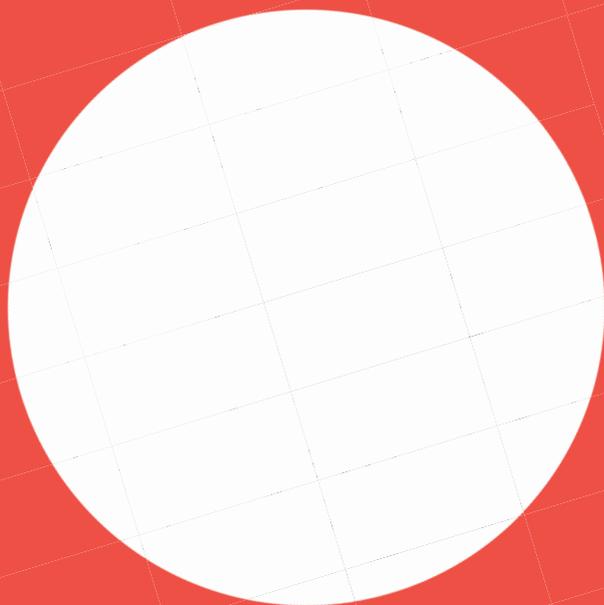


David Sumpter



FÚTBOL Y MATEMÁTICAS

**Aventuras matemáticas
del deporte rey**

Ariel

David Sumpter

Fútbol y matemáticas

Aventuras matemáticas del deporte rey

Traducción de Francisco García Lorenzana

Ariel

Título original: *Soccermatics. Mathematical Adventures in the Beautiful Game*

Esta traducción de *Fútbol y matemáticas*, 1.ª edición,
se publica en Editorial Ariel con el acuerdo de Bloomsbury Publishing Plc.

1.ª edición: junio de 2016

© 2016 David Sumpter

© 2016 de la traducción: Francisco García Lorenzana

Derechos exclusivos de edición en español
reservados para todo el mundo y propiedad de la traducción:
© 2016: Editorial Planeta, S. A. Avda. Diagonal, 662-664 - 08034 Barcelona
Editorial Ariel es un sello editorial de Planeta, S. A.
www.ariel.es

ISBN 978-84-344-2384-8
Depósito legal: B. 9.767-2016

Impreso en España por
Limpergraf, S. L.

El papel utilizado para la impresión de este libro
es cien por cien libre de cloro y está calificado como papel ecológico.

No se permite la reproducción total o parcial de este libro, ni su incorporación
a un sistema informático, ni su transmisión en cualquier forma o por cualquier medio,
sea éste electrónico, mecánico, por fotocopia, por grabación u otros métodos, sin el permiso
previo y por escrito del editor. La infracción de los derechos mencionados puede ser constitutiva
de delito contra la propiedad intelectual (Art. 270 y siguientes del Código Penal).

Dirijase a CEDRO (Centro Español de Derechos Reprográficos) si necesita
fotocopiar o escanear algún fragmento de esta obra.

Puede contactar con CEDRO a través de la web www.conlicencia.com
o por teléfono en el 91 702 19 70 / 93 272 04 47.

Índice

El saque inicial	9
----------------------------	---

PARTE I EN EL CAMPO

Capítulo 1. Nunca he predicho nada y nunca lo haré	21
Capítulo 2. Cómo el moho del cieno construyó Barcelona . . .	37
Capítulo 3. Comprueba mi ritmo	57
Capítulo 4. Brillantez estadística	79
Capítulo 5. La cohetería de Zlatan Ibrah	103

PARTE II EN EL BANQUILLO

Capítulo 6. Tres puntos para el entrenador con pájaros en la cabeza	115
Capítulo 7. El mapa táctico	135
Capítulo 8. Ciber-Dinamo total	163
Capítulo 9. El mundo en movimiento	187

PARTE III DESDE LA GRADA

Capítulo 10. <i>You'll Never Walk Alone</i>	215
Capítulo 11. Apostar contra las masas.	239
Capítulo 12. Pongo mi dinero donde tengo la boca	265
Capítulo 13. Llegan los resultados	293
El pitido final	315
<i>Agradecimientos</i>	319
<i>Índice analítico</i>	321

CAPÍTULO 1

Nunca he predicho nada y nunca lo haré

El centrocampista inglés Paul Gascoigne dijo una vez en 1996: «Nunca he predicho nada y nunca lo haré». Para mí, esta afirmación resulta tan genial como su gol en la Eurocopa de 1996 contra Escocia. En ocho palabras demuestra por qué las predicciones son inevitables: a las cuatro palabras ya estaba equivocado sobre el pasado y el presente, y después de las cuatro siguientes también se equivocaba sobre el futuro. Pero a pesar de estar tan equivocado, Gazza nos sigue diciendo algo importante. Resumía un hecho muy profundo sobre la vida: se pueden encontrar patrones en todo.

