

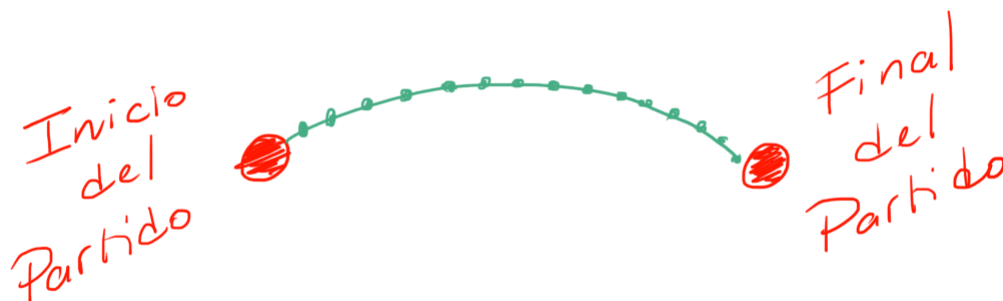
La falsa irregularidad del futbol mexicano, Parte 2

En el 2013 escribí mi primer artículo sobre futbol y confieso que, aunque me parecía bueno y original, me dejó insatisfecho y en ese momento no supe por qué. Ocho años después sigo insatisfecho, sin embargo, creo que he logrado prepararme lo suficiente como para animarme a escribir de nuevo desde un ángulo distinto, el cual, confío en que les será interesante.

Para construir el siguiente estudio y los siguientes que publicaré pedí permiso a Joseph Buchdahl de Football-data.co.uk para emplear sus datos sobre resultados deportivos y apuestas vinculadas a ellos. Se trata de una base de datos de alrededor de 150,000 partidos de 28 ligas europeas entre 2000 y 2021, a lo que añadí resultados y datos de la LigaMX. Por otro lado, la idea de este texto me saltó en la mente después de leer a Nassim Nicholas Taleb y dejar marinar en mi cabeza sus ideas durante un largo rato.

De lo que me gustaría hablar es sobre lo que hacemos cuando recordamos un partido de futbol, un ejemplo sencillo podría ser que recordamos el primer gol, el del empate, la tarjeta roja, la ocasión fallada o el penal marcado en el último minuto. En general, cuando recordamos algo tendemos a recordarlo como una sucesión de eventos seguidos unos de otros, como si se tratase de una sucesión de causas y efectos determinados, de ahí que, cuando recordamos algo como un partido de futbol, tendamos a ver las cosas de manera obvia después de que ocurrieron, porque solo contamos con esa única muestra. Dicho de otra forma: tendemos a cometer el error de pensar que los partidos solo pudieron haber ocurrido de la manera en que ocurrieron, lo que induce a confundir el todo con lo particular.

Si, para describir lo anterior de una manera gráfica, empleamos la siguiente imagen:



En esencia, lo que haremos en esta reflexión será utilizar la tecnología para construir una simulación que nos provea de una muestra lo suficientemente grande como para poder hacer observaciones y, para construir la simulación, emplearemos el mecanismo de las apuestas.

¿Qué es un momio y cómo transformarlo a una probabilidad implícita?

Un ejemplo clásico de una apuesta en el futbol (en el formato de Europa Continental) para un partido es el siguiente:

Victoria del Local	Empate	Victoria del Visitante
2	3.3	4


Lo expresado en la primera fila de la tabla representa el resultado final del partido y, por su parte, los números de la segunda fila son Los Momios o multiplicadores que establecen el pago de una apuesta exitosa. Si una persona apuesta \$100.00 a que, al final del partido el equipo local ganará, de suceder, su pago será:

$$2 \times \$100.00 = \$200.00$$

En caso de que el resultado del partido fuese distinto a la victoria del local el apostador perdería su dinero. Sin embargo, El Momio no es solo un multiplicador, pues también representa una probabilidad implícita. Para expresar un Momio en términos de probabilidad se requiere realizar la siguiente operación:

$$\text{Probabilidad} = \frac{1}{\text{Momio}}$$

Que, aplicado al ejemplo anterior, quedaría así:

Victoria del Local	Empate	Victoria del Visitante
2	3.3	4
		
Victoria del Local	Empate	Victoria del Visitante
0.5 (50%)	0.3 (30%)	0.25 (20%)

Nota: el ejemplo previo está simplificado, pues no incluye el margen de las casas de apuestas que funciona, básicamente, como una comisión. Esto será explicado en el siguiente artículo.

