

GANAR O PERDER EN EL FÚTBOL DE ALTO NIVEL. ¿UNA CUESTIÓN DE SUERTE?

Lago C.
Facultad CC. de la Educación y el Deporte. Universidade de Vigo

Resumen

En este artículo se estudian los efectos del rendimiento y el azar en la posición que ocupan los equipos en la clasificación de la Liga Española de Fútbol en la temporada 2003-2004. Para ello, se utiliza una metodología basada en la regresión lineal y la comparación de medias. Los resultados empíricos sugieren que el azar puede tener relevancia para explicar el resultado de los equipos en un único partido, pero que a partir de cierta cantidad de encuentros el rendimiento es el factor determinante para dar cuenta de los puntos que alcanzan los conjuntos.

PALABRAS CLAVE: azar, fútbol, rendimiento, resultado

Abstract

In this paper, the effects of performance and random in the position of each team in the Spanish Soccer League in the season 2003-2004 are analysed. The research is based on linear regression analysis and mean comparison tests. My findings lead me to conclude that hazard is a relevant variable for explaining the result of the teams in one only match. However, the higher the number of matches taken into account, the more important performance and the less random.

KEY WORDS: random, performance, result, soccer

1.- INTRODUCCIÓN

El resultado de un partido de fútbol no sólo está sujeto al éxito o al fracaso en la ejecución de los planes de actuación previstos por los entrenadores y jugadores, sino también por otros aspectos que poco o nada tienen que ver estrictamente con la participación de los equipos. Un empate, una victoria o una derrota por 1-0 puede esconder en realidad una clara superioridad en el juego de un conjunto frente a otro.

Es reconocido por entrenadores, jugadores y aficionados que a veces el azar es importante para entender el resultado de un partido. Veamos un ejemplo. En la Tabla 1 se presenta la estadística del partido Atlético de Madrid-Real Madrid de la Liga Española en la temporada 2004-2005. Después de examinar algunos indicadores del rendimiento de los dos equipos, pocos se atreverían a apostar que el resultado final fue de 0-3 a favor del Real Madrid. El rendimiento del Atlético de Madrid fue muy superior al de su rival y sin embargo su resultado fue peor.

Tabla 1: Indicadores de rendimiento en el partido Atlético de Madrid-Real Madrid en la Liga Española de Fútbol de la temporada 2004-2005

ATLÉTICO DE MADRID	0 - 3	REAL MADRID
124	Jugadas de ataque	100
5	Remates portería	3
10	Remates fuera	1
0	Paradas	5
44	Centros al área	3
0	Fueras de juego	3
62	Balones recuperados	52
65	Balones perdidos	75
17	Asistencias	3
21	Faltas recibidas	18
18	Faltas cometidas	21

FUENTE: GECASPORT

¿Puede resultar el azar un aspecto determinante para explicar el resultado de los equipos al final de una temporada? ¿Es realmente tan importante la suerte en un único partido? ¿Cómo se puede distinguir entonces en el resultado de un partido entre los efectos propiamente dichos del rendimiento de los equipos y los de otros acontecimientos ligados al azar pero que resultan ajenos a la propia actuación de los participantes?

La respuesta más habitual a esta última pregunta en la literatura es que no es posible o no merece la pena, de modo que se asume que todos los acontecimientos previstos o no en el juego son un efecto del rendimiento de los equipos. Si aceptamos que la suerte posee un papel relevante en el resultado de los equipos al final de una temporada, estaremos minimizando la importancia del rendimiento y, por tanto, del entrenamiento en sí mismo. Si, por el contrario, sostenemos que el azar tiene una importancia limitada para dar cuenta de la posición final de los equipos en el campeonato, estaremos reconociendo que rendimiento y resultado son la misma cosa. Claro que eso puede que no explique el resultado del partido entre el Atlético de Madrid y el Real Madrid.

La pretensión de este artículo es estimar con precisión el efecto causal del rendimiento en los puntos que alcanzan los equipos en diferentes momentos del campeonato, distinguiéndolo así del simple azar. Además, intentaremos justificar que la diferencia en los resultados obtenidos por los equipos situados en los primeros puestos de la clasificación en comparación con los demás está influenciado más por su buen rendimiento que por la suerte.

La investigación se estructura del siguiente modo. Una vez presentados los objetivos, se caracterizan el rendimiento y sus efectos y se explica la metodología planteada. A continuación, se prueba esta metodología en el análisis de los primeras 17 jornadas de la Liga Española de Fútbol de la temporada 2003-2004 y se discuten

los resultados empíricos alcanzados. En último lugar, se destacan los principales conclusiones del artículo.

2.- METODOLOGÍA

El rendimiento deportivo significa una comparación entre el valor previsto y el real, y entre las mejoras verificadas en entrenamiento y competiciones complementarias y el porcentaje de éstas expresado en las competiciones principales (Lago y Martín Acero, 2005a; Martín Acero y Vittori, 1997). Dichos rendimientos son siempre relativos. Por tanto, al analizar el nivel alcanzado tendremos presente que el rendimiento:

- es dependiente de la capacidad de cada jugador, de los compañeros del equipo y de los contrincantes;
- está sometido a la valoración propia y de otros miembros del grupo;
- está influido por las condiciones externas.