Hay patrones en lo que tardamos en llegar al trabajo por la mañana en hora punta. Hay patrones en nuestras redes de amigos y en la frecuencia con la que nos reunimos con ellos. Hay patrones en lo que comemos cada noche para cenar y en lo que compramos en el supermercado. Y también, por supuesto, hay patrones en el fútbol. El reto radica en encontrar dichos patrones y comprenderlos. En cuanto identificamos un patrón, podemos realizar una predicción.

Subbuteo al azar

Puedo remontar mi fascinación por los patrones a un gran libro naranja en tapa dura lleno de estadísticas de fútbol que me regalaron por Navidad cuando tenía ocho años. Me pasaba las horas mirando páginas llenas de números. Me gustaban las tablas que tenían los nombres de los equipos en la

parte de arriba y al lado izquierdo, y las entradas eran los resultados de los partidos que habían jugado entre ellos durante la temporada. Revisaba la tabla de arriba abajo, sumando los goles marcados y buscando partidos con resultados raros; 4-3 era mi favorito, y 5-2 también me sonaba bien.

En la actualidad no tengo demasiado tiempo para leer anuarios de fútbol, pero afortunadamente solo se tardan unos pocos segundos en encontrar en internet todos los resultados y las tablas. Si lo haces, puedes tener la sensación de impredecibilidad de la que hablaba Gascoigne. La temporada 2012/2013 de la Premier League es un buen ejemplo: hubo partidos bastante emocionantes y resultados inesperados. El Liverpool ganó dos veces por 5-0 y una vez por 6-0, pero no pudo clasificarse para Europa. La temporada terminó con la retirada de sir Alex Ferguson, el rey de los cambios de fortuna inesperados en el último minuto. Su último partido como entrenador del Manchester United no fue una excepción: un empate 5-5 en el que el West Bromwich Albion marcó tres goles en los últimos diez minutos. «¡Fútbol, maldita sea!», como lo resumió una vez Fergie.

Estos resultados fueron las excepciones excitantes de los partidos más memorables de la temporada. También hubo un buen número de aburridos empates 0-0, quizás olvidados por los aficionados, pero no por las estadísticas de la temporada. Pero si queremos comprender el patrón subyacente, también los tenemos que incluir en nuestro análisis. La figura 1.1 es un histograma del número de goles marcados en todos los partidos de la Premier League de la temporada 2012/2013. La media del número de goles marcados era ligeramente inferior a tres por partido: 2,79 para ser exactos.

Este histograma muestra la frecuencia con la que se dieron los resultados. En conjunto hubo 35 empates 0-0, que es la primera barra del histograma. El último partido de Ferguson fue uno de los dos de dicha temporada que terminó con 10 goles marcados, como se puede ver a la derecha. En el centro, el resultado más habitual fueron tres goles, y en la mayor parte de dichos partidos el resultado final fue 2-1. Ya empieza a aparecer un patrón. El siguiente paso es ver si podemos com-

prender de dónde surge este patrón, y para ello necesitamos un modelo matemático.

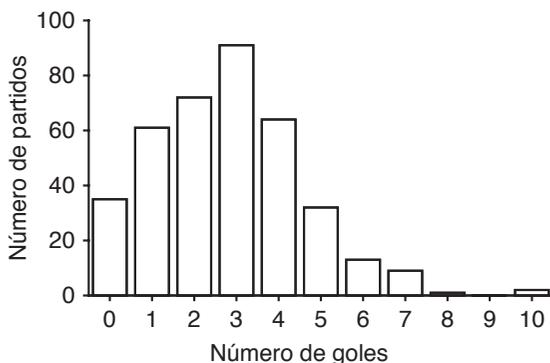


FIGURA 1.1. *Histograma del número de goles marcados durante la temporada 2012/2013 de la Premier League inglesa.*

Estoy interesado en los modelos matemáticos casi desde hace tanto tiempo como el que llevo interesado en las estadísticas. Mi otra gran afición de la época en que leía grandes anuarios de fútbol de color naranja era jugar al fútbol de mesa Subbuteo.¹ Con mi amigo David Paterson, organicé una liga de Subbuteo. Jugábamos cada día después de la escuela, completando cinco o seis partidos antes de cenar y anotando cada uno de los resultados. Pero nunca teníamos tiempo de completar los 380 partidos que formaban la competición de liga (20 equipos cada uno de los cuales juega 19 partidos en casa hacen $20 \times 19 = 380$ partidos). El día no tenía horas suficientes.

