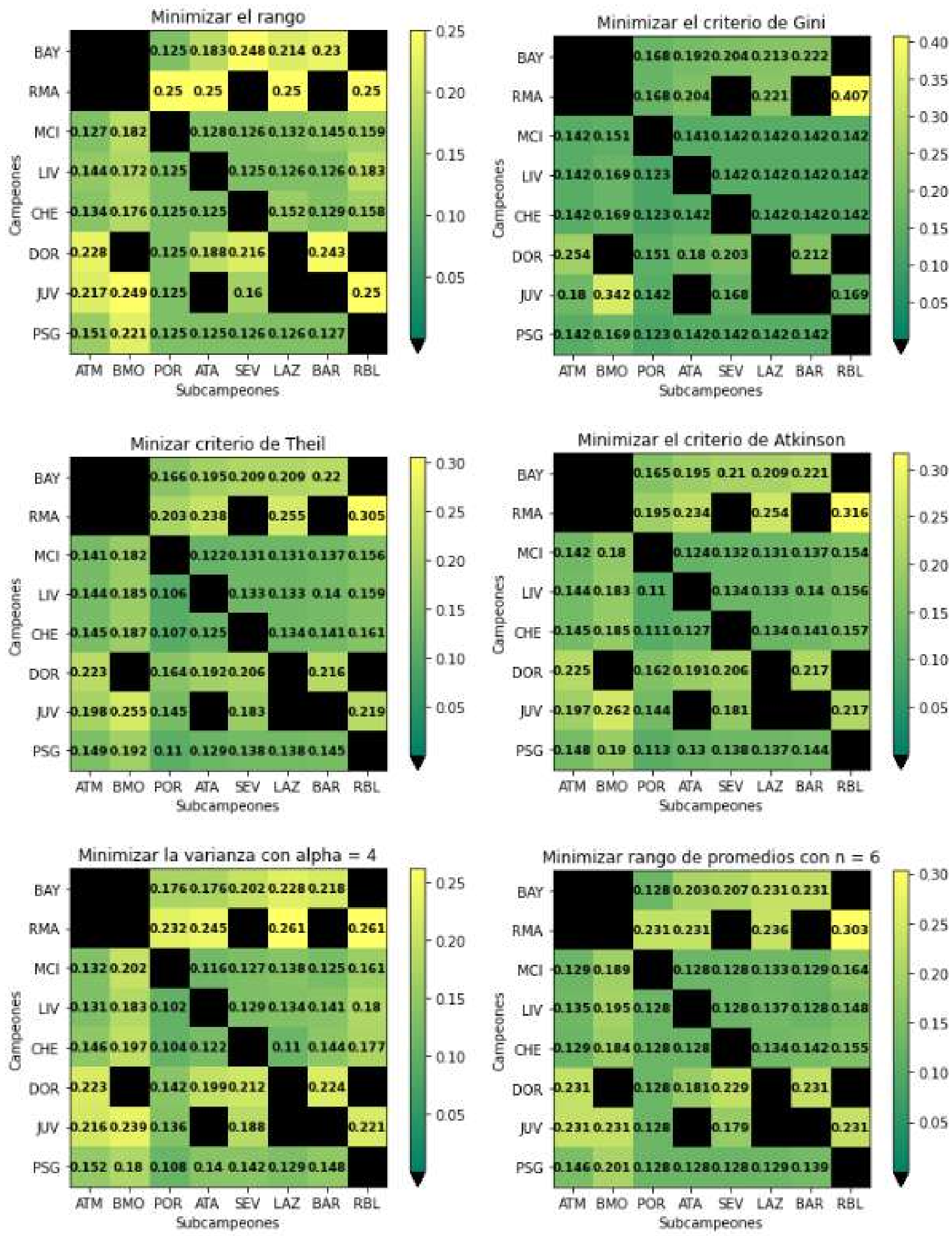

Figura 4.3: Tablas de probabilidades del año 2020 utilizando los métodos vistos en la sección 3.

El método que consideramos el mejor en términos de una mayor equidad, es el correspondiente a la siguiente figura 4.4.

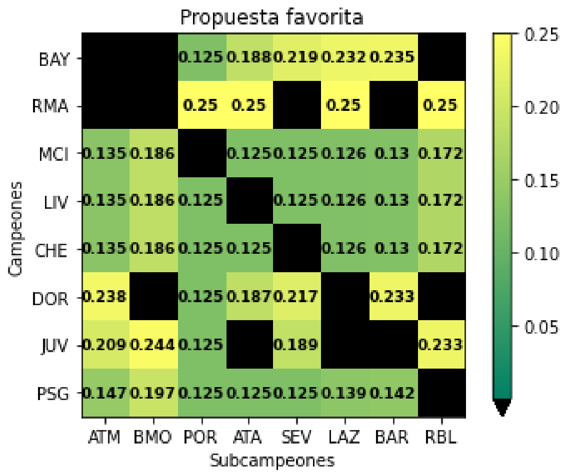

Figura 4.4: Tabla de probabilidades para el método favorito propuesto, del año 2020.

Finalmente, obtenemos la tabla comparativa final en donde se puede apreciar cómo la opción favorita es la mejor opción, pues de entre todas las tablas que minimizan al rango, también minimiza la varianza.

Procedimiento Realizado	Valor Suma de varianzas	Valor Varianza total	Valor Rango	Valor Máximo	Valor Gini	Valor Theil	Valor Atkinson
UEFA	1	1	1	1	1	1	1
UEFA Pesado	1.063	1.026	1.004	1.01	0.996	1.012	1.004
Uniforme	1.026	1.007	1.015	1.012	0.996	1.0	0.996
Minimizar Suma de varianzas	0.902	0.995	0.937	0.887	1.025	1.052	1.083
Minimizar Varianza total	0.936	0.985	0.989	0.958	1.007	1.009	1.021
Minimizar Rango	1.278	1.184	0.63	0.818	1.076	1.191	1.19
Minimizar Máximo	1.133	1.12	0.814	0.818	1.087	1.162	1.184
Minimizar Gini	2.106	1.459	1.428	1.331	0.953	1.257	1.177
Minimizar Theil	0.997	0.998	1.003	0.998	0.998	0.998	0.998
Minimizar Atkinson	1.058	1.017	1.04	1.035	0.992	1.002	0.994
Generalizar la Varianza	1.009	1.046	0.803	0.856	1.054	1.079	1.094
Diferencia de promedios	1.214	1.125	0.88	0.991	1.026	1.121	1.117
Favorito	1.169	1.128	0.63	0.818	1.057	1.144	1.148

Figura 4.5: Comparación de resultados de cada criterio para los métodos utilizados en el año 2020.