Podríamos pensar que, al registrar el nivel de éxito o fracaso que cada jugador y cada equipo expresan en cada partido, estamos observando de forma directa su máximo nivel de prestación. Sin embargo, los deportes de equipo siempre conllevan un cierto margen para lo aleatorio y lo impredecible. Supongamos que por un repentino cambio en el calendario de la competición, la Federación de Fútbol hubiera decidido repetir el partido Atlético de Madrid-Real Madrid cada semana y supongamos también que éstos fueran independientes entre sí. Aunque los planes de actuación previstos por los entrenadores y jugadores fueran siempre los mismos, cada repetición semanal no registraría la misma evolución de la competición y el mismo resultado. Podría cambiar la situación meteorológica, cometer errores el árbitro, algún jugador podría lesionarse, y todo ello afectaría al desarrollo del juego y a su resultado. De este modo, numerosos acontecimientos transitorios podrían producir resultados ligeramente diferentes. Después de todo, nuestra observación de cualquiera de estos partidos no sería una medida perfecta de la fortaleza de los equipos.

En este ejemplo, puede decirse que la varianza del resultado en diferentes partidos surge de dos factores separados: las diferencias *sistemáticas* y las *no sistemáticas*. En el ejemplo anterior, las primeras incluyen características fundamentales y predecibles de los equipos como su planteamiento táctico, la apuesta por el dominio de la posesión del balón o no, la calidad individual de los jugadores o el tipo de defensa seleccionada. En hipotéticas repeticiones semanales del mismo partido se mantendrían las diferencias sistemáticas pero variarían las no sistemáticas, como son los cambios producidos por alteraciones meteorológicas o los errores arbitrales.

Los factores sistemáticos son persistentes y sus consecuencias se repiten cuando tienen un valor determinado. Cuando equipo apuesta por tener la iniciativa en el juego, es posible que alcanzar un porcentaje alto de posesión del balón le permita mejorar su rendimiento. Los no sistemáticos son transitorios: no podemos predecir su impacto. No podemos saber si la suerte acabará por empujar o no un balón que golpea en el poste hacia el interior de la portería o si el árbitro se equivocará en una decisión clave.

Sin embargo, esto no significa que los sistemáticos representen constantes. La posesión del balón puede ser un componente sistemático a la hora de explicar el comportamiento en el juego, pero esto no significa que los valores de la posesión no varíen en cada encuentro. En el resultado de un partido, la constante de la posesión del balón es su efecto y , si es variable, cambia de forma predecible (Lago y Martín Acero, 2005b). Del mismo modo, la climatología puede ser un factor aleatorio (si los cambios intermitentes o impredecibles tienen consecuencias también impredecibles) o una característica sistemática (si el mal tiempo produce un resultado siempre favorable para un equipo). Así, por ejemplo, el hecho de jugar en casa, que en principio podría parecer un factor no sistemático del resultado, ha sido estudiado en la literatura y se ha verificado que existe un efecto real sobre el resultado que favorece a los equipos locales. De este modo, jugar en casa o fuera supone un factor sistemático que afecta al resultado de los partidos (Pollard, 1986; Thomas, Reeves & Daves, 2004).

Por lo tanto, uno de los objetivos fundamentales de la inferencia es distinguir entre el componente sistemático y el no sistemático dentro de los fenómenos que estudiamos (King, Keohane y Verba, 2000).

Así, el resultado de un partido de fútbol puede modelizarse como una función de factores sistemáticos y no sistemáticos:

$$y: f(x, \mu_t),$$

donde y es el resultado del partido, x es el rendimiento alcanzado por los equipos (factor sistemático) y μ_t es el azar propio de la competición (factor no sistemático). En cualquier caso, se asume además que las perturbaciones propias de u_t están distribuidas normalmente, de modo que: i) este factor aleatorio no sistemático desaparece en la media de las reproducciones hipotéticas $E(u_t) = 0$; ii) no se correlaciona con la verdadera variable dependiente, $C(u_t, Y^*) = 0$, ni con las independientes $C(u_t, X^*) = 0$, y iii) tienen la misma varianza $V(u_t) = \sigma^2$ en cada una de las unidades. En fin, los estimadores de y satisfacen diversas propiedades estadísticas deseables, como son la ausencia de sesgos y la varianza mínima.

Dicho de otro modo, es posible que para un único partido o un número muy limitado el azar, provoque una alteración en los acontecimientos que ocasione un

resultado imprevisto por el rendimiento alcanzado por los equipos en la competición: “*cada partido es una historia*” Sin embargo, la suerte sólo afecta al partido en que ocurre. Después de un número amplio de observaciones, la normalidad de las perturbaciones estocásticas justificaría entonces que en el análisis empírico se excluya μ_t y se exprese la influencia del rendimiento exclusivamente como el resultados de un componente sistemático o determinístico (si no fuera así, el azar sería un factor sistemático)

Uno de los aspectos más importantes a la hora de realizar cualquier investigación empírica o evaluar el comportamiento de los jugadores y equipos en un partido, consiste en seleccionar la/s variable/s que permitan cuantificar el rendimiento de los equipos en la competición (Garganta, 2000; Hughes y Bartlett, 2002).