Veamos un caso real en el partido Liverpool vs Chelsea de la EPL que ocurrió el 4 de marzo de 2021 expresado en probabilidades y añadamos el resultado:

Liverpool	Empate	Chelsea	Goles de Local	Goles de Visitante
42%	26%	32%	0	1

Aquí viene la pregunta de los 64,000: ¿cuántos partidos hubo, entre 2012 y 2021, en los que los equipos locales tuvieron 42% de probabilidad de ganar? Para ser exacto: 2484. Lo que acabamos de hacer, aunque parezca muy sencillo, es un experimento con un enorme potencial para observar muchos fenómenos de lo más interesantes. Pero vayamos poco a poco y hoy enfoquémonos en un tema de discusión omnipresente en el futbol: la cantidad de goles.

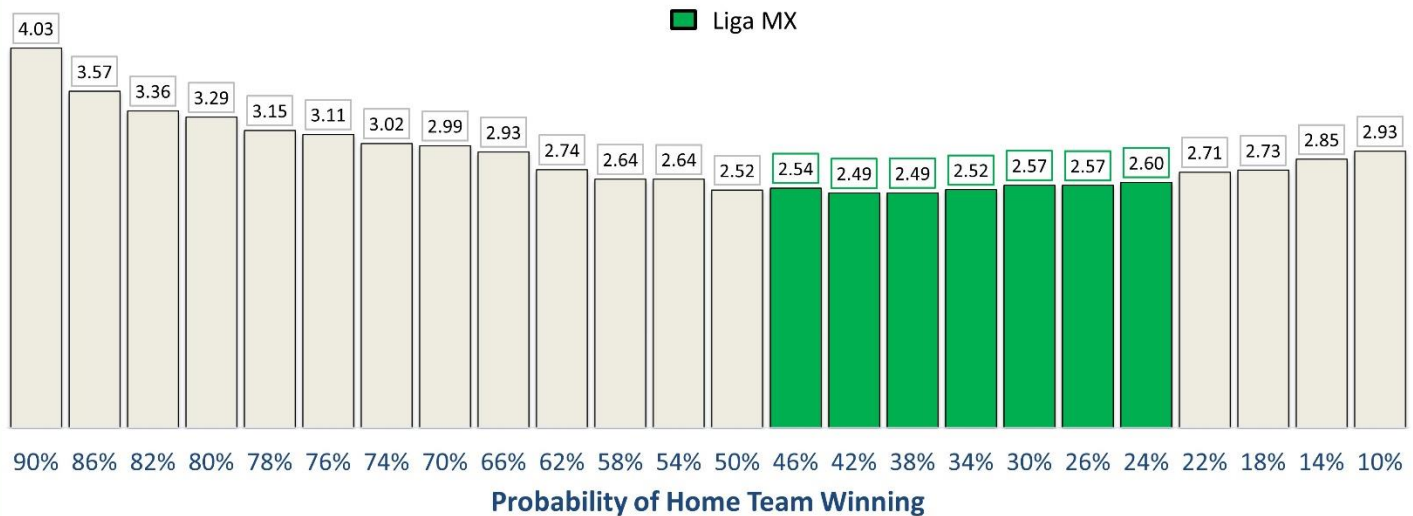
¿Cuál fue el promedio de goles de todos los partidos en ese *escenario* en particular? Habiendo anotado un total de 6185 goles en 2484 partidos el promedio fue de 2.49 goles por partido. Ahora repitamos este mismo ejercicio y hagamos la misma operación para todos los escenarios de local en el periodo analizado, es decir, a partir de la referencia que es la probabilidad de victoria de local, agrupemos, por ejemplo, todos los partidos que se consideró que el local tuvo una probabilidad de victoria del 70% y, si promediamos la cantidad de goles anotados en ese escenario obtenemos que el promedio es de 2.99 goles por partido... esto es interesante, ¿acaso los escenarios influyen en la cantidad de goles? Veamos cómo en la siguiente gráfica:

Expected Goals per Match per Probability of Home Team Winning



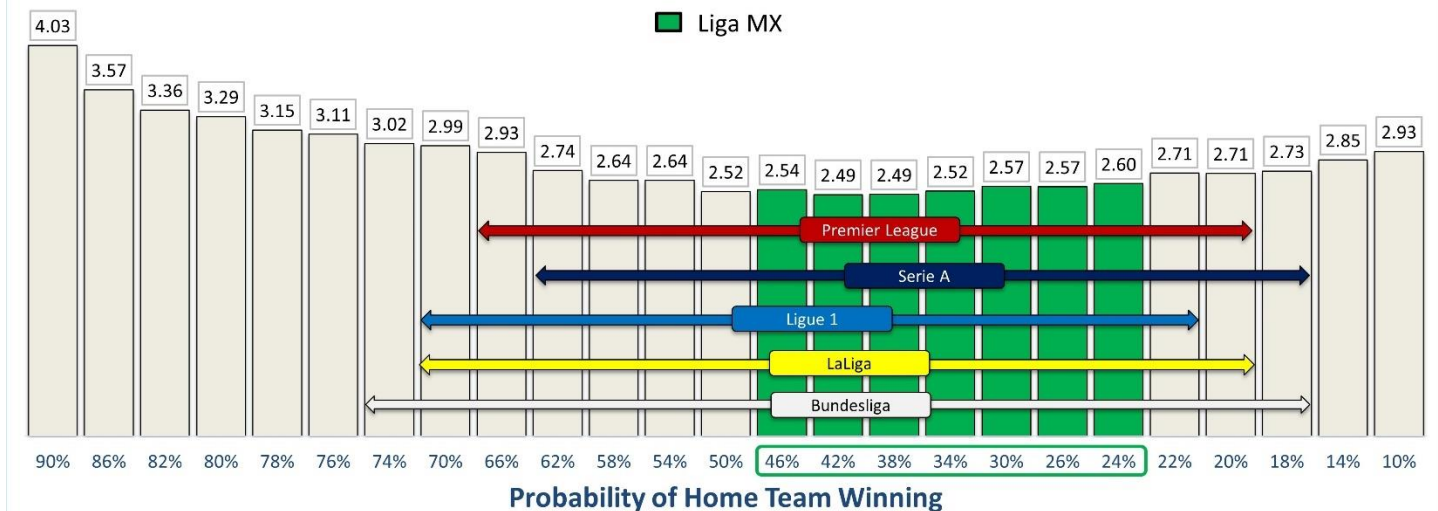
La gráfica anterior tiene unas implicaciones de lo más interesantes. En primer lugar, vemos que el promedio de goles por escenario (la altura de la barra) se comprime hacia el intervalo 42% a 38%. Para explicar el por qué deberás esperar a una entrega posterior, pero, por ahora, observemos a nuestra querida Liga MX (yo sí la quiero, es en serio es bien chida y loca). ¿Dónde se ubican los equipos de la Liga MX en este abanico de escenarios? Reuniendo los resultados y momios desde Apertura 2012 hasta Apertura 2021, los equipos de la primera división de México se ubican en la siguiente franja:

Expected Goals per Match per Probability of Home Team Winning



¿Es esto poco o mucho? ¿bueno o malo? Para que cada uno construya su juicio al respecto, veamos la siguiente gráfica para comparar a la Liga MX con las 5 principales ligas del mundo:

Expected Goals per Match per Probability of Home Team Winning



Antes de caer en el tentador error de decir “pues recorramos la Liga MX un poquitín a la izquierda” para que haya más goles, consideremos que la vida no funciona así. Cada equipo que se aleja de la media lo hará acaparando puntos que, en realidad, le está quitando a otros.

[CONEXIÓN]

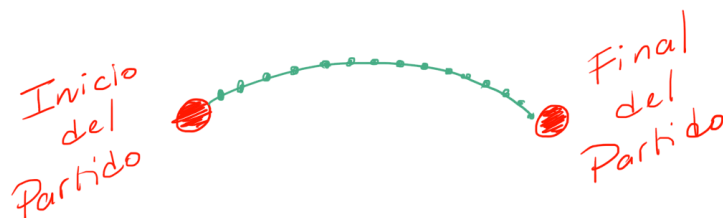
Con lo anterior en mente pensemos un poco las implicaciones en el largo plazo con un ejemplo similar al del partido entre Liverpool y Chelsea. El siguiente es un partido de la Liga MX ocurrido el 19 de agosto de 2021.

