

Información Importante

La Universidad de La Sabana informa que el(los) autor(es) ha(n) autorizado a usuarios internos y externos de la institución a consultar el contenido de este documento a través del Catálogo en línea de la Biblioteca y el Repositorio Institucional en la página Web de la Biblioteca, así como en las redes de información del país y del exterior, con las cuales tenga convenio la Universidad de La Sabana.

Se permite la consulta a los usuarios interesados en el contenido de este documento, para todos los usos que tengan finalidad académica, nunca para usos comerciales, siempre y cuando mediante la correspondiente cita bibliográfica se le dé crédito al trabajo de grado y a su autor.

De conformidad con lo establecido en el artículo 30 de la Ley 23 de 1982 y el artículo 11 de la Decisión Andina 351 de 1993, La Universidad de La Sabana informa que los derechos sobre los documentos son propiedad de los autores y tienen sobre su obra, entre otros, los derechos morales a que hacen referencia los mencionados artículos.

BIBLIOTECA OCTAVIO ARIZMENDI POSADA
UNIVERSIDAD DE LA SABANA
Chía - Cundinamarca



El presente formulario debe ser diligenciado en su totalidad como constancia de entrega del documento para ingreso al Repositorio Digital (Dspace).

TITULO	MODELO DE PREDICCIÓN DE RESULTADOS EN EL FUTBOL PROFESIONAL COLOMBNIANO		
SUBTITULO			
AUTOR(ES) Apellidos, Nombres (Completos) del autor(es) del trabajo	GÓMEZ BAYONA CARLOS ADOLFO		
PALABRAS CLAVE (Mínimo 3 y máximo 6)	FUTBOL		LOGIT ORDENADO
	PREDICCIÓN		GOLES
	RESULTADOS		EQUIPOS
RESUMEN DEL CONTENIDO (Mínimo 80 máximo 120 palabras)	<p>En esta investigación se pretende encontrar los equipos que ingresarán a la fase final del futbol profesional colombiano, implementando herramientas econométricas como logit ordenado, buscando encontrar los cortes de probabilidad y así analizarlos con los resultados obtenidos en las predicciones.</p> <p>Con este documento, se logra afianzar en el pronóstico de resultados en el deporte, logrando estar al mismo nivel de países europeos en la predicción de marcadores deportivos.</p> <p>Se logra llegar a la conclusión de que los modelos propuestos por el investiugador tienen un acierto superior al 80%.</p>		

Autorizo (amos) a la Biblioteca Octavio Arizmendi Posada de la Universidad de La Sabana, para que con fines académicos, los usuarios puedan consultar el contenido de este documento en las plataformas virtuales de la Biblioteca, así como en las redes de información del país y del exterior, con las cuales tenga convenio la Universidad.

De conformidad con lo establecido en el artículo 30 de la Ley 23 de 1982 y el artículo 11 de la Decisión Andina 351 de 1993, "Los derechos morales sobre el trabajo son propiedad de los autores", los cuales son irrenunciables, imprescriptibles, inembargables e inalienables.

UNIVERSIDAD DE LA SABANA

Modelo de predicción de resultados en el Fútbol Profesional Colombiano.

Trabajo de Grado

CARLOS ADOLFO GÓMEZ BAYONA

25/07/2013

**Asesor: Alberto José Naranjo Ramos
Escuela de ciencias Económicas y Administrativas
Economía y Finanzas Internacionales**

Palabras claves: fútbol, predicción, econometría, logit ordenado, goles, puntos, equipos, cortes de probabilidad

Resumen

El fútbol ha sido reconocido como el deporte más popular a nivel mundial, este reconocimiento ha hecho que no solo sea un deporte sino una manera eficaz para las empresas y equipos de crear dinero.

En los últimos años el pronóstico de resultados en este deporte ha crecido de manera significativa generando grandes movimientos de dinero, siendo atractivo para personas naturales e inversionistas, que ven una forma llamativa de aumentar sus ingresos, distinta a inversiones conocidas. Es por esto, que este documento busca obtener el pronóstico de los resultados de los próximos partidos en el fútbol profesional Colombiano (FPC) y saber que equipos ingresaran a la fase final.

La metodología econométrica utilizada para el pronóstico de resultados es por medio de un modelo Logit Ordenado, ya que permite interpretar la variable binaria de una mejor forma. Así, mismo se busca obtener los cortes de probabilidad para así ser utilizados como guía, para determinar en qué área cae el resultado pronosticado. Luego se calculan los resultados por medio de tres modelos (Modelo pronóstico de resultados, Modelo pronóstico goles anotados y Modelo pronóstico goles recibidos).

Abstract

The football has been recognized as the most popular sport worldwide, this award has done that football does not be only a sport but an effective way for businesses and teams can create money.

In recent years forecasting results in this sport has grown significantly generating large movements of money, being attractive to individuals and investors, who see a striking way to increase your income, other than investment known. It is for that reason that this this paper seeks to get the forecast of the results of the next games in the Colombian professional football (FPC) and know that teams from entering into the final phase.

The econometric methodology used for forecasting results is through Ordered Logit model, allowing interpret the binary variable in a better way. Thus, it is desired to obtain the probability cuts so as to be used as guidance to determine in which area falls predicted result. Results are then calculated using three models (forecast results model, forecast goals scored model and forecast goals conceded model).

1. Introducción

Desde la creación del fútbol a finales de la edad media, este deporte ha tomado gran importancia en la sociedad, logrando la creación de importantes campeonatos que han capturado la atención no solo de la sociedad sino de grandes empresas que ven este deporte como una gran zona de inversión.

Según Dobson y Goddard (2001) el interés académico en la economía del deporte se remonta a mediados de los años 50, publicando numerosos libros y artículos acerca de la competitividad y predicción de resultados. A medida que ha pasado el tiempo la investigación sobre la economía del deporte se ha vuelto cada vez más sofisticada, implementando metodologías econométricas para modelar y crear predicciones con probabilidades de acierto más altas. La mayoría de los economistas están de acuerdo dado su importancia social, cultural y económica, el deporte profesional es un área legítima de interés para los investigadores teóricos y empíricos¹.