Es evidente que el indicador de éxito en el juego más significativo lo constituye, sin duda, los goles anotados o recibidos. Sin embargo, a veces un empate, una victoria o una derrota puede esconder una clara superioridad en el juego de un conjunto frente a otro que no se refleja en el resultado final. Además, la limitada varianza en el número de goles que se produce en un partido limita la inferencia estadística de la investigación. Por ello, la *proxy* de rendimiento utilizada en este trabajo consiste en la diferencia entre los lanzamientos realizados y recibidos por cada equipo en cada partido. Es decir, si en un encuentro el equipo A efectúa 15 lanzamientos a la portería rival y recibe 10, su rendimiento será de 5; mientras que el rendimiento de su oponente será de -5. Lógicamente, el rendimiento positivo de un equipo se corresponde con un rendimiento negativo de la misma magnitud por parte de su rival: la suma de ambos rendimientos debe ser siempre por ello igual a 0. Hemos recogido en la variable lanzamientos a portería, tanto los remates que se producen entre los tres palos como aquellos que van fuera de la portería con el fin con el fin de maximizar el número de observaciones y facilitar la inferencia estadística. El elevado número de lanzamientos que se suceden, en comparación con la baja cantidad de goles que se producen, y la evidente relación que aquéllos guardan con el marcador final (a más lanzamientos realizados y menos recibidos parece lógico que se incremente la posibilidad de ganar) justifican el indicador del rendimiento seleccionado para este trabajo en Hughes y Bartlett (2002).

La investigación empírica que se presenta a continuación tiene tres grandes objetivos: i) estimar cuál es el efecto causal del rendimiento sobre los puntos que obtienen los equipos en la competición; ii) comprobar si el efecto del rendimiento es igual en diferentes momentos de la temporada: después de un número limitado de partidos o de una serie amplia de observaciones; iii) constatar si existen diferencias

en el rendimiento de los equipos que se sitúan en los primeros puestos de la clasificación con respecto a los demás.

Para llevar a cabo el análisis empírico se han tomado un total de 340 partidos, correspondientes a las primeras 17 jornadas de la Liga Española de la Temporada 2003-2004. De cada partido se han recogido datos de cada uno de los dos equipos que se enfrentan entre sí.

Para dar cuenta de los dos primeros objetivos de la investigación se han realizado diferentes análisis de regresión lineal. En cada regresión la variable dependiente es el número de puntos que alcanzan los equipos. El coeficiente positivo o negativo para el rendimiento (variable independiente o explicativa) implica que ésta tiene respectivamente influencia positiva o negativa sobre los valores de la variable dependiente. Además, para aceptar la relevancia estadística de las variables es preciso rechazar la hipótesis nula de que su coeficiente es igual a 0.

Desde los pioneros estudios cuantitativos de McGarry y Franks (1994), el análisis de regresión se ha convertido en los últimos años en una de las técnicas estadísticas más valoradas en el análisis del rendimiento en los deportes de equipo (Álvarez, 2005; Atkinson y Neville, 2001; Carmichael, Thomas y Ward, 2001; Dawson, Dobson y Gerrard, 2000; Ensum, Pollard y Taylor, 2004; Hadley, Poitras, Ruggiero y Knowles, 2000; Hughes, 2003; Hughes, Evans y Wells, 2001; James, Jones y Mellalieu, 2004; Lago y Martín Acero, 2005b; McGarry y Franks, 2003; Neville, Atkinson, Hughes y Cooper, 2002; Pollard y Reep, 1997).

Para constatar si existen diferencias en el rendimiento de los equipos que se sitúan en los primeros puestos de la clasificación con respecto a los demás se han dividido a los 20 equipos que participan en la Liga Española de Fútbol en cuatro grupos en función del puesto que ocupan en la jornada 17. El Grupo de la Liga de Campeones (G_{LC}) está conformado por los equipos que ocupan los 4 primeros puestos de la clasificación (Real Madrid, Valencia, Deportivo y Osasuna). El Grupo de la UEFA (G_{UEFA}) está integrado por los conjuntos que se sitúan entre los puestos 5 y 9 (Villarreal, Atlético, Málaga, Athletic y Barcelona). Los equipos que van entre los puestos 10 y 17 (Valladolid, Mallorca, Racing de Santander, Sevilla, Albacete, Zaragoza, Betis y Celta) forman el Grupo Intermedio ($G_{INTERMEDIO}$). Finalmente, los conjuntos que ocupan los puestos de descenso (Real Sociedad, Murcia y Espanyol) integran el Grupo de Descenso ($G_{DESCENSO}$).

Para analizar si dos grupos difieren en su rendimiento se utilizará una comparación de medias. El fundamento de esta técnica es muy sencillo: se calculan las medias de la variable dependiente (el rendimiento) para los grupos que forman las variables independientes (G_{LC} , G_{UEFA} , $G_{INTERMEDIO}$ y $G_{DESCENSO}$) y se estudia las diferencias que se observan. Para decidir si el estadístico obtenido es significativo se

utilizan los contrastes, pruebas o test de hipótesis. En los contrastes de diferencias vamos a tener hipótesis nula y además una hipótesis alternativa. La hipótesis nula siempre será que la diferencia de medias de la población (D) es igual a 0 ($H_0: D=0$). Como hipótesis alternativa para nuestro trabajo manejaremos que:

- Hipótesis alternativa A : las medias son distintas:

$$H_1: D \neq 0$$

Como hipótesis de la investigación se plantea que el rendimiento en el juego es un factor que determina los puntos que obtienen los equipos en la competición y que su importancia se hace mayor a medida que se suceden las jornadas. Además se presume que existen diferencias significativas en el rendimiento de los equipos que se sitúan en los primeros puestos de la clasificación con respecto a los demás, y que éstas son mayores a medida que la distancia en la clasificación entre los equipos es más grande.