Obligados por unos padres que parecían creer que teníamos que hacer cosas como comer y dormir, Patzi y yo teníamos que encontrar una manera diferente de completar la liga. La respuesta estaba en los dados. Patzi lanzaba un dado por un equipo y yo lo lanzaba por el otro. Entonces tomábamos el número en cada uno de los dados para obtener el resultado.

1. Parece que el Subbuteo ha desaparecido completamente de nuestra cultura, aunque he oído que es posible que esté a punto de disfrutar de un renacimiento. Se trata de un juego de fútbol de mesa en el que se mueven unos jugadores en miniatura con base de metal sobre un campo de tela.

Así, si el Arsenal jugaba contra el Manchester City, él tiraba un dado rojo y yo tiraba uno azul. Si el dado rojo mostraba un cinco y el azul un 3, entonces el Arsenal ganaba 4-2. Este modelo puede generar partidos de 0 a 10 goles, como en el histograma de la Premier League.

Después de un montón de tiradas de dados, y algunos pequeños ajustes para favorecer a nuestros clubes favoritos, completábamos los resultados basándonos en los números que aparecían en los dados. Compilamos las ligas y las estadísticas y las apuntamos limpiamente en papel pautado. Creo que siempre estuve destinado a convertirme en matemático (y el otro David es en la actualidad un contable de éxito).

El lanzamiento de dados es un ejemplo muy sencillo de un modelo matemático, pero presenta algunos problemas. El Chelsea batió al Aston Villa por 8-0 poco antes de las Navidades de 2012, lo que no podía ocurrir en nuestro modelo de tirada de dados. Otro problema es que el empate 0-0 ocurre con mucha frecuencia en el fútbol real. Para los dados, un resultado de 0-0 es como un 5-5, pero en el histograma 0 goles es casi 20 veces más probable que 10 goles. El modelo no funciona. Los partidos de fútbol no tienen resultados al azar, como ocurre con la tirada de dados.

Pero los partidos de fútbol son aleatorios de otra manera. Lo que hace que el fútbol y otros deportes de equipo sean apasionantes es su impredecibilidad. Si estás mirando un partido y apartas la vista durante unos pocos segundos, te puedes perder una jugada importante y un gol inesperado. Como creador de modelos, esto me dice algo importante. Un gol se puede marcar en cualquier momento durante el partido. Mientras que existe toda una serie de factores que determinan la cadencia con la que marcan los equipos, el momento de los goles es más o menos aleatorio.

Esta suposición la podemos convertir en una simulación. Imagina que un partido de fútbol está formado por 90 períodos individuales de un minuto, en cada uno de los cuales existe la misma probabilidad de que se marque un gol. Con una media de 2,79 goles por partido, la probabilidad de que se marque un gol en cualquiera de estos períodos es de $2,79/90 = 0,031$. Esto

significa que las posibilidades de ver un gol en un minuto escogido al azar es de aproximadamente 1 de 32. No es muy probable, pero sí suficiente para estar seguros de que sigues mirando.

Utilizando este modelo podemos poner en marcha una simulación informática de 90 minutos, en la que en cada minuto simulado existe un 0,031 de probabilidad de que se marque un gol. Si mantenemos en marcha la simulación para un gran número de partidos, podemos descubrir cuál es el aspecto de una temporada típica. Esta temporada simulada se representa en la figura 1.2 como una línea, superpuesta sobre el histograma de la temporada real 2012/2013 de la Premier League.

La correspondencia entre el modelo y la realidad es muy buena. Recuerda toda la complejidad que entra en juego. Todos los gritos del entrenador desde la banda. Los aficionados intentando animar a su equipo o, con mayor frecuencia de la deseada, gritándoles lo inútiles que son. Los pensamientos en la cabeza de los jugadores mientras se dicen que ahora ha llegado la oportunidad de marcar. Ninguno de estos factores parece afectar la distribución de los goles marcados. Al contrario, son todos estos factores actuando juntos los que generan el tipo de aleatoriedad asumida en el modelo. A más factores implicados, mayor es la aleatoriedad de los goles, y mayor la igualdad de nuestro histograma simulado con la realidad.

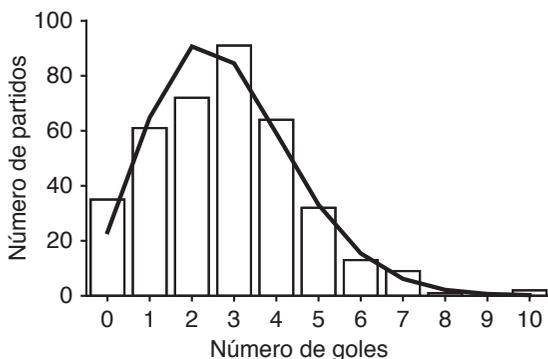


FIGURA 1.2 *El histograma del número de goles marcados durante la temporada 2012/2013 de la Premier League inglesa (barras del histograma) comparado con la distribución de Poisson (línea).*