La importancia económica en el fútbol se está volviendo cada vez más evidente. La gran evolución que ha tenido este deporte ha hecho que la sociedad se integre cada vez más generando un flujo constante de dinero, reactivando una parte importante en la economía de una nación “el gasto”. Rodríguez y Cañaveral (2001) “Es un producto ambivalente en un doble sentido: como bien de consumo (los participantes y espectadores son sus destinatarios finales) y bien intermedio (los fabricantes lo utilizan para la elaboración de otros productos), por un lado, y como bien público y privado, por otro.”² Con esto, el fútbol une dos restricciones implícitas el gasto y el tiempo. Según Heinemann (1998) El gasto se relaciona con el gasto producido por los hogares y el tiempo con otras alternativas de ocio. “Su demanda dependerá de los precios de los bienes y servicios demandados, del tipo, calidad, tecnología y distribución espacial de los bienes y servicios ofertados; de la información existente sobre las posibilidades de practicarlo; de la cantidad y estructura del tiempo disponible y de las condiciones institucionales generales en las que se toman, coordinan y controlan las decisiones”.³

Por tal motivo, la incidencia del fútbol y de cualquier otro deporte en la economía hace equivalencia a tres niveles de análisis: El enfoque macro, micro y meso económico.

¹ Tomado del paper “The economics of Football” realizado por Stephen Dobson y John Goddard en el año 2001. Impreso por la Universidad de Reino Unido, Cambridge.

² Tomado del libro **Deporte y cambio social en el umbral del siglo XXI, Volumen 2. (2001)** Páginas 453-454

³ Tomado del libro **Introducción a la economía del deporte (1998)**. Páginas 91-92

El macroeconómico evalúa la importancia del deporte en este caso el futbol en la economía, el microeconómico indaga sobre el comportamiento de las personas, hogares, empresas y estado y el mesoeconómico analiza el comportamiento del futbol en términos de cadena de producción.

Las predicciones de resultados deportivos dan un punto de partida sobre la recolección de datos que sirvan para analizar no solo el entorno del deporte si no también el ámbito social, las preferencias de los consumidores, el entorno en el que se distribuye este bien, modelar nuevos datos entre otras. Por otro lado, estos pronósticos ayudan a que las personas vean una oportunidad de negocio, favoreciendo reactivar el flujo de dinero y poder ayudar para que la economía no solo dependa de los factores conocidos si no que el deporte y las predicciones tengan un espacio en el ámbito económico.

La gran atención que ha generado el futbol en estos últimos años y la gran importancia que algunos partidos generan, han incrementado las casas de apuestas y la predicción de resultados. Las tres casas de apuestas más importantes en el mundo como lo son BWIN, William Hill y BET 365 están generando grandes movimientos de dinero gracias a las expectativas que están imponiendo diferentes partidos.

En la casa de apuesta BWIN cuenta con más de 30.000 apuestas al día, convirtiéndola como la casa de apuesta número uno a nivel mundial. La mayoría de apostantes tienen una mayor preferencia por el futbol, alcanzando cifras importantes de apuestas en partidos de gran nivel. BWIN aproximadamente posee más de 5 millones de clientes en todo el mundo donde la apuesta mínima es de 0,50 € por jugador logrando ingresos alrededor de 2.500.000 € por partido. El último reporte económico dado por BWIN muestra un aumento del 3% de sus ingresos en los primeros seis meses del año 2012 totalizando 410 millones de euros, BET 365 cuenta con más de 4 millones de clientes en 140 países. En su esquema de negocio la apuesta mínima es de 0,20 € por jugador logrando más de 600.000 € por encuentro y William Hill cuenta con más de 500.000 clientes alrededor del mundo. Para poder ser parte de esta empresa se necesita realizar un depósito de 2 € y la apuesta mínima es de 0,01 €, mostrando ingresos mínimos aproximados de 5.000 euros por encuentro y por suscripción ha recolectado más de 1 millón de euros.⁴

A medida que la competitividad de las ligas aumente la predicción de resultados se vuelven cada vez más complejas, dejando de ser intuición para ser pronósticos apoyados por datos históricos, probabilidad y econometría.

⁴ Tomado de la página oficial de BWIN www.bwin.com , <http://www.bet365.com/es/>, <http://www.williamhill.com/> y también tomado de <http://www.apuestas-deportivas.es/casas-de-apuestas-deportivas.html> el día 25 de octubre de 2012 a las 8:45 pm

Ahora, la predicción de los resultados no solo es fuerte en Europa, en el continente Americano, las ligas de Brasil, Argentina, Uruguay, Ecuador y Colombia son ligas que no son ajenas a la evolución que ha tenido la predicción de partidos. Por ende las casas de apuestas más importantes en el mundo han puesto sus ojos en estas ligas que están tomando gran posicionamiento a nivel internacional.

Debido al gran impacto que ha generado las apuestas deportivas en Europa, Colombia no ha querido quedarse rezagada en este tema, como es el caso de la creación de Gana Gol sistema que ofrecía a las personas opciones y accesibilidad de pronosticar resultados en el torneo colombiano y algunos partidos a nivel internacional.

Por esta razón, este trabajo tiene como objetivo la creación de un modelo de predicción de resultados de los equipos que ingresaran a las finales del fútbol profesional colombiano (FPC)⁵.

Para la creación del modelo de predicción se usara la metodología Logit ordenado y probabilidad, para así obtener los cortes de probabilidad y ver en qué área caen los resultados de cada equipo (Gana, empata o pierde). Así mismo, se aplicara la misma regla de los cortes de probabilidad en el modelo de predicción de goles anotados y goles recibidos, debido que la diferencia de goles juega un papel importante en la clasificación de los equipos, ya que se ha visto que algunos equipos al final del torneo todos contra todos quedan con los mismo puntos y la única forma se saber qué equipo obtiene cupo para la fase final es por medio de la diferencia de goles⁶.

Los modelos están estructurados por las variables que logren capturar la mayor información posible⁷ y poder explicar la variable explicativa. En este trabajo se realizaran tres modelos de pronósticos (Pronostico de puntos, pronóstico de goles recibidos y pronóstico de goles anotados).

Algunos autores han implementado técnicas para poder modelar los datos y llegar a predicciones más acertadas sobre los resultados en los deportes.