3.- RESULTADOS Y DISCUSIÓN

En la Tabla 2 se compara la influencia del rendimiento de los equipos sobre los puntos que alcanzan en las jornadas 3, 5, 10 y 17. Es preciso recordar que la victoria otorga 3 puntos, el empate 1 y la derrota 0. De este modo, el número máximo de puntos que un equipo podría alcanzar en las jornadas 3, 5, 10 y 17 es respectivamente 9, 15, 30 y 51. En las tres primeras jornadas, el rendimiento de los equipos no es estadísticamente significativo para explicar los puntos obtenidos, aunque el signo positivo esperado (cuanto mejor es el rendimiento, se obtienen más puntos). El modelo apenas explica el 12 por 100 de la varianza de los puntos obtenidos por los equipos. La constante del modelo es significativa al 1 por 100 y se refiere a los puntos que cabría esperar que obtendría un equipo con un rendimiento de 0: en este caso 4. El coeficiente del rendimiento alcanza un valor de 0,15, es decir, que cada punto de mejora en el rendimiento otorga 0,15 puntos más en la clasificación.

Así, por ejemplo, un rendimiento positivo de 6, supondría que un equipo sume prácticamente 1 punto más en las tres primeras jornadas con respecto a un equipo con un rendimiento de 0. Del mismo modo, con un rendimiento de -6 , cabría esperar que ese equipo tendría 1 punto menos. En la jornada 5, el rendimiento pasa a ser significativo al 5 por ciento y explica el 27 por 100 de los puntos obtenidos por los equipos. Además, el coeficiente del rendimiento pasa a ser de 0,39. La constante es significativa al 1 por 100, con un valor de 6,76. A partir de la jornada 10, el rendimiento es significativo al 1 por ciento y explica cada vez más los puntos que se alcanzan en el campeonato. En la jornada 10 el coeficiente del rendimiento es de 1,01 y en la 17 llega a 1,70. Así, si consideramos que la diferencia de rendimiento,

por ejemplo, del Valencia y el Espanyol es de 10 puntos, cabría esperar que los primeros tuvieran 17 puntos de ventaja con respecto a los segundos en la jornada 17. La constante sigue siendo significativa al 1 por 100.

De este modo, parece posible concluir dos cuestiones básicas acerca de la relación entre resultado, rendimiento y azar que verifican las hipótesis planteadas anteriormente: i) el rendimiento de los equipos es importante para dar cuenta de los puntos que alcanzan los equipos en el campeonato; ii) la influencia del rendimiento se incrementa conforme pasan las jornadas, de modo que el azar puede tener relevancia para explicar el resultado de los equipos en un único partido o en un número limitado de ellos, pero no después de un número amplio de jornadas.

Tabla 2. Análisis de la influencia del rendimiento en los puntos obtenidos por los equipos en las jornadas 3, 5, 10 y 17 de la Liga Española de Fútbol en la temporada 2003/2004

Variable Dependiente: Puntos	Modelos			
	Jornada 3	Jornada 5	Jornada 10	Jornada 17
Rendimiento	0,15 (0,10)	0,39** (0,15)	1,01* (0,22)	1,70* (0,51)
Constante	4,04* (0,49)	6,76 (0,63)	13,68* (0,77)	23,07 (1,39)
R ²	0,12	0,27	0,54	0,38
Número de observaciones	20	20	20	20

Notas: Aparecen en primer lugar los coeficientes de regresión estimados, seguidos por las desviaciones típicas de los parámetros calculadas a partir de la matriz de varianzas y covarianzas estimada mediante el método de mínimos cuadrados ordinarios (MCO) y los errores robustos. R₂ es el coeficiente de determinación.

*p<0,01. **p<0,05.

En la Tabla 3 se presenta una simulación de los puntos que cabría esperar que alcanzase un equipo en las jornadas 3, 5 10 y 17 de la Liga Española de Fútbol, de acuerdo con los resultados presentados en la Tabla 2. Como puede apreciarse, a medida que se suceden las jornadas del campeonato las diferencias en los puntos obtenidos por los equipos con un rendimiento positivo o negativo se incrementan notablemente. Las limitadas diferencias de puntuación en la jornada 3, a pesar de las notables disparidades de rendimiento entre los equipos (un equipo con un rendimiento de +15 únicamente tendría el doble de puntos que un conjunto con un rendimiento de -10), se amplían a medida que avanza el campeonato (en la jornada 17 el equipo con mejor rendimiento tendría 8 veces los puntos del equipo con peor rendimiento).