La línea en la figura 1.2 generada por mi simulación se conoce como distribución de Poisson. Este tipo de distribución sugiere siempre que el momento en que se producen los acontecimientos anteriores no afecta a los eventos futuros. Esto es exactamente lo que asumí en mi simulación, y es lo que ocurre realmente en el fútbol: ni el número de goles marcados hasta el momento ni la cantidad de tiempo jugado influyen en la probabilidad de marcar otro gol. La distribución de Poisson resultante captura con bastante éxito la forma general del histograma de goles.² Los acontecimientos hacen que cada minuto de un partido de fútbol sea impredecible, y así aparece la distribución de Poisson. Se trata de un patrón que surge del puro azar.

No tomé la decisión de centrarme en la Premier League de fútbol porque supiera de antemano que seguiría el modelo de Poisson. Simplemente me gusta el fútbol. Podría haber escogido cualquier deporte en el que se pueden marcar goles en cualquier momento. Para asegurarme, comprobé todos los resultados de los partidos de hockey sobre hielo de la temporada 2012/2013 de la NHL. Hubo una media de 5,2 goles durante los 60 minutos de partido. La figura 1.3 muestra un histograma del número de goles en los 720 partidos que forman la temporada. La línea es la distribución de Poisson correspondiente.

La media de goles más alta mueve la cima del histograma hacia la derecha, pero la simulación vuelve a corresponderse

2. El modelo coincide bastante bien, pero existen diferencias entre la realidad y el modelo. El test estadístico de chi al cuadrado de Pearson se basa en

$$X^2 = \sum_{i=0}^{10} \frac{(O_i - E_i)^2}{E_i}$$

donde O_i es el número de partidos en los que se marcaron i goles y E_i es la predicción del modelo. La suma $X^2 = 26,3$, que es estadísticamente significativa en un nivel del 0,5 % con 10 grados de libertad. Este valor alto de X^2 aparece principalmente a causa de los dos partidos de 10 goles que se han visto durante la temporada. Según el modelo, partidos de 10 goles solo deberían ocurrir una vez cada cuatro años. Si agrupamos todos los partidos con 9 o más goles al comparar el modelo y los datos, entonces la suma $X^2 = 14,6$, que es significativa a un nivel del 10 % con 9 grados de libertad.

La otra desviación entre el modelo y los datos es para los empates 0-0. Volveré más tarde sobre este tema en el texto.

con los datos. Los datos y el modelo no son muy diferentes en su conjunto, y las pequeñas diferencias en cuanto a los partidos con cuatro goles se pueden explicar por las fluctuaciones de una temporada a otra.³ En el hockey sobre hielo se marcan más goles, pero no son más o menos aleatorios que en el fútbol.

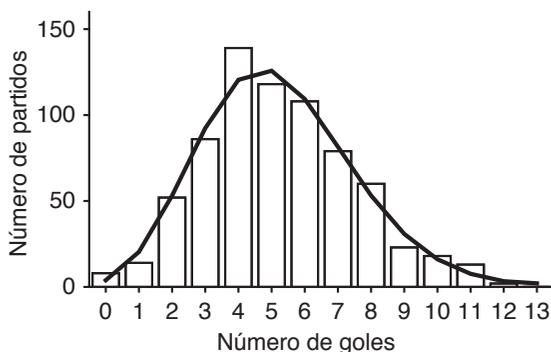


FIGURA 1.3. *Histograma del número de goles marcados durante la temporada 2012/2013 de hockey sobre hielo de la NHL (barras) comparado con la distribución de Poisson (línea).*

Recibir la coza de un caballo

Si empiezas a pensar en términos de simulaciones aleatorias y distribuciones de Poisson, entonces las verás por todas partes. En la asignatura de estadística de la licenciatura, la mejor (y la única) broma del profesor es que las llegadas del autobús están sometidas a la distribución de Poisson. La empresa de autobuses empieza con un horario, pero aparecen toda una serie de factores aleatorios: un anciano se toma su tiempo para subir al vehículo, o un ciclista ocupa el centro del carril bus. Otro ejemplo clásico es el número de bombillas que tienes que cambiar en casa cada año. Cada vez que enciendes la luz hay una pequeña posibilidad de que la bombilla se funda. Suma todos los fundidos y obtienes una distribución de Poisson.

3. La estadística para los datos de la NHL es $X^2 = 19,6$, que no es estadísticamente significativa por los 13 grados de libertad de los datos.