Cruz Azul	Empate	Monterrey	Goles de Local	Goles de Visitante
42%	26%	32%	1	1

Sabemos que, a partir de la muestra tomada, los partidos en los que los locales tienen 42% de probabilidad de victoria el promedio de goles por partido es de 2.49, pero este promedio tiene que componerse de la suma de los promedios de los goles anotados por los equipos locales y el promedio de los anotados por los visitantes:

$$\begin{array}{lcl}
 \text{Promedio de} & & \text{Promedio de} \\
 \text{goles en el} & = & \text{goles de} \\
 \text{escenario} & & \text{Cruz Azul} \\
 42\% & & 1.40 \\
 & + & \\
 & & \text{Promedio de} \\
 & & \text{goles de} \\
 & & \text{Monterrey} \\
 & & 1.09 \\
 & = & 2.49
 \end{array}$$

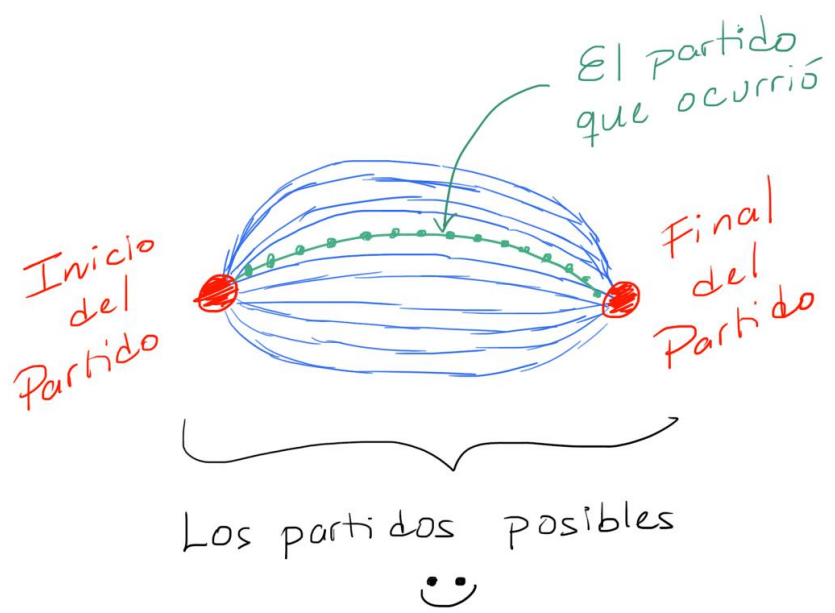
Ahora recordemos el diagrama anterior



Con el diagrama en mente preguntémosnos, ¿cuántos goles tenderían a anotar ambos equipos si jugaran un total de 100 partidos en las mismas circunstancias?

$$\begin{array}{ccccccccc}
 \text{Promedio} & & \text{Promedio de} & & \text{Número de} & & \text{Promedio de} & & \text{Número de} \\
 \text{de goles} & & \text{goles del} & & \text{partidos} & & \text{goles del} & & \text{partidos} \\
 \text{en el} & = & \text{Local} & \times & & + & \text{Visitante} & \times & \\
 \text{escenario} & & 1.40 & & 100 & & 1.09 & & 100 \\
 42\% & & & & & & & & \\
 & & & & & & & & = 249
 \end{array}$$