Tabla 3. Estimación de los puntos previstos en las jornadas 3, 5, 10 y 17 de la Liga Española de Fútbol en según el rendimiento alcanzado

Rendimiento	Puntos			
	Jornada 3	Jornada 5	Jornada 10	Jornada 17
-10	3	3	4	6
-5	3	5	9	15
0	4	7	14	23
5	5	9	19	31
10	6	11	24	40
15	6	13	29	49

Nota: las puntuaciones de los equipos se han redondeado con el fin de ofrecer valores enteros reales de la competición.

En la Tabla 4 se presenta la comparación de medias en el rendimiento entre los cuatro grupos que recogen a los equipos de la Liga Española de Fútbol.

En primer lugar, se constata que no existen diferencias estadísticamente significativas entre el rendimiento del G_{LC} y el G_{UEFA} . Si bien el signo tiene el valor esperado (los equipos del G_{LC} tienen un mejor rendimiento que los del G_{UEFA}), no es posible descartar la hipótesis nula de igualdad de medias. En la comparación entre el G_{LC} y el $G_{INTERMEDIO}$ la diferencia de rendimiento es de 3,65 puntos y además es estadísticamente significativa al 1 por ciento. Finalmente, la diferencia en el rendimiento entre G_{LC} y el G_{UEFA} también es significativa estadísticamente al 1 por 100 y alcanza los 5,48 puntos. Es decir, se confirma que el rendimiento es un buen predictor de la posición que ocupan los equipos en la clasificación. La imposibilidad de rechazar la hipótesis nula en la comparación entre el G_{LC} y el G_{UEFA} nos lleva a pensar que quizás no exista una gran diferencia en el rendimiento obtenido por ambos grupos (no hay una gran diferencia entre ser 4 ó 6 en la clasificación) o que es posible que algún equipo tenga un rendimiento superior o inferior al que cabría esperar por su posición.

Tabla 4.

Tests de comparación de medias

	Todas las observaciones	$G_{LC}-G_{UEFA}$	$G_{LC}-G_{INTERMEDIO}$	$G_{LC}-G_{DESCENSO}$
Observaciones	340	153 (68+85)	204 (68+136)	119 (68+51)
N_h (media)	0,06 (7,01)			
N_{GLC} (media)		2,56 (0,88)	2,56 (0,88)	2,56 (0,88)
N_{GUEFA} (media)		1,58 (0,76)		
N_{GRESTO} (media)			-1,09 (0,58)	
$N_{GDESCENSO}$ (media)				-2,92 (0,82)
Diferencia		0,98 (1,16)	3,65 (1,03)	5,48 (1,24)
T		0,84	3,56	4,41
p>t		0,40	0,00*	0,00*

Notas: Desviaciones típicas entre paréntesis

*p<0,01

En la Figura 1, se representa la relación existente entre rendimiento y resultado en la Jornada 3, 5, 10 y 17. Si las cuatro rectas fuesen paralelas, el efecto del rendimiento en todos los casos sería el mismo. La pendiente más acusada de las rectas conforme se suceden las jornadas del campeonato indica que la influencia del rendimiento es cada vez mayor sobre los puntos obtenidos.

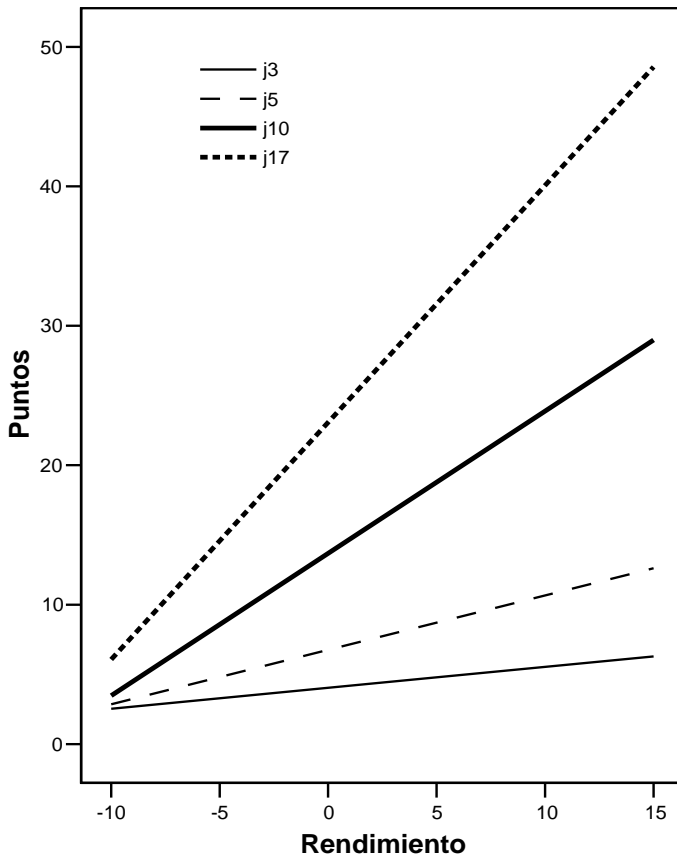


Figura 1. Relación entre el rendimiento alcanzado por los equipos en las Jornadas 3, 5, 10 y 17 y los puntos obtenidos en Liga Española de Fútbol en la temporada 2003/2004

En la Figura 2 se representa la relación existente entre el rendimiento alcanzado por los 20 equipos de la Liga Española de Fútbol y la posición que ocupan en el Campeonato después de disputar las primeras 17 jornadas. Como norma general, se confirma que los equipos con un rendimiento bajo se sitúan en los puestos inferiores de la clasificación. Llama la atención el rendimiento del Deportivo, el Barcelona y el Athletic. El Deportivo tiene una media de rendimiento inferior a varios equipos que se sitúan por detrás en la clasificación. En el caso del Athletic y el Barcelona, su rendimiento es el segundo y el tercero del Campeonato y, a pesar de ello ocupan los puestos 8 y 9. En cualquier caso, hay que recordar que en esa temporada el Barcelona acabó siendo segundo y el Athletic sexto. Es posible que estos casos sean responsables de no haber encontrado diferencias estadísticas entre el G_{LC} y el G_{UEFA} .