La distribución de Poisson recibe su nombre de Siméon Denis Poisson, un francés que fue el primero en describirla a principios del siglo XIX. No obstante, su descripción se centra en las ecuaciones matemáticas que subyacen a la distribución y no en cómo se puede utilizar para crear un modelo del mundo real. La primera aplicación en este sentido fue obra de un polaco, Ladislaus Bortkiewicz, que trabajaba en Alemania en 1898,⁴ analizando dos series de datos. La primera era una macabra serie estadística que encontró en el número de niños menores de 10 años que se suicidaron a lo largo de un período de 24 años. La segunda serie de datos, solo un poco menos inquietante, se refería a los soldados que habían muerto accidentalmente después de recibir una cox u otro tipo de golpe por parte de un caballo. Bortkiewicz analizó 14 regimientos diferentes a lo largo de 20 años, registrando cuántos soldados habían muerto de esta manera. Obviamente, no era consciente que solo unos pocos años antes se fundó la Liga Inglesa de Fútbol. Esta le podría haber proporcionado todos los datos que necesitaba, sin necesidad de sumergirse en las cifras de mortalidad alemanas.

En ambas series de datos, Bortkiewicz descubrió una correspondencia bastante buena con la distribución de Poisson. Las muertes causadas por las coces de los caballos eran poco habituales. De los 280 regimientos estudiados, en 144 no se produjo ningún fallecimiento. Pero en dos regimientos desafortunados hubo cuatro muertes en un solo año. Al comparar con la distribución de Poisson, Bortkiewicz pudo demostrar que estos regimientos no trataban a sus caballos necesariamente peor que cualquier otro regimiento: solo habían tenido mala suerte ese año. El fútbol puede o no ser más importante que la vida y la muerte, pero los tres se rigen por las mismas reglas.

La comparación con la distribución de Poisson es lo primero que hago cuando se me presenta una nueva serie de

4. Una historia completa de la obra de Bortkiewicz se puede encontrar en la StatProb Encyclopedia (statprob.com/enciclopedia/LadislausVanBortkiewicz.html). Su libro sobre la «ley de los pequeños números» y la aplicación de la distribución de Poisson se puede encontrar en el original alemán en la California Digital Library (archive.org/details/dasgesetzderlei00bortrich).

datos. A veces un colega entra en mi despacho con resultados experimentales que acaba de reunir. «Resulta extraño», dice. «La mayor parte de los peces nunca nadan cerca de un depredador, pero hay un pez que ha pasado cuatro veces a su lado. Debe de ser un tipo de personalidad atrevida o algo por el estilo.» Tres minutos después estoy calculando la distribución de Poisson y superponiéndola a los datos de mi colega. «No, tu pez no es especialmente atrevido», le explico. «Solo era una necesidad estadística.» Que un depredador te persiga una y otra vez es como recibir una paliza de 5-0. Resulta terrible cuando ocurre, pero le puede pasar a cualquiera.

La distribución de Poisson es nuestro primer ejemplo de una analogía matemática. Funciona en muchos contextos. Funciona para los partidos de fútbol, funciona para las bombillas y funciona para las muertes por culpa de los caballos. Siempre que sea razonable asumir que los acontecimientos pueden ocurrir de manera inesperada, en cualquier momento, independientemente de cuántos eventos hayan tenido lugar antes del siguiente, entonces es razonable esperar una distribución de Poisson.

Muy lejos del fútbol, la mayoría de las aplicaciones modernas de la distribución de Poisson siguen la tradición iniciada por Bortkiewicz. Parece que los estadísticos sienten una fascinación perversa por la muerte, las heridas y los accidentes. O posiblemente solo sea que les pagamos para que analicen las cosas malas que nos pueden ocurrir, para que no tengamos que pensar en ellas. Sea cual sea la razón de su interés por las desgracias, los estadísticos han encontrado distribuciones de Poisson en los accidentes de coche, colisiones de camiones, heridas en la cabeza, averías en los motores de los aviones, bancarrotas, suicidios, asesinatos, accidentes laborales y los numerosos peligros de una obra de construcción.⁵ Incluso la han encontrado en el número de guerras iniciadas entre 1480 y 1940. Y cuando han acabado con la muerte y las heridas, encuentran la distri-

5. Algunos de estos ejemplos se presentan con más detalle en Letkowsky, J., 2012, «Application of Poisson probability distribution», *Proceedings of the 2012 Academic and Business Research Institute Conference*, San Antonio.

bución de Poisson en los errores de imprenta, los defectos de fabricación, los fallos de la red, los ataques de virus informáticos y los divorcios. Ya sea que se trate de muerte o destrucción, de mala suerte o errores, siempre se puede encontrar el mismo patrón de aleatoriedad.