Los hermanos Leeds (2007) realizaron un documento llamado *“International Soccer Success and National Institutions”*. En este documento se desarrolló un modelo de éxito nacional, donde el éxito es medido por el número de puntos FIFA que un equipo nacional

⁵ Los equipos más importantes del FPC son: Millonarios, América, Atlético Nacional y Santa Fe por los grandes logros obtenidos a nivel nacional e internacional.

⁶ Diferencia de Goles se obtiene de la resta de los goles anotados por el equipo y los goles recibidos del equipo.

⁷ La mayor información posible en los modelos se puede ver en la significancia que cada variable obtiene al correr el modelo, buscando que el P-valor de la variable se encuentre por debajo de 0.05 a un nivel de confianza del 95%, para que se pueda rechazar la hipótesis nula (Ho: la variable no es significativa).

ha ganado. Ellos usan la ayuda de Mínimos Cuadrados Ordinarios (MCO) y regresión Poisson, donde corroboraron estudios anteriores y encontraron que el éxito se incrementa con el ingreso, la población y haber acogido una competición de copa mundial. También se encuentra que las instituciones políticas del país afecta el rendimiento en el fútbol, así mismo el éxito de los clubes que juegan en la liga de ese país es un buen predictor del éxito que lograra el equipo nacional. Ellos al final del trabajo concluyen que el éxito de los clubes refleja la voluntad del país y su capacidad para financiar el éxito del fútbol.

Esta investigación aporta cosas importantes a la investigación porque me hace ver que variables tomar para la construcción de los modelos. Los hermanos Leeds toman como referencia los puntos según la FIFA estipula, y usan como variable dependiente los puntos obtenidos a lo largo del torneo. La diferencia que existe entre el documento presentado y la investigación es que los puntos acumulados de cada equipo no será la variable dependiente sino formara parte de las variables que expliquen o tengan alguna incidencia en la variable resultado de equipo (variable dependiente) y se le dará un nuevo esquema a los puntos como (1, 0.5, 0) y dejando a un lado los puntos como la FIFA señala (3,1, 0). A su vez se mejorara el modelo propuesto por ellos no solo para ver la incidencia de las variables si no la oportunidad de crear predicción en los resultados de los próximos partidos.

Goddard y Asimakopoulos (2002) desarrollaron un trabajo acerca de modelar los resultados de los partidos de fútbol "*Modelling football match results and the efficiency of fixed-odds betting*". Este trabajo presenta un modelo de regresión basado para explicar y predecir los marcadores de una forma más extensa y completa que otros trabajos realizados anteriormente, tanto en relación con la variable explicativa que se incorporan en los resultados anteriores y la extensión del conjunto de datos que se utiliza para estimar el modelo. Durante la realización de la investigación los dos autores utilizaron el modelo de regresión Probit ordenado sobre datos de los resultados de la liga de fútbol de Inglaterra de los últimos 15 años para poder predecir los próximos resultados de la liga. Goddard y Asimakopoulos resaltan la importancia de los resultados pasados de cada equipo, partidos de liga para los resultados finales, la participación de los equipos en la competición de copa, la distancia geográfica entre los lugares de equipo local y el equipo visitante, estas variables contribuyen al desempeño del modelo. Ellos dan valor de 1 a la victoria, 0.5 al empate y 0 a la derrota. El aporte a la literatura que hacen estos autores es reemplazar la predicción Poisson realizada por Maher (1982) por la metodología Probit. Ya que con la forma de Maher no producía predicción de resultados o marcadores. Se toma de este documento la estructura que le dan a la variable dependiente y algunas variables que son incorporadas para correr el modelo.

Forrest, Goddard y Simmons (2005) publicaron un documento llamado *“odd-setters as forecasters: the case of English football”*. Los investigadores toman una muestra de los juegos disputados entre 1998 y 2003 del fútbol inglés y toman como base un modelo de referencia estadística que incorpora un gran número de variables cuantificables. Los principales hallazgos de este estudio desafían el consenso de que la predicción subjetiva de los expertos será normalmente inferior a las previsiones de los modelos estadísticos. Para realizar el modelo de predicción es usado el modelo Probit ordenado. El modelo pronostica basado únicamente en la información histórica antes de cada partido, dando a las victorias el valor de 1, a los empates el valor de 0.5 y las derrotas el valor de 0. Esta investigación aporta una mejor forma de pronosticar los futuros resultados ya que dan un valor diferente a las victorias, empates y derrotas. Al analizar este documento se hace más evidente que la construcción de la variable dependiente toma los valores 1, 0.5 y 0, aportando más seguridad de implementarla en la investigación. Así mismo, la investigación se diferencia respecto a este estudio el hecho de soportar los modelos con pruebas que aseguren que tan adecuado son los modelos presentados.

El gran aporte que realizaron estos autores a esta investigación es el poder desarrollar de una mejor forma el modelo Logit y Probit en la predicción de los resultados en el fútbol profesional colombiano. Logrando pronosticar resultados en ligas europeas como la liga inglesa, la liga Portuguesa, la liga de Noruega y en el fútbol sudamericano como la liga de Brasil.

En esta investigación se utiliza el método Logit ordenado ya que permite realizar una regresión con una variable binaria que toma un valor de más (0.5), diferente entre 1 y 0 como variable dependiente, para seguir con la misma línea de los artículos anteriores; además analizando los documentos antes mencionados lo distinto de esta investigación es que se usa el mismo modelo Logit pero busca pronosticar los cortes de probabilidad para así analizar con los resultados reales en que área cae el pronóstico obtenido, además un punto importante en este documento es que se busca que tan bueno son los modelos por medio de una pruebas especiales para esta metodología. Otro punto importante es la inclusión de más variables haciendo un poco más robusto los modelos en comparación a los modelos realizados por los otros autores.

Para la creación del modelo de predicción de resultados en el FPC se tomara como documento base el trabajo realizado por Alves, Correia Soares de Mello, Ramos y Sant’Anna (2010)⁸.