Además, destaca la diferencia de rendimiento entre el Real Madrid y el Valencia, primer y segundo clasificado. Cabría esperar que el equipo con el mejor rendimiento de la competición ocupase el primer puesto. En cualquier caso, resulta sugerente apuntar que el Valencia acabó siendo el campeón de Liga esa temporada.

Por último, en la Figura 3 se recogen comparativamente las diferencias de rendimiento entre los equipos que participan en la Liga Española tomando como referencia los valores del Valencia. Todos los equipos presentan diferencias estadísticamente significativas con respecto al Valencia, salvo el Real Madrid, el Osasuna, el Athletic, el Barcelona y el Betis. Las diferencias de rendimiento se hacen cada vez mayores conforme los equipos se encuentran en los puestos finales de la clasificación.

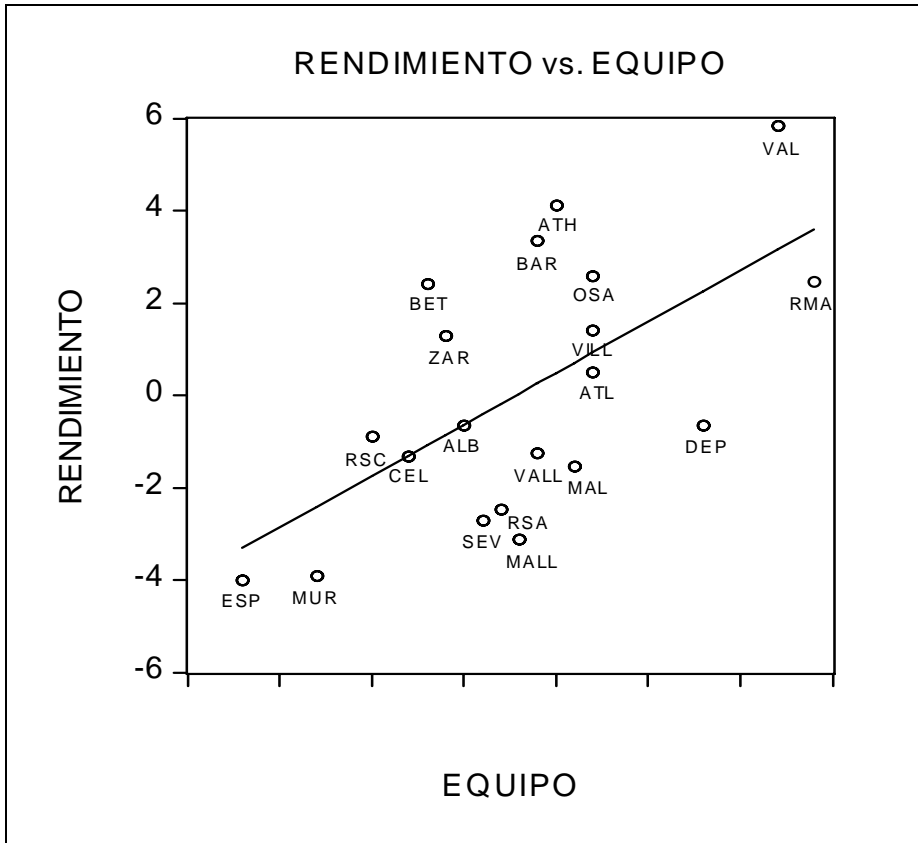


Figura 2.
Correlación entre el rendimiento y la posición ocupada por los equipos
de la Liga Española de Fútbol

Notas: RMA = Real Madrid; VAL = Valencia; DEP = Deportivo; OSA = Osasuna; VILL = Villarreal; ATL = Atlético de Madrid; MAL = Málaga; ATH = Athletic de Bilbao; BAR = Barcelona; VALL = Valladolid; MALL = Mallorca; RSA = Racing de Santander; SEV = Sevilla; ALB = Albacete; ZAR = Zaragoza; BET = Betis; CEL = Celta; RSC = Real Sociedad; MUR = Murcia; ESP = Espanyol.