4.2. 2019

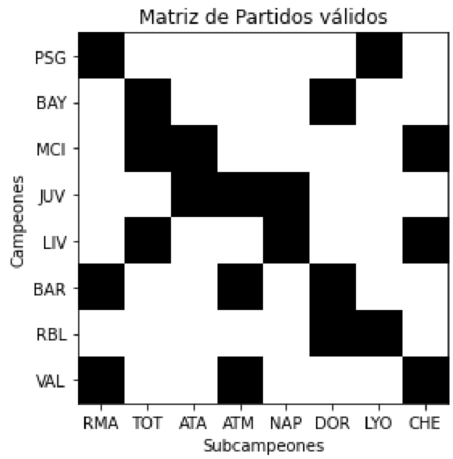

Figura 4.6: Tabla con todos los partidos posibles en 2019. Los recuadros negros representan los partidos que no pueden jugarse porque rompen alguna de las tres condiciones.

El número total de posibles emparejamientos para la fase de eliminación es de $2,002$.

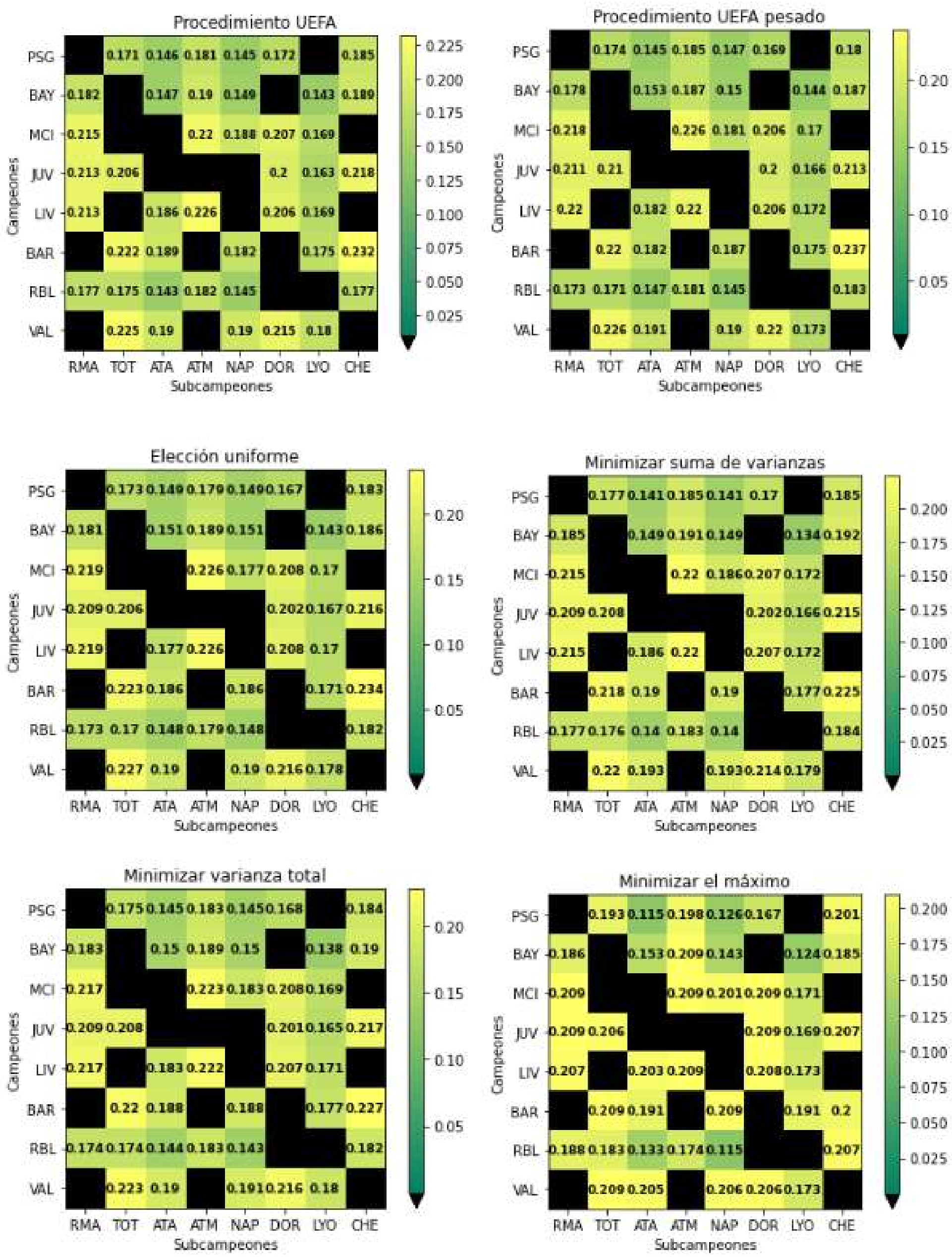

Figura 4.7: Tablas de probabilidades del año 2019 utilizando los métodos vistos en la sección 3.