En 2015, Cristian Tomasetti, un matemático aplicado, y Bert Vogelstein, un médico, utilizaron un argumento estadístico para demostrar que las dos terceras partes de los casos de cáncer se deben a la «mala suerte».⁶ Aunque ciertos cánceres se pueden relacionar con el estilo de vida, por ejemplo el cáncer de pulmón y fumar, solo son una parte de la historia. La parte más importante tiene que ver con la inevitable división celular que tiene lugar en nuestro cuerpo. Cada vez que se divide una célula, existe una posibilidad muy pequeña de una mutación genética que puede provocar un cáncer. Lo que descubrieron Cristian y Bert fue que en las partes del cuerpo en las que las células se dividen con más rapidez es más probable que se desarrolle un cáncer, y llegaron a la conclusión de que el cáncer se explica principalmente por estas mutaciones aleatorias.

Este estudio provocó algunas controversias. Si el cáncer solo aparece al azar, entonces ¿por qué debemos gastar tanto dinero para investigar sus causas? Para justificar el uso del término «mala suerte» y con el objetivo de explicar mejor sus conclusiones, Cristian y Bert usaron una analogía con los accidentes de tráfico. Cuanto más tiempo pasas en el coche conduciendo, afirmaron, más probable es que te veas envuelto en un accidente. Cómo conduces el coche es un factor, pero también es importante el tiempo que pasas al volante.

Una analogía futbolística también viene al caso e incluso lo explica un poco mejor. Puedes pensar en cada división celular en tu cuerpo como el equivalente a un minuto concreto de un partido de fútbol. Cuando una célula se divide, existe una probabilidad (muy) escasa de una mutación cancerígena aleatoria, de la misma manera que existe la probabilidad (mucho mayor) de que te marquen un gol en un partido de fútbol. En

6. Tomasetti, C., y Vogelstein, B., 2015, «Variation in cancer risk among tissues can be explained by the number of stem cell divisions», *Science* 347(6217): 78-81.

este sentido se puede pensar en el cáncer como mala suerte. A veces nuestro equipo acaba el partido sin que le hayan marcado un gol y con suerte pasamos por la vida sin padecer un cáncer. Y como a veces perdemos porque el contrincante era bueno, nadie puede negar que la suerte juega un papel importante en cualquier partido. Nuestra salud es como un sábado por la tarde mirando desde las gradas, y no todos los goles son evitables.

No todo lo que nos ocurre se debe al azar. Muchas enfermedades son prevenibles si tenemos un estilo de vida sano, y que te marquen goles con frecuencia se debe a una mala defensa. Pero ser conscientes de que mucho de lo que nos ocurre es aleatorio, a veces nos puede ayudar a aceptar los retos que nos plantea la vida. No todo en la vida es predecible.

Explicado por el azar

La impredecibilidad de un partido de fútbol de un minuto al siguiente genera la distribución de Poisson después de 90 minutos. Conocemos el número medio de goles marcados en un partido, pero el momento en que se marcan es impredecible. Como consecuencia, algunos resultados se vuelven mucho más probables que otros. La paradoja es que los resultados se explican mediante el azar. El hecho de que los goles sean muy aleatorios en el tiempo hace que el patrón de los resultados sea predecible. Resulta una idea difícil de aceptar, pero es cierta. A menudo, el propio hecho de que algo sea extremadamente aleatorio nos ayuda a explicarlo y a predecir la frecuencia con la que ocurrirá. El azar nos permite realizar todo tipo de predicciones sobre el futuro.

Los matemáticos utilizan continuamente este truco. Al principio de una nueva temporada de fútbol, o en los prolegómenos de una Copa del Mundo o de los Oscars, los periódicos publican con frecuencia historias sobre un «genio» matemático que ha predicho la probabilidad de victoria de unos equipos o de unas películas en particular. Con frecuencia estas predicciones parecen aleatorias y a veces son correctas. Pero ¿de dónde salen?

Os voy a explicar un secreto. Normalmente estos genios están haciendo algo muy sencillo con la distribución de Poisson y con un poco de información sobre los equipos o películas que compiten. Un truco para crear modelos de los resultados de fútbol es calcular la tasa de goles marcados y la tasa de goles encajados por cada equipo y después simular partidos entre ellos. Por ejemplo, durante la temporada 2012/2013 de la Premier League, el Arsenal marcó una media de 2,47 goles cuando jugaba en casa y 1,32 cuando jugaba fuera. Le marcaban una media de 1,21 goles en casa y 0,74 fuera. Al reunir estadísticas similares para cada equipo y después simular los partidos entre cada pareja, podemos generar predicciones para la temporada siguiente. Un ejemplo de dichas predicciones se presenta en la tabla 1.1, en la que he utilizado los datos de la temporada 2012/2013 y un modelo para predecir los cuatro primeros de la temporada 2013/2014.⁷

Esta predicción no se aleja demasiado de lo que ocurrió en realidad. En el mundo real, el Manchester City fue el campeón, dos puntos por delante del Liverpool y el Chelsea quedó tercero. Pero este solo es uno de los muchos posibles cuatro primeros simulados que se obtienen como resultado cuando aprieto «enter» en mi ordenador. Cada vez que activo la simulación, los equipos se encuentran tanto en casa como fuera y los resultados se generan al azar, con una media basada en sus tasas de marcar y encajar goles, y compilo una tabla de la liga basada en los resultados. Cada activación da resultados diferentes, algunos muy diferentes. La tabla 1.2 es otro ejemplo.