⁸“ Logit models for the probability of winning football games”

Alves, Baptista Soares de Mello, Ramos y Parracho Sant'anna (2010) realizaron un estudio sobre la probabilidad de victorias en el futbol "*Logit models for the probability of winning football games*". La investigación es hecha para el campeonato profesional de futbol brasileño con datos del 2008. El objetivo de este trabajo es emplear una simulación del número de puntos que los equipos ganaran y anticipar la clasificación final. El modelo escogido es un modelo Logit ordenado, la cual fue implementada a los 20 equipos que conforman la liga de futbol de Brasil. El resultado encontrado muestra que la probabilidad de una victoria o un empate no depende de la fuerza de los equipos en el largo plazo si no en la fortaleza que está mostrando en el momento. Así mismo, se basan en el modelo realizado por Lawal y Sundheim (2007) y Alves (2008) para mejorar la predicción de los resultados próximos en el futbol Brasileiro. Para la construcción del modelo se usaran los datos de los equipos que conforman la liga colombiana 2013⁹.

Alves, Baptista, Ramos y Parracho cambian la idea de trabajos presentados anteriormente donde la muestra es de varios años. Ellos pronostican dependiendo la actualidad del equipo y torneo y hacen que los pronósticos sean más cercanos a la realidad del club y de la liga.

Se toma de este documento la importancia que se le da a la fortaleza o desempeño que los equipos están teniendo en la actualidad, dejando a un lado los datos pasados. La diferencia con esta investigación es el uso que se le dará a los cortes de probabilidad para analizar en qué área cae la predicción y ver si el equipo consigue un buen resultado e ir sumando puntos en la clasificación general.

En la investigación se realizara la predicción de resultados ex-ante es decir se pronosticara los resultados de las fechas 16, 17 y 18 con los datos actuales y no siguiendo la metodología que la gran mayoría de autores toman de armar datos de años atrás y así pronosticar resultados que ya pasaron.

La siguiente sección describe la estrategia econométrica utilizada, el análisis de los datos, el desarrollo de los tres modelos utilizados para el pronóstico de las próximas fechas y los resultados obtenidos. En la sección 3 se realizara la discusión de la investigación, ya que comprende los pronósticos logrados de los modelos desarrollados y en la última sección se podrá observar las conclusiones conseguidas a lo largo de la exploración de los datos y el desarrollo del modelo en el pronóstico de resultados en el futbol profesional colombiano.

⁹ (Millonarios Futbol Club, Independiente Santa Fe, Atlético Nacional, Deportivo Independiente Medellín, Atlético Junior, Patriotas, Boyacá Chico, Alianza Petrolera, Deportivo Pasto, Quindío, Cúcuta Deportivo, Deportes Tolima, Deportivo Cali, Envigado, Once Caldas, Equidad, Atlético Huila, Itagüí). La liga colombiana de futbol se encuentra dividida en dos torneos cortos (torneo Apertura y torneo Finalización).

2. Estrategia Econométrica

En esta sección se busca implementar la metodología en los datos recolectados del primer semestre del futbol profesional colombiano y así poder obtener los resultados para lograr pronosticar que equipos ingresaran a la segunda fase del torneo colombiano.

2.1 Contexto

Desde la creación de la división mayor del futbol profesional colombiano en 1948 se han entregado 76 títulos oficiales, posicionando la liga colombiana entre una de las ligas más llamativas a nivel internacional.

En lo corrido de los años el esquema de la competición ha cambiado, agregando nuevos equipos a la liga. En el año 2002 la asamblea general de clubes aprobó el ingreso de dos nuevos equipos, contando con 18 equipos profesionales que se ha mantenido hasta el día de hoy. Así mismo, en ese mismo año se decidió dividir el torneo en dos competiciones (torneo apertura y torneo finalización). El campeón de cada torneo logra un cupo directo al torneo más importante de América (La copa Libertadores) y los demás equipos por cantidad de puntos pueden llegar a obtener otro cupo en la Copa Libertadores o la Copa Sudamericana (segundo torneo más importante).

El Futbol Profesional Colombiano (FPC) está conformado de la siguiente manera para los dos torneos (Apertura y Finalización)¹⁰:

Primera Fase:

FPC cuenta con 18 equipos donde se enfrentaran en la primera fase partido de ida y vuelta, donde cada equipo jugara un total de 36 encuentros, conocido como “todos contra todos”.

Segunda Fase (Octogonales):

Los mejores 8 equipos pasaran a la siguiente fase del torneo (conocido como OCTOGONALES). Se conformaran dos grupos cada uno de 4 equipos, los dos mejores equipos de la fase anterior serán los líderes de cada grupo. Así mismo se jugaran partido de ida y vuelta todos contra todos.

¹⁰ Esquema tomado del máximo patrocinador del futbol profesional colombiano <http://www.ligapostobon.com.co/> el día 28 de mayo de 2013 a las 4:12 pm

Tercera Fase (Final):

Los equipos que hayan obtenido el mayor número de puntos en cada grupo en los octogonales, disputaran la final de FPC. La final será de ida y vuelta (un partido en la ciudad de cada equipo).

El torneo también cuenta con un torneo completo, la cual es la suma del torneo Apertura y el torneo Finalización llamado la reclasificación. Este torneo general, indica que equipos descenderán de forma directa (últimos puestos) o que equipo disputara un juego de ida y vuelta con el campeón de la categoría B, para decidir quién entra a la máxima categoría de la Liga Postobon.

2.2 Análisis de datos

Los datos que se utilizaron para el pronóstico de resultados fueron tomados de diferentes fuentes de información deportiva en Colombia¹¹.

Para la creación de los pronósticos se generó una base de datos con 9 variables, la cual corresponde a las variables incluidas en los tres modelos; en la siguiente tabla (Estadística descriptiva) se analizaron la varianza, desviación estándar, media, simetría y curtosis.