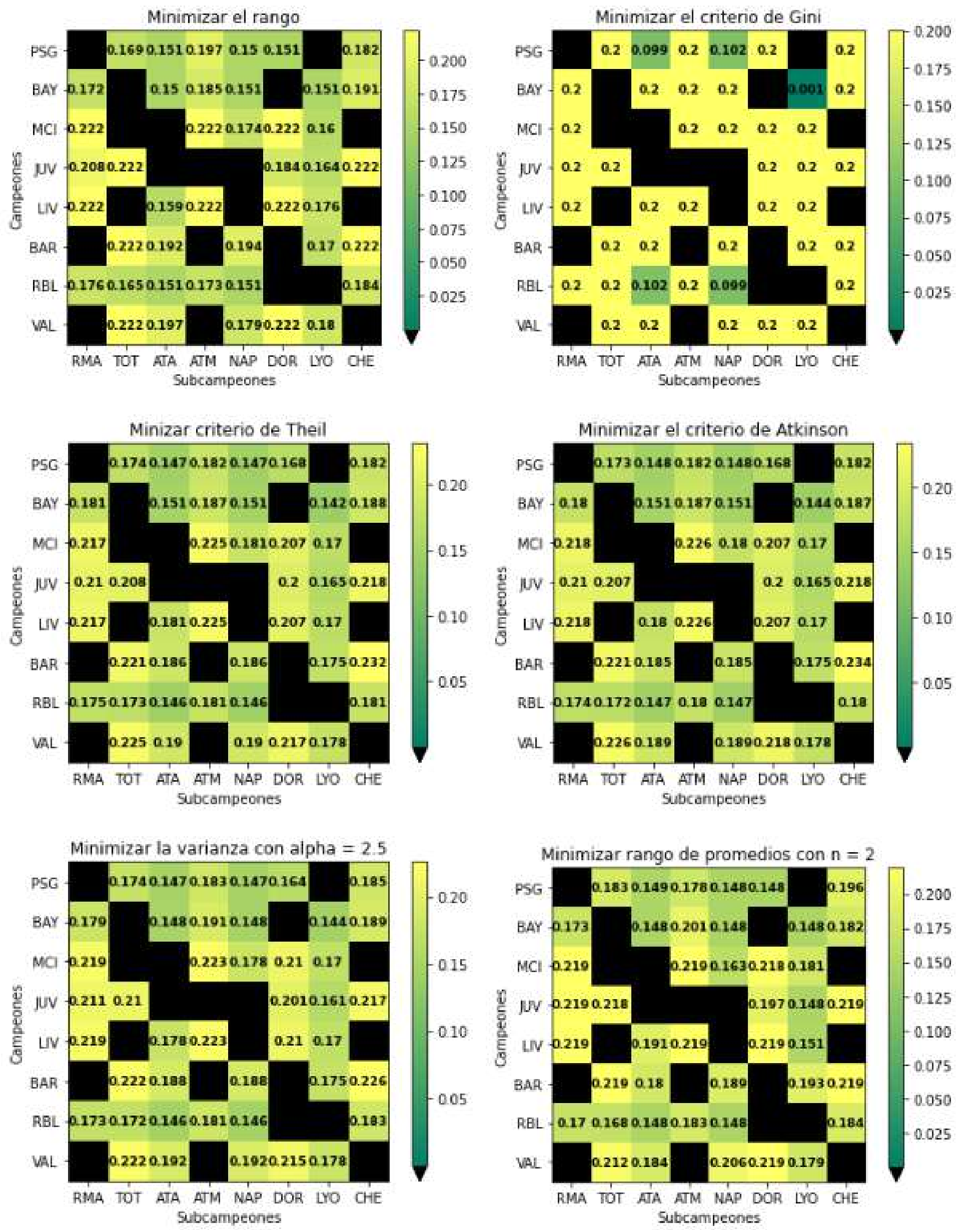

Figura 4.8: Tablas de probabilidades del año 2019 utilizando los métodos vistos en la sección 3.

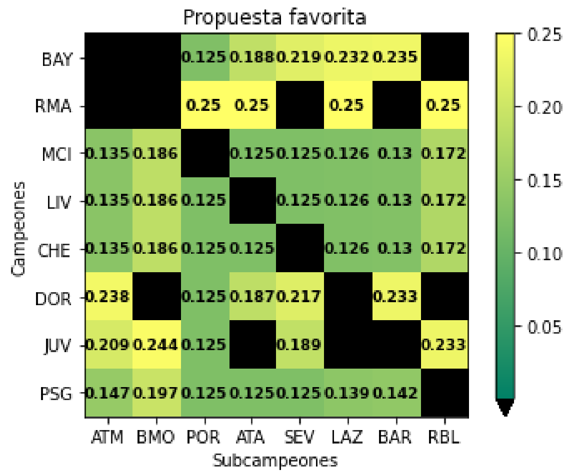

Figura 4.9: Tabla de probabilidades para el método favorito propuesto, del año 2019.

Procedimiento Realizado	Valor Suma de varianzas	Valor Varianza total	Valor Rango	Valor Máximo	Valor Gini	Valor Theil	Valor Atkinson
UEFA	1	1	1	1	1	1	1
UEFA Pesado	1.026	1.006	1.044	1.023	1.002	0.998	0.993
Uniforme	1.032	1.007	1.021	1.008	1.004	0.995	0.99
Minimizar Suma de varianzas	0.965	0.996	1.022	0.969	0.988	1.013	1.021
Minimizar Varianza total	0.978	0.991	1.004	0.981	0.993	0.995	0.997
Minimizar Rango	1.241	1.117	0.803	0.958	1.052	1.103	1.095
Minimizar Máximo	1.368	1.29	1.062	0.903	1.025	1.381	1.432
Minimizar Gini	3.391	2.653	2.24	0.863	0.885	4.068	6.151
Minimizar Theil	1.002	0.995	1.007	0.999	0.996	0.99	0.988
Minimizar Atkinson	1.021	1.001	1.01	1.007	0.998	0.991	0.986
Generalizar la Varianza	1.005	0.999	0.916	0.974	1.004	0.998	0.997
Diferencia de promedios	1.24	1.132	0.805	0.946	1.06	1.135	1.136
Favorito	1.12	1.051	0.803	0.958	1.019	1.041	1.035

Figura 4.10: Comparación de resultados de cada criterio para los métodos utilizados en el año 2019.

4.2.1. 2018

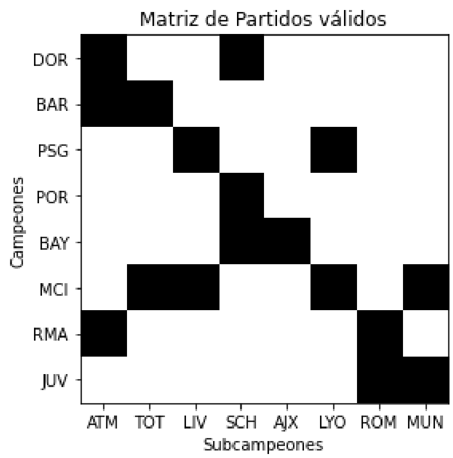

Figura 4.11: Tabla con todos los partidos posibles en 2018. Los recuadros negros representan los partidos que no pueden jugarse porque rompen alguna de las tres condiciones.

Hay un total de 3,694 fases de eliminación posibles.