7. En este modelo utilizo cuatro parámetros para cada equipo: el número medio de goles marcados en casa (M_C), goles encajados en casa (E_C), goles marcados fuera de casa (M_F) y goles encajados fuera de casa (E_F). Se trata de estimaciones a partir de los goles marcados y recibidos en la temporada 2012/2013. Cuando se encuentran dos equipos en la liga en mi temporada 2013/2014 simulada, primero genero goles para el equipo de casa. Estos goles están distribuidos según Poisson con una media igual a $\frac{1}{2}(M_C + E_F)$, que tiene en cuenta tanto la fuerza del ataque del local como la fuerza defensiva del visitante. Los goles del visitante se determinan mediante una distribución Poisson con una media igual a $\frac{1}{2}(E_C + M_F)$. El mismo procedimiento se repite en casa y fuera para todos los equipos para completar una temporada simulada.

TABLA 1.1. *Los cuatro primeros equipos después de la primera simulación de la temporada 2013/2014, basada en las tasas de goles de los clubes durante 2012/2013.*

<i>Equipos</i>	<i>P</i>	<i>G</i>	<i>E</i>	<i>P</i>	<i>F</i>	<i>C</i>	<i>Ptos.</i>
Manchester City	38	22	7	9	71	42	73
Liverpool	38	22	5	11	64	43	71
Chelsea	38	21	5	12	74	51	68
Manchester United	38	19	7	12	61	45	64

TABLA 1.2. *Los cuatro primeros equipos después de la segunda simulación de la temporada 2013/2014, basada en las tasas de goles de los clubes durante 2012/2013.*

<i>Equipos</i>	<i>P</i>	<i>G</i>	<i>E</i>	<i>P</i>	<i>F</i>	<i>C</i>	<i>Ptos.</i>
Liverpool	38	23	7	8	68	37	76
Chelsea	38	22	8	8	75	52	74
Manchester United	38	22	5	11	72	43	71
Manchester City	38	19	8	11	64	42	65

Como seguidor del Liverpool, ¡esta me gusta mucho más! Representa una realidad alternativa en la que Steven Gerrard no cayó en el partido crucial contra el Chelsea y el Liverpool consiguió ganar su primer título de liga en casi 25 años. Es posible que Gerrard hubiera utilizado su energía positiva para que Inglaterra ganase la Copa del Mundo y lo hubieran nombrado caballero como sir Stevie G. Como existen un montón de realidades alternativas simuladas, resulta normal que escogiese la que más me gusta.

Desgraciadamente, el científico objetivo que hay en mí siente que debe informar de los resultados completos de todas las simulaciones. Mi portátil tarda un par de minutos en jugar 10.000 veces la Premier League, y cada vez el resultado es algo diferente. Por muy interesantes que puedan ser estas realidades alternativas, individualmente no tienen importancia. Lo importante es resumir lo que ocurre en las 10.000 veces. ¿Con

qué frecuencia ganan la liga los diferentes equipos? Cuando hacemos esto, vemos que el Liverpool solo la ganó en el 11,5 % de las simulaciones. El Manchester United, que había ganado el título en la temporada anterior, ganó el 26,2 % de las veces. El Chelsea ganó el 19,2 %, el Arsenal el 17,6 %, el Manchester City el 12,8 % y el Tottenham Hotspur el 6,0 %.

En retrospectiva, podemos ver que estas predicciones estaban un poco desajustadas. El Manchester United cambió de entrenador y realizó una temporada terrible. El Manchester City y el Liverpool dominaron, y los dos equipos marcaron más de 100 goles. Pero esto no es lo importante. Desde luego, no voy a pretender que ya he creado el mejor modelo para el fútbol. Solo estamos al principio de nuestra historia y no les voy a explicar de buenas a primeras todos mis trucos para crear modelos.

Lo importante es que, aunque no es totalmente cierto, el modelo basado en el azar tampoco es completamente erróneo. Los equipos que se predicen como los ganadores probables de la liga son los que habitualmente lo hacen bien, y las tablas de la liga que he mostrado parecen resultados potencialmente probables de una temporada, o al menos no son demasiado diferentes de lo que podríamos esperar. Y lo hemos conseguido sin un razonamiento real. Solo hemos simulado goles que se marcan aleatoriamente, con cada equipo con una tasa de marcaje diferente, y el resultado fueron unos cuatro primeros que parecían razonables. Esto es casi lo contrario de la imagen que daba Paul Gascoigne de un fútbol impredecible. El fútbol es muy predecible. Cada fin de semana durante la temporada de la Premier League, más de 400 jugadores pasan 90 minutos corriendo y dando patadas a un balón, y al final de la temporada gana un club grande de Londres o Manchester.