Tabla Estadística descriptiva:

Variable	Observaciones	Media	Desviación estándar	Varianza	Simetría	Curtosis
Goles anotados partido anterior	324	1.22	1.062377	1.12	.8918	3.61
Goles recibidos partido anterior	324	1.22	1.062377	1.12	.8918	3.61
Puntos acumulados	324	12.53	8.047978	64.76	.4729	2.50
Resultado equipo	324	.5	.4031609	.1625	0	1.54
Goles acumulados	324	11.61	7.249426	52.55	.3657	2.19

¹¹ Futbolred (www.futbolred.com), El deportivo (www.eldeportivo.com.co), GolGolGol.net (www.golgolgol.net), El colombiano (www.elcolombiano.com) y El Tiempo (www.eltiempo.com)

a favor contrincante						
Posición equipo fecha anterior	324	8.99	5.513556	30.39	-.00011	1.78
Local equipo	324	.5	.5007734	.2507	0	1
Equipo en torneo internacional	324	.166	.37325	.1393	1.7888	4.2
Suma de resultados 5 fechas anteriores	324	1.80	1.369448	1.875	-.06581	1.76

Las variables tienen la siguiente estructura resultado del equipo (1= gana, 0.5= empate y 0= pierde), local (1 si es local y 0 si es visitante), equipo en torneo internacional (1= SI está en torneo internacional y 0= NO está en torneo internacional), y si los partidos fueron transmitidos por televisión cerrada o abierta (1= tv abierta y 0= tv cerrada).

Las variables incluidas en los tres modelos, al analizarlas mostraron que se encuentran con una estructura adecuada, ya que no se modificó ninguna variable porque se quiere encontrar un pronóstico donde los datos no tengan ninguna transformación y ver un resultado sin modificaciones.

2.3 Metodología

Para el pronóstico de los próximos resultados en el fútbol profesional colombiano, se tendrá como base el modelo Logit y Probit, ya que es una herramienta que nos permite conocer las probabilidades de suceso de cada variable independiente ante un cambio en la variable dependiente (binaria).

Llano y Mosquera (2006) en su trabajo de grado¹² *“El modelo Logit, permite obtener estimaciones de la probabilidad de un suceso, identificar los factores de riesgo que determinan dichas probabilidades, así como la influencia o peso relativo que estos tienen*

¹² El modelo Logit una alternativa para medir probabilidad de permanencia estudiantil en Junio de 2006

sobre la misma”¹³ (.....)”La modelización Logit es similar a la regresión tradicional salvo que utiliza como función de estimación la función logística en vez de la lineal”.

Pérez, Kizys y Manzanedo¹⁴ (2006) describen el modelo Logit como una posible solución a las inconsistencias que presentaban el modelo de probabilidad lineal, es decir explicar el comportamiento de una variable dependiente binaria sin necesidad de usar Mínimos Cuadrados Ordinarios (MCO). Ellos argumentan que este tipo de modelos no resulta posible interpretar directamente las estimaciones de los parámetros betas (β) ya que son modelos no lineales. “Lo que haremos en la práctica es fijarnos en el signo de los estimadores. Si el estimador es positivo, significara que incrementos en la variable asociada causan incrementos en $P(Y=1)$. Por el contrario, si el estimador muestra un signo negativo, ellos supondrá que incrementos en la variable asociada causaran disminuciones en $P(Y=1)$ ”.

Función Logística:

$$f(x) = \frac{\exp(x)}{1 + \exp(x)}$$

Se tendrá que:

$$E[Y] = P(Y = 1) = \frac{\exp(\beta_1 + \beta_2 X_2 + \dots + \beta_k X_k)}{1 + \exp(\beta_1 + \beta_2 X_2 + \dots + \beta_k X_k)}$$

El modelo Logit presenta los odds, la cual son los que miden la probabilidad de ocurrencia de que paso algo sobre la variable dependiente:

$$Odds = \frac{P(Y = 1)}{1 - P(Y = 1)} = \exp(\beta_1 + \beta_2 X_2 + \dots + \beta_k X_k)$$

Para volver esta fórmula en una expresión lineal se sacan los logaritmos:

¹³ <http://www.bdigital.unal.edu.co/1038/1/laurarosallanodiaz.2006.pdf> tomado el 13 de noviembre de 2012 a la 1:07 am.

¹⁴ <http://www.uoc.edu/in3/emath/docs/RegLogistica.pdf> tomado el 13 de noviembre de 2012 a la 1:23 am en el documento llamado “Regresión Logística Binaria”

$$\text{Logit} [P(Y = 1)] = \text{Ln} \left(\frac{P(Y=1)}{1-P(Y=1)} \right) = \exp(\beta_1 + \beta_2 X_2 + \dots + \beta_k X_k)$$

De acuerdo a Pérez, Kizys y Manzanedo (2006), “cuando se hace referencia al incremento unitario en una de las variables explicativas del modelo, aparece el concepto de odds-ratio como el cociente entre los dos odds asociados”

$$\text{Odds - Ratio} = \frac{\text{Odds 2}}{\text{Odds 1}} = \exp(\beta_1)$$

“Con la formula anterior, se deduce que un cociente β_1 cercano a 0, equivalentemente, un odds-ratio cercano a 1 significara que cambios en la variable explicativa X_1 asociada no tendrá efecto alguno sobre la variable dependiente Y .”¹⁵

Los modelos multinomiales tienen el siguiente esquema:

$$P (X_1 = x_1, X_2 = x_2, X_3 = x_3 \dots) = \frac{n!}{X_1! * X_2! * X_3! * \dots} P_1^{x_1} * P_2^{x_2} * P_3^{x_3} * \dots$$

Dado el esquema que se incorporó en la investigación del valor que se le daba a la victoria, empate y derrota (1, 0.5, 0) no es posible correr los datos de forma multinomial ya que el programa base (STATA) no reconoce los valores dados y son tomados como un modelo Logit ordenado. El esquema tomado se debe a la utilización de este en los artículos bases.

Al correr los modelos pronóstico de goles recibidos y pronóstico de goles realizados, se busca pronosticar la cantidad de goles que se anotaran en el encuentro e ir sumando en la tabla de posiciones los resultados, ya que una característica que tiene el balompié colombiano por su estructura de juego, es que varios equipos al final de la temporada finalizan con la misma cantidad de puntos, por tal motivo, la diferencia de goles es la

¹⁵ Tomado de la pagina de internet http://www.uoc.edu/in3/emath/docs/T10_Reg_Logistica.pdf el dia 8 de noviembre del año 2012 a las 8:15 pm

siguiente etapa que se tiene en cuenta para determinar que equipos pasaran a la siguiente fase del torneo.