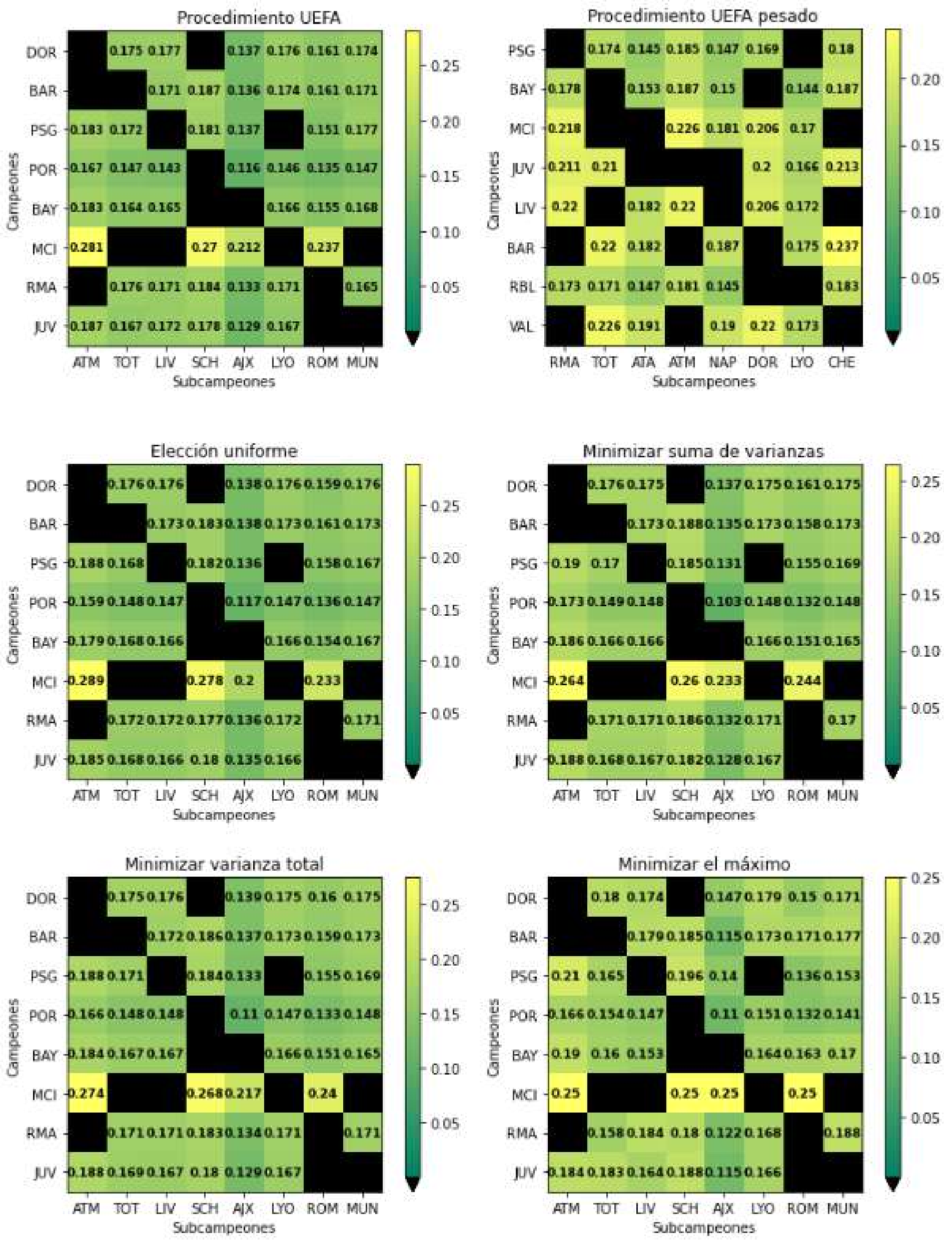

Figura 4.12: Tablas de probabilidades del año 2018 utilizando los métodos vistos en la sección 3.

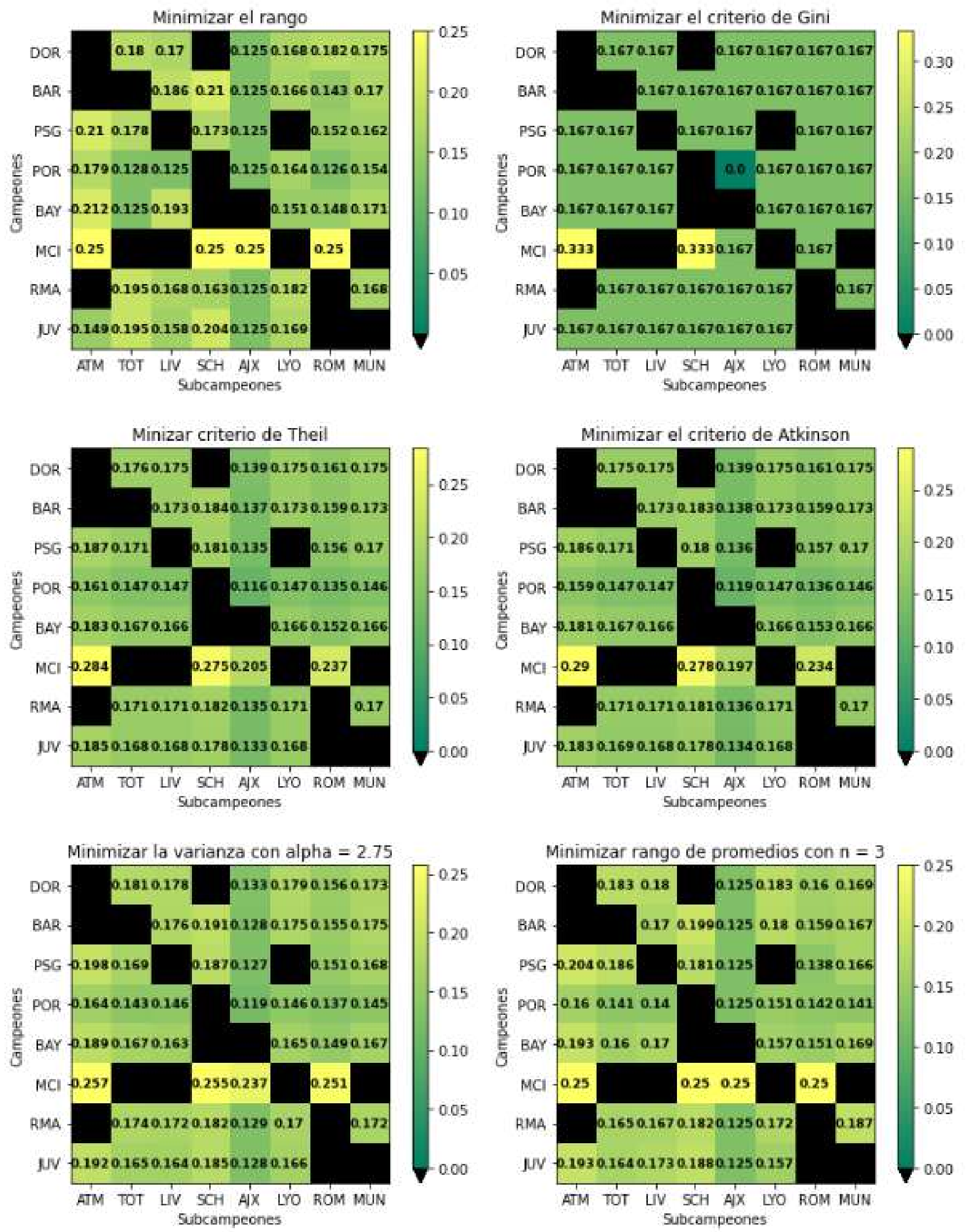

Figura 4.13: Tablas de probabilidades del año 2018 utilizando los métodos vistos en la sección 3.

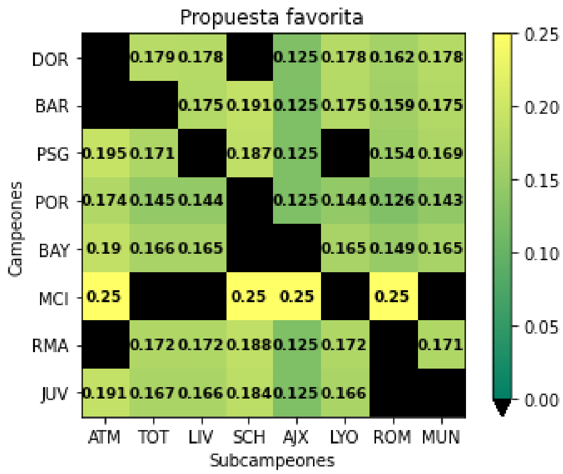

Figura 4.14: Tabla de probabilidades para el método favorito propuesto, del año 2018.