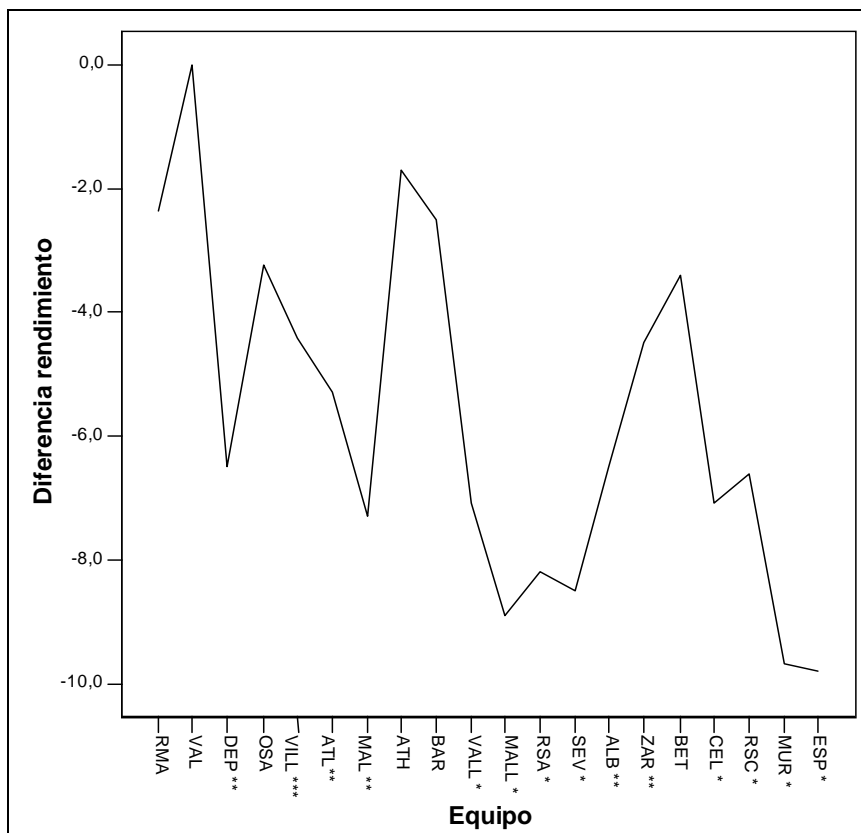


Figura 3. Diferencia de rendimiento entre los equipos que participan en la Liga Española de Fútbol

Notas: RMA = Real Madrid; VAL = Valencia; DEP = Deportivo; OSA = Osasuna; VILL = Villarreal; ATL = Atlético de Madrid; MAL = Málaga; ATH = Athletic de Bilbao; BAR = Barcelona; VALL = Valladolid; MALL = Mallorca; RSA = Racing de Santander; SEV = Sevilla; ALB = Albacete; ZAR = Zaragoza; BET = Betis; CEL = Celta; RSC = Real Sociedad; MUR = Murcia; ESP = Espanyol.

* $p < 0,01$. ** $p < 0,05$. *** $p < 0,10$.

4.- CONCLUSIONES

En este artículo hemos calculado cuáles fueron los efectos del rendimiento sobre los puntos obtenidos por los equipos en las primeras 17 jornadas de la Liga Española de Fútbol de la temporada 2003-2004. Después de presentar diferentes análisis de regresión lineal, hemos examinado la influencia del rendimiento y el azar en distintos momentos de la competición. Además, a través de una comparación de

medias, hemos justificado que la diferencia de puntos obtenidos por los equipos situados en los primeros puestos de la clasificación en comparación con los demás, está influenciada más por el rendimiento que por la suerte. Los resultados empíricos demuestran que el azar puede tener relevancia para explicar el resultado de los equipos en un único partido o, a lo sumo, en un número muy limitado de ellos, pero que a partir de cierta cantidad de encuentros el rendimiento es determinante para dar cuenta de los puntos que alcanzan los conjuntos.

En las simulaciones realizadas a partir de los modelos anteriores se ha constatado, por ejemplo, como en la jornada 17 del Campeonato cabría esperar que un equipo con un rendimiento 10 puntos superior a otro, tuviese 17 puntos más en la clasificación. Los entrenadores y los jugadores deberían prestar por ello más atención a valorar el comportamiento del equipo por su rendimiento en aquellas facetas del partido que le permiten desarrollar su modelo de juego y no tanto por el resultado final de un único partido, puesto que, como se ha demostrado rendimiento y resultado, a la larga, van de la mano.

De todos modos, los resultados de estas simulaciones deben tomarse con cierta cautela, puesto que la capacidad predictiva del modelo no es perfecta. Existe un margen relativamente amplio para que otras variables no contempladas en este trabajo o, simplemente, el azar repercuta en el resultado de los equipos. La predicción de las puntuaciones de los equipos pretende mostrar las tendencias o posibilidades de evolución en el campeonato que cabría esperar si se dan ciertas condiciones de rendimiento.

Otra conclusión a destacar es que en este trabajo hemos seleccionado una variable del rendimiento en el juego estrechamente relacionada con el resultado y no tanto con el modelo de actuación individual de los equipos. Parece razonable que, una vez que se ha justificado que para todos los equipos el rendimiento es determinante para dar cuenta de los puntos alcanzados, se traslade la atención hacia el comportamiento individual de los equipos en competición y cómo el modelo de juego que propone cada conjunto y su ejecución tiene consecuencias para su rendimiento. Para ello resultará clave seleccionar variables específicas del juego para cada equipo. Así, por ejemplo, para un conjunto que actúa a la expectativa, puede que sea poco importante la posesión del balón: su rendimiento debe ser juzgado, por ejemplo, por los remates que recibe en su portería, por las situaciones de 1x1 que gana en su propio campo o por la velocidad en la transición defensa/ataque cuando recupera la pelota y realiza un contraataque. Para un equipo que pretende la iniciativa, es posible que la posesión del balón y el número de lanzamientos que realiza sean claves. En cualquier caso, la selección de estos

indicadores del rendimiento deberá ser individual para cada equipo dependiendo de su modelo de juego.