Los modelos son conformados por las siguientes variables:

Modelo pronóstico de puntos (victoria-empate-derrota)

Ologit Resultado del equipo

$$\begin{aligned} &= \beta_1 \text{ Local equipo} + \beta_2 \text{ Goles anotados partido anterior} \\ &+ \beta_3 \text{ Equipo en torneo Internacional} \\ &+ \beta_4 \text{ Goles acumulados a favor del contricante} \\ &+ \beta_5 \text{ Puntos acumulados del equipo} \\ &+ \beta_6 \text{ Suma de los resultados 5 fechas anteriores} \\ &+ \beta_7 \text{ Posicion del equipo en la fecha anterior} \end{aligned}$$

Modelo pronóstico de goles anotados

Ologit Goles anotados partido

$$\begin{aligned} &= \beta_1 \text{ Goles recibidos partido} + \beta_2 \text{ Local equipo} \\ &+ \beta_3 \text{ Resultado del Equipo} + \beta_4 \text{ Equipo en torneo Internacional} \end{aligned}$$

Modelo goles recibidos

Ologit Goles recibidos partido

$$\begin{aligned} &= \beta_1 \text{ Local Equipo} + \beta_2 \text{ Goles anotados Partido} \\ &+ \beta_3 \text{ Goles acumulado a favor contricante} \\ &+ \beta_4 \text{ Puntos Acumulados del equipo} \\ &+ \beta_5 \text{ Posicion equipo fecha anterior} + \beta_6 \text{ Resultado del Equipo} \end{aligned}$$

2.4 Análisis de resultados

A continuación se analizaran los resultados conseguidos buscando que los modelos cumplan con las pruebas específicas de un modelo correcto. Así mismo, se usaran los resultados en la predicción de los próximos partidos.

2.4.1 Modelo pronostico puntos

En cualquier deporte a nivel mundial, los puntos hacen parte de la estructura de cada competición. En el caso del futbol, los puntos son distribuidos en: tres puntos (3) para la victoria, un punto (1) para el empate y cero puntos (0) para la derrota; es por esto que el modelo pronóstico de puntos es el más importante en este trabajo, ya que nos indicara que equipos logran la victoria, empate o derrota y así sumar en la tabla de posiciones los puntos obtenidos.

Numero de Observaciones	270
LR Chi2	220.26
Prob > Chi2	0.0000
Pseudo R2	0.3723
Log Likelihood	-185.66514

Variable Dependiente (Resultado Equipo)				
Variables	Coficiente	Std. Err.	Z	P> z
Local Equipo	.8111283	.27801	2.92	0.004
Goles anotados partido	1.672686	.1846074	9.06	0.000
Equipo en torneo Internacional	.2369499	.3946388	0.60	0.548
Goles acumulados a favor contrincante	-.2786224	.0361555	-7.71	0.000
Puntos acumulados	.2946759	.0437697	6.73	0.000
Suma de resultados 5 fechas anteriores	-.4384276	.1545298	-2.84	0.005
Posición equipo fecha anterior	.1252119	.0314928	3.98	0.000

Cortes	Coficiente
Cut 1	1.736615

Cut 2	4.790841
-------	----------

Los cortes que se ven en la regresión, se usaran más adelante en el pronóstico de los resultados de las próximas fechas en la liga Colombiana de futbol.

Como se observa en el modelo las variables ingresadas son significativas, ya que su P-valor con un nivel de confianza del 95% y un error tipo I del 5%, se rechaza la hipótesis nula (H_0 : los betas no son significativos).

Al analizar los coeficientes de cada variable, vemos que todos los coeficientes son acordes a la teoría del futbol. (Ver anexo 1)

Al ver que el modelo es consistente, donde la mayoría de las variables son significativas y tienen sentido en el impacto de la variable dependiente, se procede a buscar si el modelo es robusto, es decir, si lo que se pronosticó tiene un porcentaje alto de cobertura. Se correrá el mismo modelo, pero sin tener en cuenta los cortes que el modelo anterior nos daba, buscando obtener los odds ratio ya mencionados. (Ver anexo 2)

Como se observa en el modelo, los odds ratio nos indica la probabilidad de que se genere el evento, confirmando lo pronosticado en el primer modelo.

Cuando el coeficiente es mayor a 1, indica que la probabilidad de ocurrencia es alta y cuando el coeficiente es menor a 0 señala que la probabilidad de acierto es menor.

Una característica de los modelo Logit es que los errores aleatorios pueden no seguir una distribución normal (No se logra normalidad) y los errores no son homoscedasticos, incumpliendo dos de los supuesto de Gauss Markov, los errores tienen varianza constante y normalidad en los residuales. (Ver anexo 3)

La prueba estat classification proyecta que tan sólido se encuentra el modelo, cuando esta prueba muestra un valor por encima del 80% indica que el modelo a capturado una muy buena proporción de información, apoyando la credibilidad de los resultados señalados por el modelo.

Al realizar la prueba estat classification, nos muestra que el 81.48% de los datos fueron correctamente clasificados, dando como resultado un modelo creíble y fuerte para poder empezar a crear los pronósticos de las próximas fechas del Futbol Profesional Colombiano (FPC). (Ver anexo 4). Así mismo, se realiza la prueba Estat Gof (Goodness-of-fit test), para ver si el modelo se encuentra bien especificado. En esta prueba se busca no rechazar la hipótesis nula (H_0 : El modelo se encuentra bien especificado). (Ver anexo 5)

2.4.2 Modelo Pronostico Goles Anotados

El modelo goles anotados es un modelo que nos permite identificar cuantos goles realizara el equipo durante la próxima fecha. Con el resultado obtenido se sumaran los goles pronosticados en la tabla de posiciones e ir viendo que equipos estarán en los primeros lugares.

Numero de Observaciones	270
LR Chi2	262.65
Prob > Chi2	0.0000
Pseudo R2	0.3417
Log Likelihood	-252.9996

Variable Dependiente (Goles Anotados Partido)				
Variables	Coficiente	Std. Err.	Z	P> z
Goles recibidos partido	1.94631	.1936877	10.05	0.000
Local equipo	.6403446	.2536829	2.52	0.012
Resultado equipo	7.776075	.623323	12.48	0.000
Equipo en torneo Internacional	.5656927	.3291213	1.72	0.086

Cortes	Coficiente
Cut 1	4.966506
Cut 2	7.748903
Cut 3	10.43796
Cut 4	12.56454
Cut 5	15.25861

El modelo pronostico goles anotados muestra que las variables incluidas son significativas a un nivel de significancia del 95% y un error tipo I del 5%. Rechazando la hipótesis nula (Ho: Los betas no son significativos).