Procedimiento Realizado	Valor Suma de varianzas	Valor Varianza total	Valor Rango	Valor Máximo	Valor Gini	Valor Theil	Valor Atkinson
UEFA	1	1	1	1	1	1	1
UEFA Pesado	1.032	1.006	0.988	1.003	0.999	1.0	0.997
Uniforme	1.084	1.014	1.043	1.03	0.969	0.995	0.986
Minimizar Suma de varianzas	0.926	1.003	0.973	0.939	1.035	1.03	1.044
Minimizar Varianza total	0.956	0.991	1.0	0.978	1.003	1.0	1.004
Minimizar Rango	1.475	1.302	0.76	0.89	1.283	1.361	1.388
Minimizar Máximo	1.131	1.13	0.851	0.89	1.158	1.179	1.203
Minimizar Gini	2.191	1.886	2.025	1.187	0.687	2.449	3.843
Minimizar Theil	1.03	1.001	1.018	1.011	0.978	0.991	0.986
Minimizar Atkinson	1.093	1.015	1.039	1.032	0.963	0.993	0.984
Generalizar la Varianza	0.965	1.027	0.842	0.916	1.083	1.053	1.064
Diferencia de promedios	1.091	1.104	0.759	0.89	1.151	1.141	1.157
Favorito	1.002	1.06	0.76	0.89	1.102	1.098	1.115

Figura 4.15: Comparación de resultados de cada criterio para los métodos utilizados en el año 2018.

4.2.2. 2017

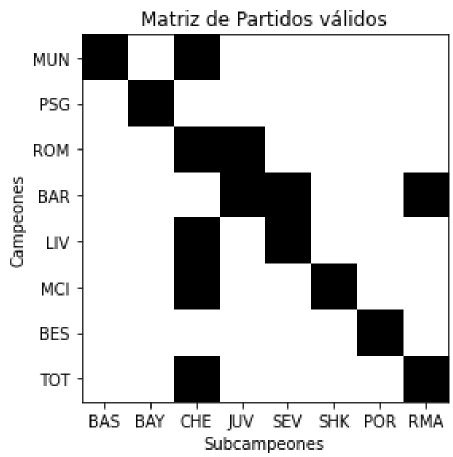

Figura 4.16: Tabla con todos los partidos posibles en 2017. Los recuadros negros representan los partidos que no pueden jugarse porque rompen alguna de las tres condiciones.

Hay un total de 4,238 posibles fases de eliminación.

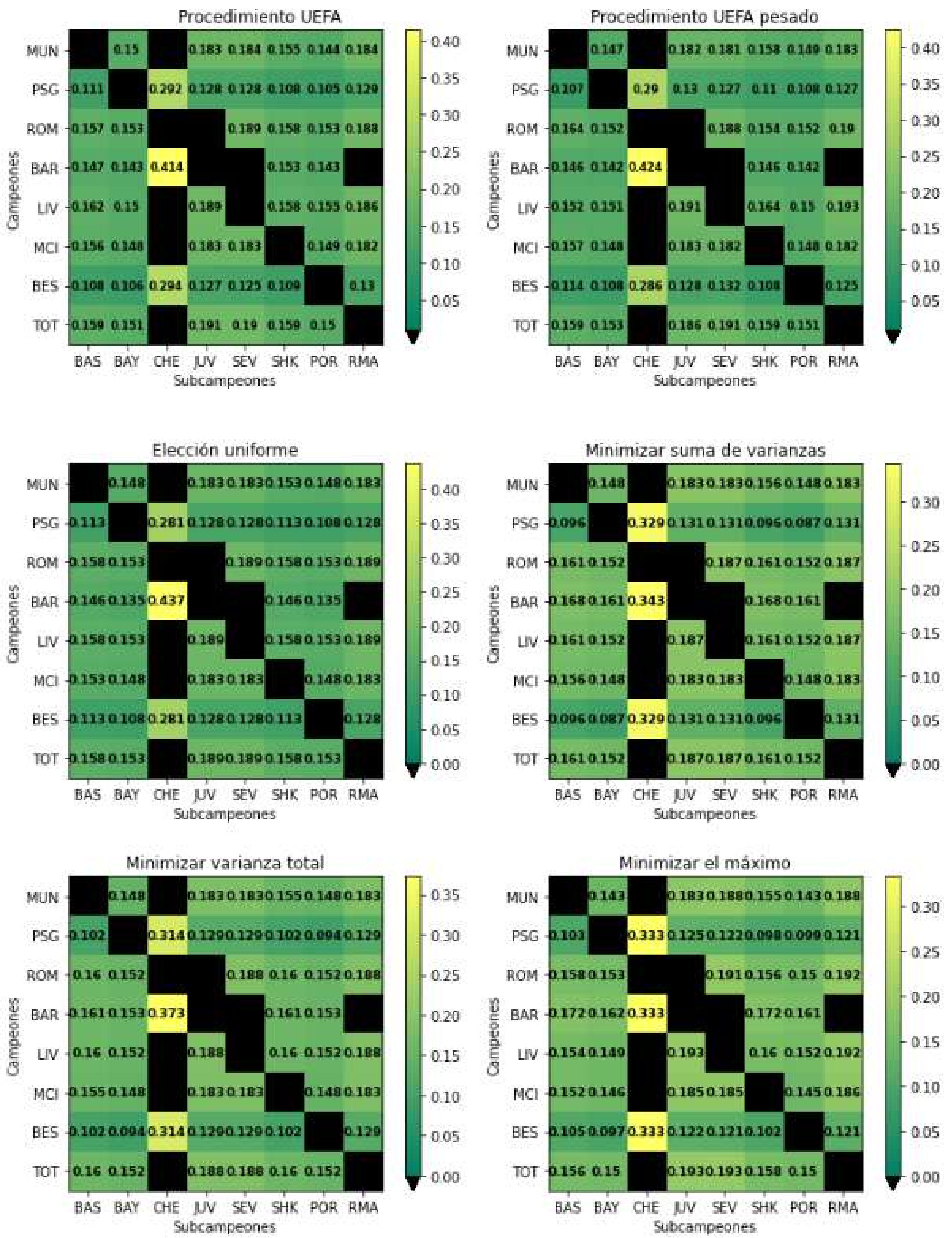

Figura 4.17: Tablas de probabilidades del año 2017 utilizando los métodos vistos en la sección 3.