REFERENCIAS

- Álvaro, J. (2005). El análisis de la competición como instrumento para la toma de decisión de los entrenadores: un estudio de la Liga Española de Fútbol Profesional de Primera División de la Temporada 2003-2004, Tesis Doctoral inédita, Madrid: Universidad Europea de Madrid.
- Atkinson, G., Neville, A. (2001). Selected issues in the design and analysis of sport performance research, *Journal of Sports Science*, 19, 811-827.
- Carmichael, F., Thomas, D., Ward, R. (2001). Production and Efficiency in Association Football. *Journal of Sports Economics*, 2 (3), 228-243.
- Castelo, J. (1996). *Fútebol. A organizaçao de jogo*, Lisboa: Edição do autor.
- Dawson, P., Dobson, S., Gerrard, B. (2000). Stochastic Frontiers and the Temporal Structure of Managerial Efficiency in English Soccer. *Journal of Sports Economics*, (1) 4, 24-32.
- Ensum, R., Pollard., Taylor, S. (2004). Applications of logistic regression to shots at goal in association football: calculation of shots probabilistics quantification of factors and player/team, *Journal of Sports Science*, (22) 6, 504.
- Garganta, J. (2000). Análisis del juego del fútbol. El recorrido evolutivo de las concepciones, métodos e instrumentos”. *Revista de Entrenamiento Deportivo*, Tomo XIV, 2, 6-13.
- Hadley, L., Poitras, M., Ruggiero, J. Knowles, S (2000). Performance Evaluation of National Football League Teams. *Managerial and Decision Economics*, (21), 4, 45-56.
- Hughes, M. (2003). Notational analysis, in *Science and Soccer* (eds T. Reilly and M. Williams), Routledge: London, 245-264.
- Hughes, M.D, Bartlett, R. (2002). Performance indicators in performance analysis, *Journal of Sports Science*, 20, 738-754
- Hughes, M.D.; Dawkins, N., David, R.; Mills, J. (1998). The perturbation effect and goal opportunities in soccer. *Journal of Sports Science*, 16, 20-28.
- Hughes, M.D., Evans, S.; Wells, J. (2001). Establishing normative profiles in performance analysis. *Electronien International Journal of Performance Analysis of Sports*, 1, 1-25.
- James, N., Jones, P.D., Mellalieu, S.D. (2004). Possession as a performance indicator in soccer as a function of successful and unsuccessful teams. *Journal of Sports Science*, (22) 6, 507-508.

- King, G., Keohane, R. O., Verba, S. (2000). *El diseño de la investigación científica. La inferencia científica en los estudios cualitativos*, Madrid: Alianza.
- Lago, C., Martín Acero, R. (2005a). *Deportes de equipo: comprender la complejidad para elevar el rendimiento*, Barcelona: INDE.
- Lago, C., Martín Acero, R. (2005b). Determinantes en el fútbol de alto rendimiento: el tiempo de posesión del balón (abriendo la caja negra del fútbol), *Revista de Entrenamiento Deportivo*, Tomo XIX, N. 2, 13-19.
- Lees, A. (2002). Technique analysis in sports: a critical review. *Journal of Sports Science* 20, 813-828.
- Lees, A., Nolan, L. (1998). The biomechanics of soccer: a review. *Journal of Sports Science* 16, 211-234.
- McGarry, T., Anderson, D.I., Wallace, S., Hughes, M.D., Franks, I. (2002). Sport competition as a dynamical self-organizing system. *Journal of Sports Science*, 20, 771-781.
- McGarry, T., Franks, I. (1994). A stochastic approach to predicting competition squash match-play, *Journal of Sports Science*. 12, 573-584.
- McGarry, T., Franks, I. (2003). The science of match analysis, in *Science and Soccer* (eds T. Reilly and M. Williams), Routledge: London, 265-275.
- Martín Acero, R; Vittori, C. (1997). Metodología del rendimiento deportivo (I): Sentido, definición y objeto de estudio, *Revista de Entrenamiento Deportivo*, Tomo XI, N.1, 5-10.
- Neville, A.; Atkinson, G.; Hughes, M., Cooper, S.M. (2002). Statistical methods for analysing discrete and categorical data recorded in performance analysis. *Journal of Sports Science*, 20, 829-844.
- Pollard, R. (1986). Home advantage in soccer: a retrospective analysis. *Journal of Sports Science*, 4, 237-246.
- Pollard, R., Reep, C. (1997). Measuring effectiveness of playing strategies at soccer. *The Statistician*, 46, 541-550.
- Thomas, S., Reeves, C., Davies, S. (2004). An analysis of home advantage in the English Football Premiership. *Perceptual and Motor Skill*, 99, 1212-1216.
- White, H. (1980). A Heterokedastic-Consistent Covariance Matrix Estimator y a Direct Test for Heterokedasticity, *Econometrica*, 48, 817-838.