Un punto importante del modelo Logit es que nos brinda los cortes de probabilidad, ya que con estos cortes podremos ubicar los resultados de los equipos y dar un pronóstico más acertado.

Las variables usadas en el modelo de predicción de goles anotados resultaron ser teóricamente bien estipuladas (Ver anexo 6)

Ahora se realiza el mismo modelo quitando los cortes de probabilidad para poder ver los Odds-Ratio y comprobar que el modelo pasado estuvo bien pronosticado y ver la probabilidad de evento que cada variable ejerce sobre la variable (Goles anotados partido). (Ver anexo 7)

Cada variable del modelo de los goles anotados partido, tienen los odds ratio por encima de cero, queriendo decir que la probabilidad de ocurrencia es alta, garantizando los resultados del modelo pronosticado con los cortes de probabilidad.

Para ver si realmente el modelo explica la variable dependiente de una forma adecuada y corroborar que el modelo propuesto es creíble, se realiza la prueba estat classification. (Ver anexo 8)

El 86.67% de los datos fueron clasificados correctamente, mostrando que el modelo realizado se encuentra como un modelo confiable.

Se realiza la prueba goodness-of-fit (estat gof), buscando no rechazar la hipótesis nula de que el modelo es encuentra correcto. (Ver anexo 9)

2.4.3 Modelo Pronostico Goles Recibidos

El modelo pronostico goles recibidos nos proporciona una idea de cuantos goles los equipos recibirán durante el próximo partido. Esto es importante porque al tener los goles anotados y los goles recibidos podemos obtener la diferencia de goles de cada equipo y esta diferencia de goles ayudara que equipos ingresaran en la fase final del futbol profesional colombiano ya que los puntos no son la única referencia del posicionamiento de los equipos.

Numero de Observaciones	270
LR Chi2	270.85
Prob > Chi2	0.0000
Pseudo R2	0.3504

Log Likelihood	-266.35247
-----------------------	------------

Variable Dependiente (Goles Recibidos Partido)				
Variabes	Coeficiente	Std. Err.	Z	P> z
Local equipo	-.6572	.2561	-2.57	0.010
Goles anotados partido	1.868	.1970	9.48	0.000
Goles acumulados a favor contrincante	.08588	.0313	2.74	0.006
Puntos acumulados	-.03967	.0316	-1.25	0.210
Posición equipo fecha anterior	-.04773	.0278	-1.71	0.087
Resultado equipo	-7.4612	.6500	-11.48	0.000

Cortes	Coeficiente
Cut 1	-3.58611
Cut 2	-.682361
Cut 3	2.00035
Cut 4	3.95498
Cut 5	6.83764

El modelo comprendido por estas variables, mostro que las variables son significativas, y teóricamente están bien a un nivel de significancia del 95% y un error tipo I del 5%. Rechazando la hipótesis nula (Ho: Los betas no son significativos), ya que el objetivo principal con este modelo es pronosticar cuantos goles recibirá el equipo en la próxima fecha.

El R-cuadrado obtenido en el modelo, refleja que las variables independientes están explicando en gran proporción la variable dependiente (Goles recibidos partido), con un porcentaje de 40.69%. Así mismo, la interpretación de las variables es acorde. (Ver anexo 10)

Después de analizar los resultados presentados por el modelo pronóstico de goles recibidos, se procede a estudiar los odds ratio para comprobar que los resultados demostrados anteriormente se encuentren bien especificados. (Ver anexo 11)

Se realiza la prueba estat classification y la prueba estat gof, para ver si el modelo es adecuado para el pronóstico de los goles recibidos de cada equipo durante la siguiente jornada. (Ver anexo 12)

Al ver el resultado de la prueba estat classification, se observa que el modelo se encuentra por encima del 80% dando señal de un modelo adecuado para el pronóstico de los goles recibidos de cada equipo.

Así mismo, la prueba estat gof, nos indica que el modelo cumple los parámetros de un modelo correcto, ya que no se rechaza la hipótesis nula a un nivel de confianza del 95% y un error tipo I del 5% (valor obtenido > 0.005).

Ya obtenidos los resultados, y representar los modelos bien definidos, para poder pronosticar las tres próximas fechas del Fútbol Profesional Colombiano se tienen en cuenta los cortes de cada modelo para llevar como guía ante los pronósticos de la fecha 16, 17 y 18.

3. Discusión

Para empezar a pronosticar que equipos ingresaran a la segunda fase del FPC conocida como octogonales, se usan los cortes de probabilidad ya encontrados anteriormente en cada modelo. Los cortes de probabilidad tienen como función ubicar el resultado obtenido de los próximos encuentros y señalar si el equipo lograra la victoria, empate o derrota. Así mismo, para los goles recibidos y los goles anotados.

Cortes de probabilidad:

Cortes	Modelo pronostico Puntos	Modelo Pronostico Goles anotados	Modelo Pronostico Goles recibidos
1	1.736615	4.966506	-3.58611
2	4.790841	7.748903	-0.68236
3		10.43796	2.00035
4		12.56454	3.95498
5		15.25861	6.83764

Modelo pronóstico de puntos:

Si el resultado está por debajo de 1.736615 el equipo perderá el partido, si el resultado está entre 1.736615 y 4.790841 el equipo empatara el partido y si está por encima de 4.790841, el equipo ganara el partido.

Modelo pronóstico de goles anotados:

Si el resultado está por debajo del corte 1, el equipo anotara 0 goles, entre el corte 1 y el corte 2, el equipo hará 1 gol, entre el corte 2 y el corte 3, el equipo hará 2 goles, entre el corte 3 y el corte 4, el equipo hará 3 goles, entre el corte 4 y el corte 5, el equipo hará 4 goles y si está por encima del corte 5, el equipo hará más de 5 goles.