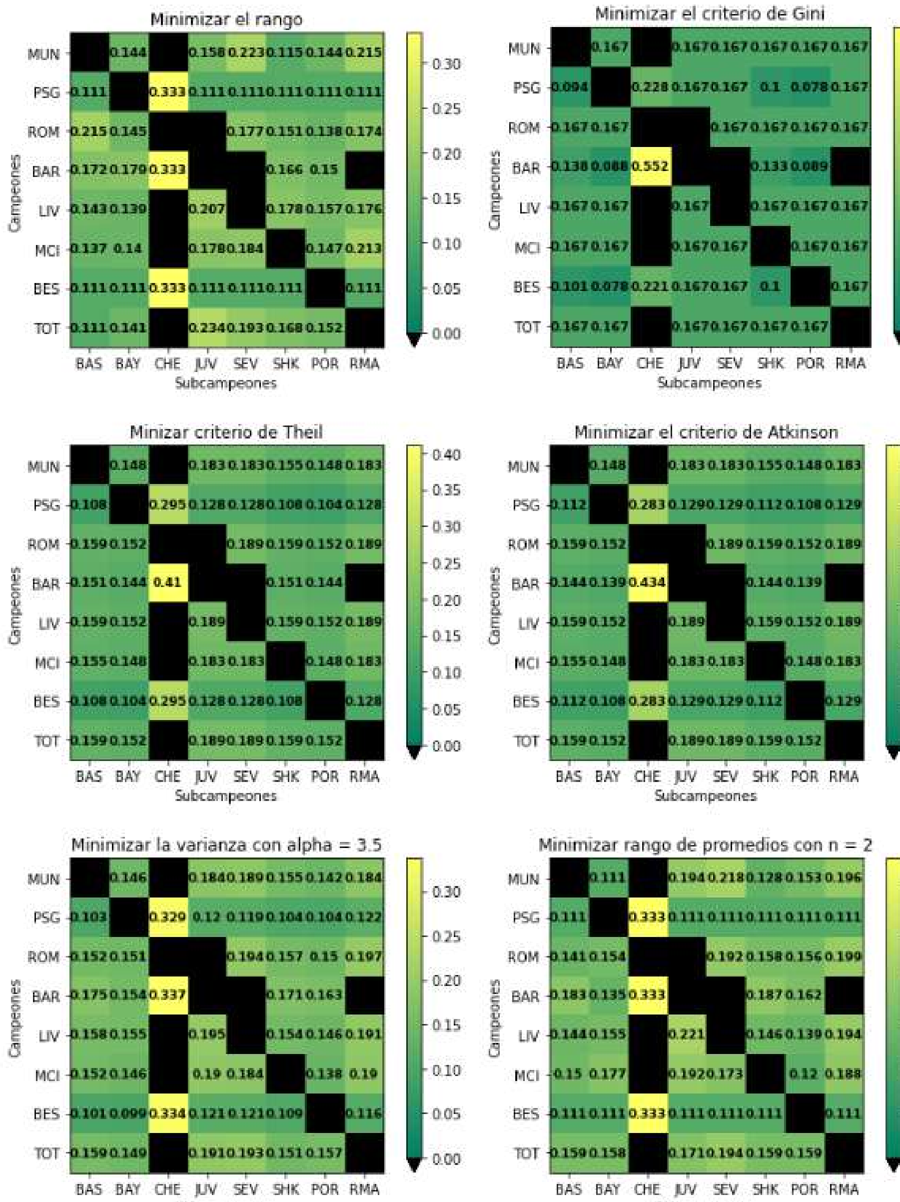

Figura 4.18: Tablas de probabilidades del año 2017 utilizando los métodos vistos en la sección 3.

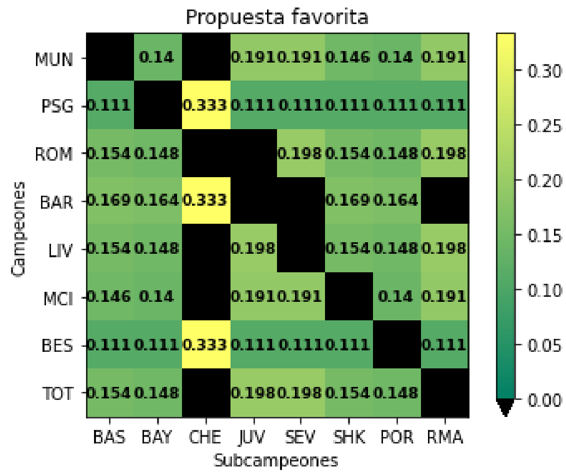

Figura 4.19: Tabla de probabilidades para el método favorito propuesto, del año 2017.

Procedimiento Realizado	Valor Suma de varianzas	Valor Varianza total	Valor Rango	Valor Máximo	Valor Gini	Valor Theil	Valor Atkinson
UEFA	1	1	1	1	1	1	1
UEFA Pesado	1.057	1.015	1.027	1.024	0.998	1.003	0.999
Uniforme	1.142	1.036	1.066	1.057	0.985	1.007	0.996
Minimizar Suma de varianzas	0.809	0.986	0.828	0.828	1.042	1.056	1.091
Minimizar Varianza total	0.842	0.973	0.901	0.9	1.018	1.016	1.037
Minimizar Rango	1.042	1.139	0.72	0.806	1.206	1.216	1.246
Minimizar Máximo	0.826	0.999	0.766	0.806	1.079	1.06	1.087
Minimizar Gini	2.51	1.533	1.535	1.333	0.912	1.386	1.36
Minimizar Theil	0.979	0.994	0.993	0.991	0.997	0.999	1.001
Minimizar Atkinson	1.116	1.028	1.056	1.048	0.987	1.005	0.995
Generalizar la Varianza	0.837	1.002	0.773	0.815	1.085	1.06	1.085
Diferencia de promedios	0.965	1.086	0.72	0.806	1.167	1.159	1.189
Favorito	0.878	1.029	0.72	0.806	1.106	1.087	1.111

Figura 4.20: Comparación de resultados de cada criterio para los métodos utilizados en el año 2017.

4.2.3. 2016

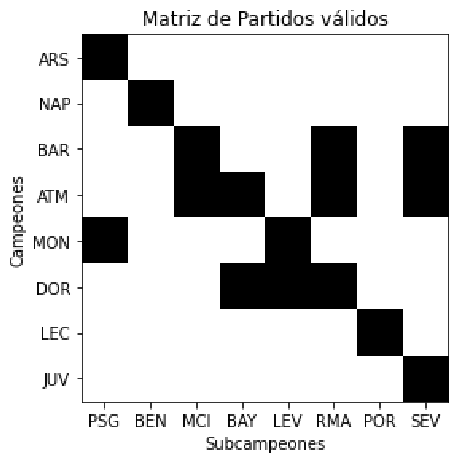

Figura 4.21: Tabla con todos los partidos posibles en 2016. Los recuadros negros representan los partidos que no pueden jugarse porque rompen alguna de las tres condiciones.

Hay un total de 3,715 posibles fases de eliminación.