La predicción basada en el azar es una parte importante de cómo se utilizan actualmente las matemáticas en la sociedad. Cuando está esperando al teléfono, un analista ya ha estudiado el ritmo con el que las llamadas entran en la centralita y ha deducido cuanto tiempo está dispuesta la gente a esperar en línea. Cuando el banco presta dinero a un pequeño negocio o concede una hipoteca a un particular, ya ha estudiado la

probabilidad de impago y ha aplicado la distribución de Poisson para descubrir a cuántos impagos tendrán que hacer frente en los próximos años.

La predicción no consiste en decir exactamente qué club ganará la liga, exactamente cuánto tiempo tendrá que esperar al teléfono o qué empresa va a quebrar. Se trata de utilizar la frecuencia de los acontecimientos del pasado para calcular las probabilidades de los acontecimientos en el futuro. Todas estas predicciones surgen de un modelo matemático basado originalmente en los soldados alemanes coceados por caballos. Si prefiere una analogía verbal, se podría decir que esperar que el Liverpool marque es lo mismo que esperar que llegue el autobús número 19 en un lunes festivo: no ves nada durante una eternidad y después llegan dos o tres al mismo tiempo. A través del modelo he conseguido que esta analogía sea útil. Las matemáticas nos permiten descubrir los rasgos que tienen en común las llegadas de los autobuses, los partidos de fútbol, las quiebras, los casos de cáncer y las llamadas telefónicas. Y nos permiten predecir con qué frecuencia tendrán lugar.

La historia real

Aunque los goles se marquen al azar, las matemáticas pueden encontrar un camino para realizar predicciones. Pero Gascoigne tiene algo de razón. Las historias reales en el fútbol no tienen nada que ver con el azar, sino que se refieren a situaciones por encima de la aleatoriedad. Se trata de los fracasos y las recuperaciones. Cuando sir Alex Ferguson se retiró en 2012, David Moyes dirigió al Manchester United en su peor temporada en más de 20 años, y esto no puede explicarse por una racha de mala suerte. Cuando Alemania destrozó a Brasil con cinco goles en 18 minutos en la semifinal de la Copa del Mundo de 2014, no se trató simplemente de una secuencia aleatoria de goles. Brasil desapareció bajo la presión y Alemania se aprovechó de ello.

El éxito de Fergie o de la selección alemana de fútbol no se puede explicar en términos aleatorios: hay que pensar en su

trabajo interno. La ironía es que los acontecimientos que no son aleatorios son los más difíciles de comprender y más difíciles de predecir, pero también son mucho más interesantes.

En mi trabajo de investigación es la falta de aleatoriedad lo que plantea los retos más importantes. Mi colega biólogo vino a verme unas semanas después y me dijo: «Cuando no hay ningún depredador cerca, los peces se distribuyen al azar, pero cuando ven un depredador forman un grupo compacto que gira sobre sí mismo». Ahí hay un verdadero misterio. ¿Un solo pez inicia la rotación? ¿A qué velocidad giran, y algunos peces prefieren posiciones concretas? ¿Por qué es el grupo compacto y en rotación la mejor formación para evadir a un depredador? Las preguntas adquieren interés cuando falla el modelo aleatorio.

A medida que profundice en la creación de modelos en los capítulos siguientes, los problemas que analizaré serán menos aleatorios. Los movimientos de los jugadores están muy sincronizados, su red de pases está estructurada, el balón se mueve según las leyes de la física y los entrenadores piensan estratégicamente en la táctica. Los modelos que veremos serán muy diferentes, pero el enfoque básico que adoptaré será siempre el mismo. Realizo observaciones, y estas me permiten formular una serie de supuestos. Convierto estos supuestos en ecuaciones y las investigo utilizando simulaciones informáticas y soluciones matemáticas. Después comparo las propiedades del modelo con los datos del mundo real.

El reto para un matemático aplicado es la elección del modelo correcto para la cuestión que le interesa. Si estamos interesados en predecir el número de goles a lo largo de una temporada, entonces el azar con frecuencia resulta suficiente. Pero si queremos comprender las formaciones, los movimientos y las habilidades, entonces necesitamos comprender la estructura. Personalmente no estoy satisfecho con la explicación aleatoria: quiero descubrir qué ocurre en realidad. Para ello necesito acercarme a los jugadores y observar con atención lo que hacen. Y eso es exactamente lo que haremos a continuación.