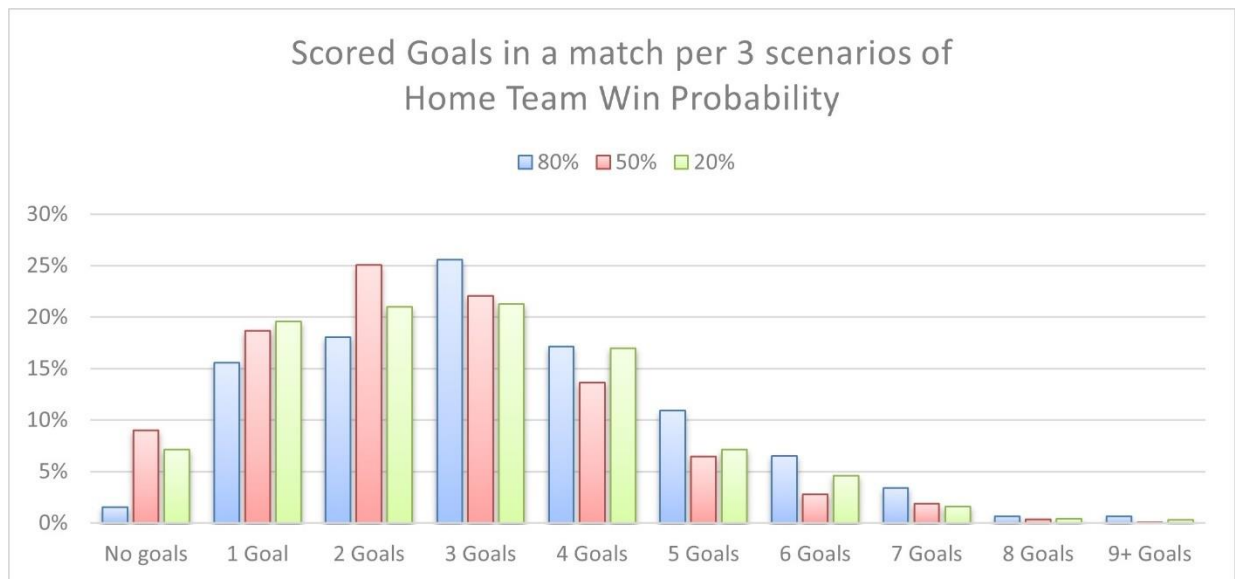
Esta sencilla operación puede hacerse para cualquier partido y así obtener una referencia de los valores a los que deben tender los goles anotados por los equipos, solamente acumulando los valores esperados en cada escenario y, al final, hacer una comparación. Por supuesto, esta referencia se vuelve cada vez más sólida entre más partidos incorporemos a nuestro ejercicio. Resumiendo en una imagen lo que acabamos de hacer fue completar el diagrama anterior, pasando a lo siguiente:



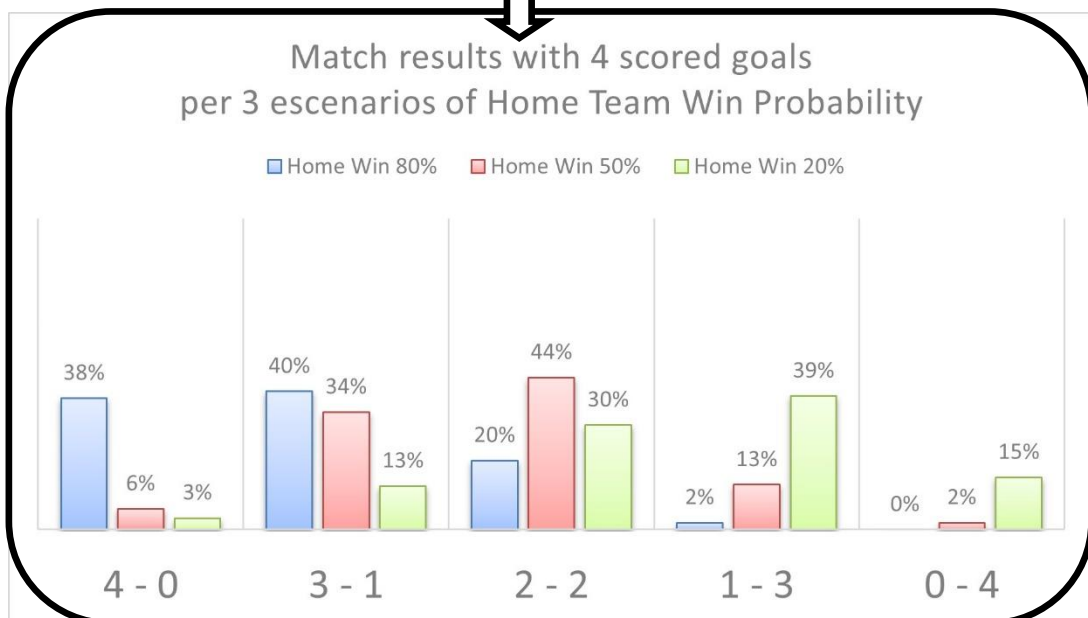
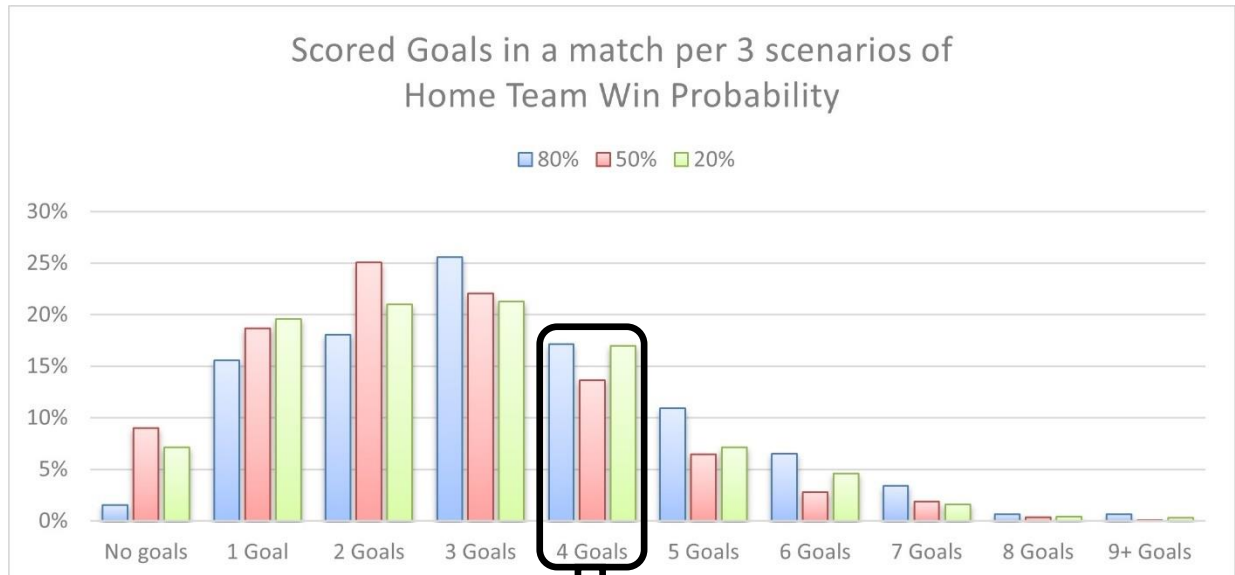
Pero ahora demos un pequeño salto hacia arriba: imagina que en el diagrama anterior cambiamos la palabra “partido” por “torneo”. ¿Cómo se plasman estas nuevas circunstancias en el largo plazo? ¿Cuánto afecta la disparidad de los partidos al número de goles? Para intentar responder estas preguntas usaremos tres escenarios:

Probabilidad de victoria del Local	Probabilidad de victoria del Local	Probabilidad de victoria del Local
90%	46%	10%

Ahora observemos la siguiente gráfica que muestra cómo se distribuye la cantidad de goles según cada escenario:



Ahora hagamos un zoom in a los partidos que terminaron con 4 goles y observemos cómo se distribuyen los goles en los posibles resultados de 4 goles:



El escenario en color rojo es el más cercano a lo que acontece en la LigaMX, lo cuál debe darnos una idea de cómo se comportan, en cuanto a goles anotados, los equipos de la liga. En la primera división del futbol mexicano, entre los torneos Apertura 2012 y Apertura 2021, se anotaron 7425 goles de 7395 esperados en fase regular, es decir, se anotaron 100.4% de goles. Sin embargo, si bien en el largo plazo la LigaMX es un torneo muy estable en cuanto a goles anotados, el torneo Apertura 2021 pareció empeñarse en lo contrario. Observemos la siguiente gráfica:

Liga MX Apertura 2021 Regular Season Goals vs Expected Goals