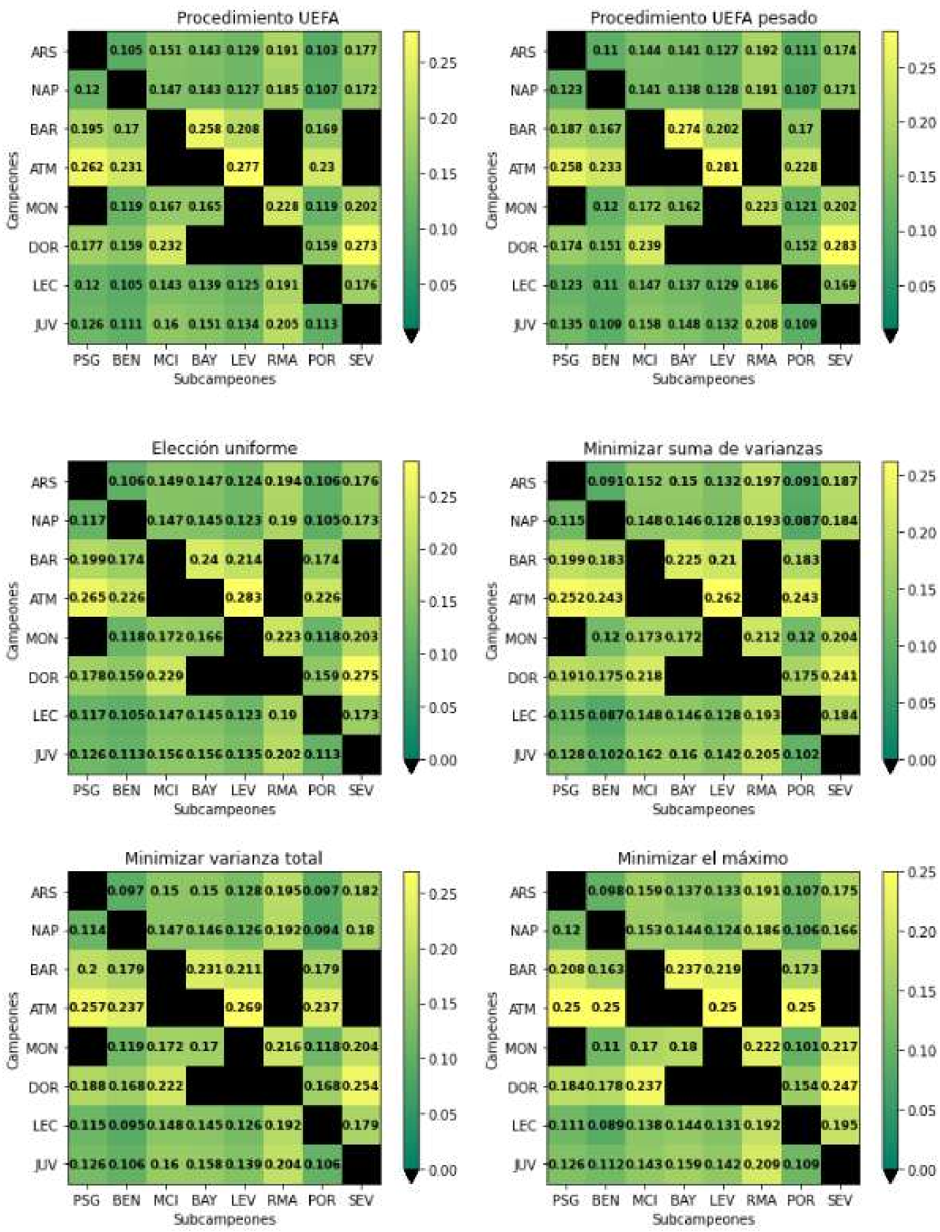

Figura 4.22: Tablas de probabilidades del año 2016 utilizando los métodos vistos en la sección 3.

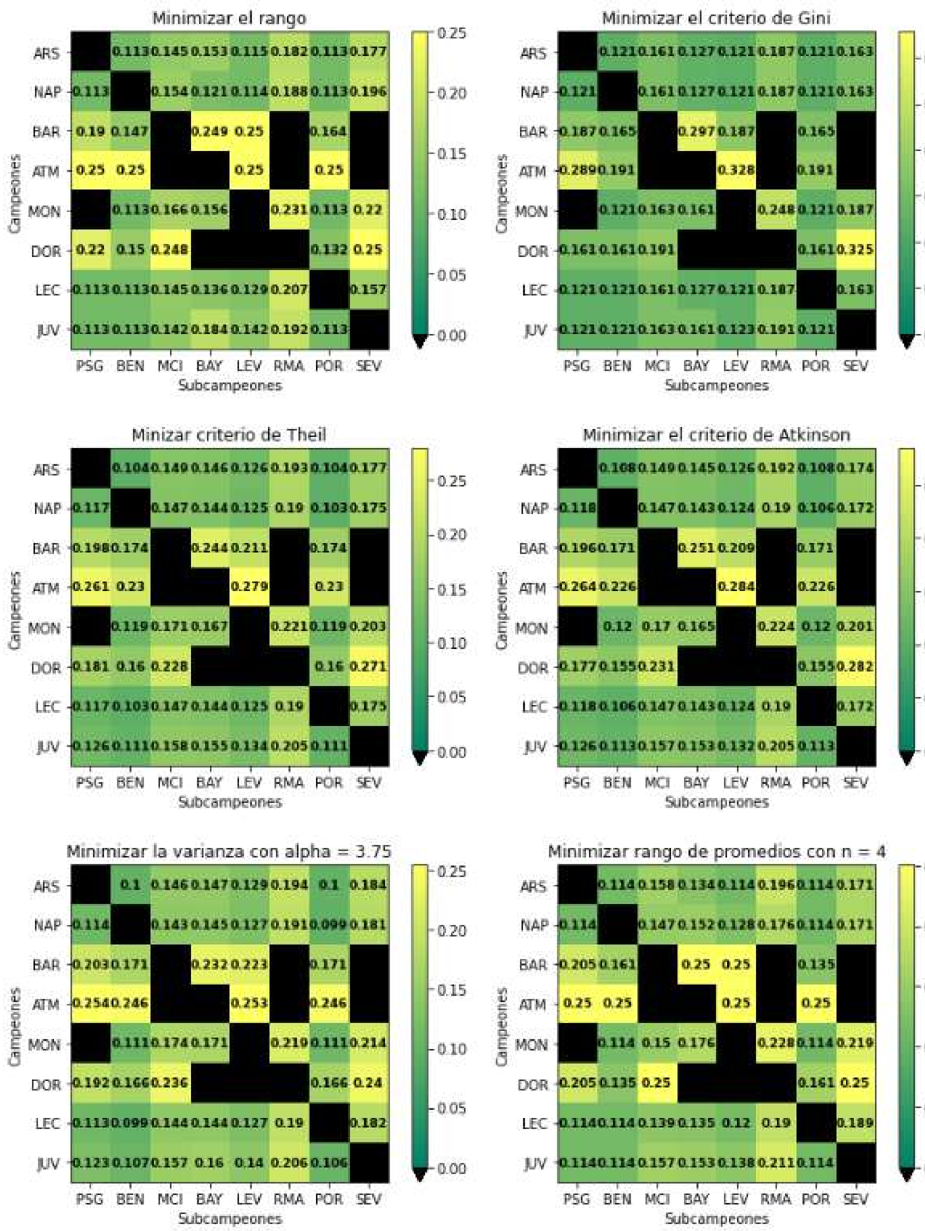

Figura 4.23: Tablas de probabilidades del año 2016 utilizando los métodos vistos en la sección 3.

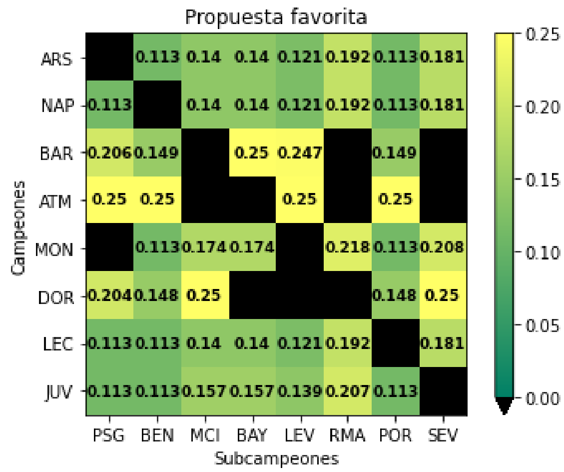

Figura 4.24: Tabla de probabilidades para el método favorito propuesto, del año 2016.

Procedimiento Realizado	Valor Suma de varianzas	Valor Varianza total	Valor Rango	Valor Máximo	Valor Gini	Valor Theil	Valor Atkinson
UEFA	1	1	1	1	1	1	1
UEFA Pesado	1.071	1.033	1.01	1.023	1.0	1.016	1.007
Uniforme	0.988	0.992	1.021	1.02	0.997	0.997	0.999
Minimizar Suma de varianzas	0.897	0.98	1.004	0.943	1.013	1.039	1.071
Minimizar Varianza total	0.911	0.972	1.006	0.972	1.007	1.009	1.028
Minimizar Rango	1.141	1.096	0.784	0.902	1.045	1.104	1.106
Minimizar Máximo	0.973	1.011	0.925	0.902	1.028	1.044	1.061
Minimizar Gini	1.532	1.247	1.19	1.183	0.981	1.149	1.107
Minimizar Theil	0.969	0.986	1.011	1.006	0.999	0.995	1.0
Minimizar Atkinson	1.02	1.005	1.02	1.025	0.996	0.999	0.996
Generalizar la Varianza	0.929	0.986	0.891	0.917	1.017	1.022	1.039
Diferencia de promedios	1.112	1.077	0.785	0.903	1.035	1.082	1.083
Favorito	1.066	1.051	0.784	0.902	1.023	1.058	1.06

Figura 4.25: Comparación de resultados de cada criterio para los métodos utilizados en el año 2016.

Viendo estos resultados obtenidos para 5 años distintos (más los del 2021), se confirman los puntos que se mencionaron al final del capítulo 3 sobre las ventajas y desventajas de los métodos. Pues en estas tablas se tienen resultados cualitativamente similares a los del 2021 y llevan a las mismas conclusiones.

Capítulo 5

Conclusión

Las conclusiones más importantes de este trabajo son:

- Las condición que impone la UEFA sobre la distinta nacionalidad de los equipos que se enfrentan es muy restrictiva y hace imposible mantener la equidad de los partidos, pues necesariamente habrán partidos más probables que otros.

- El problema fundamental para encontrar la tabla de probabilidades más equitativa no es tanto un problema computacional, sino que es un problema con la definición que se le quiera dar a equidad. Dada una definición de equidad (definida a partir de un criterio), es fácil encontrar la tabla de probabilidades que minimiza dicho criterio usando métodos de optimización computacionales. A veces hay muchas tablas que minimizan a un criterio y se puede aprovechar esto para intentar luego minimizar otro.

  Sin embargo, eso significa que si se definiera claramente qué es la equidad matemáticamente, optimizarla sería fácil y se terminaría con el problema. Sin embargo, como hay muchas posibles definiciones de equidad, el problema se puede resolver de varias formas distintas y ninguna es más correcta que las demás si no hay un consenso sobre la definición de equidad.

- Propusimos varios criterios para medir la equidad de una tabla de probabilidades y basados en ellos construimos muchos métodos para realizar sorteos. Cada método tiene sus ventajas y desventajas, tal como se discute al final del capítulo 3.

- Una parte fundamental de todos los métodos, excepto de los primeros 3, es la realización de cómo se puede utilizar el teorema de Birkhoff-von Neuemann para que, dada una tabla de probabilidades, se pueda construir un procedimiento para realizar un sorteo que dé como resultado a esa matriz. El sorteo resultante no es tan espectacular como el de la UEFA, ya que de una sola vez se conocen automáticamente todos los partidos. Sin embargo, se puede hacer más emocionante si el presentador revela los partidos poco a poco y mantiene la sorpresa.

- El método que consideramos mejor es el que llamamos “método favorito”, cuyo procedimiento se presenta en el capítulo 3. La ventaja de este método es que minimiza al rango y al máximo (tal como se puede ver en todas las tablas de los 6 años estudiados). Minimizar estas cantidades es muy importante, ya que es fundamental que no haya partidos que sean muy probables o muy poco probables. Además, este método después minimiza la varianza bajo las condiciones de seguir manteniendo el rango lo más bajo posible. De esta forma, nos aseguramos de que no haya valores muy altos o muy bajos y que todas las probabilidades estén distribuidas de la forma más uniforme posible.

- Finalmente, recalcamos que es imposible encontrar un método que sea totalmente equitativo bajo las condiciones de la UEFA y que lo único que se puede hacer es comparar qué tan equitativos son algunos métodos respecto a otros, pero la respuesta siempre dependerá del criterio que se use. No obstante, hemos encontrado que hay métodos que son superiores al de la UEFA en varios criterios para todos los años analizados y que entonces podrían ser más convenientes y llevar a competencias más equitativas.

Bibliografía

[1] UEFA. UEFA Champions League kernel description. `https://es.uefa.com/uefachampionsleague/`, 2022.

[2] Glenn Hurlbert. A short proof of the birkhoff-von neumann theorem. *Bioresource technology*, 143:360–368, 2013.

[3] Javier Girón Enrique Castillo. Problema del sorteo del torneo de campeones de la uefa. *Universidad Icesi*, 2016.