Modelo pronóstico de goles recibidos:

Si el resultado está por debajo del corte 1, el equipo recibirá 0 goles, entre el corte 1 y el corte 2, el equipo recibirá 1 gol, entre el corte 2 y el corte 3, el equipo recibirá 2 goles, entre el corte 3 y el corte 4, el equipo recibirá 3 goles, entre el corte 4 y el corte 5, el equipo recibirá 4 goles y si está por encima del corte 5, el equipo recibirá más de 5 goles. Para predecir la fecha 16, se toman las variables que se usaron para el modelo pronóstico de puntos con sus respectivos coeficientes y se busca la información real de los equipos. Seguido se multiplica cada coeficiente por cada dato real y se suman los resultados, el dato total se compara con los cortes obtenidos y se decide en que zona cae (victoria, empate o derrota). Así mismo, para los goles anotados y los goles recibidos.

$$Pronostico = \sum \beta_t * dato conocido$$

Los resultados obtenidos de la fecha 16 se pueden ver en el (Anexo 13).

En la tabla de la predicción de puntos, la fila predicción tiene el siguiente esquema:

Victoria se identifica con una (G), empate con una (E) y derrota con una (P).

Analizado los datos obtenidos se construye la próxima tabla de posiciones de la fecha 16 del Fútbol Profesional Colombiano. Se suman los puntos conseguidos a cada equipo a la tabla de clasificación de la fecha 15, tres puntos al equipo que obtenga la victoria, un punto al equipo que consiga el empate y cero puntos al equipo que logre la derrota, de esa manera se empieza a tener una idea de que equipos conseguirán un cupo a los mejores ocho equipos de la temporada.

También se realiza el pronóstico de goles anotados y goles recibidos de cada encuentro y se suman a la nueva tabla (fecha 16). (Ver anexo 14)

Al tener los pronósticos y la nueva tabla de posiciones de la fecha 16 se procede a realizar los nuevos pronósticos de la fecha 17 y la fecha 18. (Ver anexo 15)

La tabla de posiciones de la fecha 17 y la fecha 18 de pueden observar en el (Anexo 16). Los resultados de las dos últimas fechas fueron pronosticados de igual forma como se pronosticó la fecha 16.

Así los ocho equipos que ingresaran a la segunda fase (octogonales) del torneo Colombiano son los siguientes:

Tolima, Itagüí, Santa Fe, Deportivo Cali, Millonarios, Once Caldas, Nacional y Cúcuta.

Al analizar los resultados y hallar los equipos que conformaran la segunda fase, se hubiese deseado incluir más variables para poder hacer que el error sea lo más pequeño posible, ya que mucha más información puede ayudar a controlar este error y pronosticar los resultados más acertadamente.

4. Conclusiones

Los modelos de predicción, se han convertido cada vez más complejos, logrando recoger la mayor información disponible para poder predecir de forma más precisa y confiable los resultados próximos.

Los modelos Logit y Probit se han vuelto herramientas importantes para usar en la creación de pronósticos tanto para la economía del deporte como para el entorno económico. El modelo Logit ordenado, se ha convertido en una gran herramienta, ya que genera unos cortes de probabilidad, la cual serán utilizados como guía en el pronóstico de resultados.

El pronóstico de resultados ha tomado gran importancia en el mundo, que hoy en día se puede observar grandes esfuerzos en poder predecir con una mayor certeza eventos futuros. En esta investigación el modelo Logit ordenado es la herramienta fundamental para pronosticar que equipos ingresaran a la segunda fase del futbol profesional colombiano, logrando capturar con una gran certeza que equipos ganaran, empataran o perderán sus próximos partidos, así mismo, pronosticar cuantos goles anotaran y recibirán cada equipo en el próximo encuentro.

El modelo logra recoger la mayor cantidad de variables posibles, para así poder generar unos pronósticos más confiables. Algunos datos no fueron posibles encontrarlos, ya que hacen parte de la información confidencial de los equipos y la federación colombiana de futbol.

Con este trabajo se afianza cada vez más, en la aplicación de la economía y la probabilidad en asuntos deportivos, gracias a la gran importancia que ha estado generando el futbol en los últimos tiempos.

La gran contribución hecha en este trabajo es el pronóstico de los goles que serán anotados y recibidos, ya que pronosticar los equipos que ingresaran a la fase final, no solo pasan por puntos si no por la diferencia de goles.

Así mismo, la investigación tuvo dificultades, ya que la liga colombiana no posee la cantidad de datos necesarios para realizar la investigación, así mismo, uno de los objetivos principales en este trabajo es ingresar la mayor cantidad de datos posibles en el modelo, para que este tenga un mayor porcentaje de acierto. Desafortunadamente algunas variables incluidas en los modelos no resultaron ser significativas y por consecuencia hace

que el modelo no tenga el máximo porcentaje de acierto deseado. Otro limitante que se encontró en la investigación es controlar el error generado por el modelo, ya que predecir los resultados tiene variables que no son posibles cuantificar.

Los modelos propuestos proporcionan confiabilidad ya que se ha analizado el mismo modelo quitando algunas variables y se observa que las pruebas de confiabilidad no superan los porcentajes requeridos para obtener un modelo adecuado.

Bibliografía

Revista El Colombiano (2013), Sección Deportes, www.elcolombiano.com.co

Revista El Deportivo (2013), www.eldeportivo.com.co (versiones digitales)

División Mayor del Fútbol Colombiano (DIMAYOR), www.dimayor.com, estadísticas y Liga Postobon.

Página oficial del fútbol profesional colombiano (Liga Postobon), www.ligapostobon.com.co, estadísticas.

FútbolRed, www.futbolred.com, tabla de posiciones liga Postobon

GolGolGol.net, www.golgolgol.net, página encargada de recopilar estadística de los equipos y fútbol profesional colombiano.

Long y Freese (2001), "Regression models for categorical dependent variables using stata", Estados Unidos – Texas.

Leeds (2007), "International Soccer Success and National Institutions"

Brillinger (2006), "Modelling some Norwegian soccer data".

Goddard y Asimakopulos (2003) "Modelling football match results and the efficiency of fixed-odds betting" <http://www.stat.berkeley.edu/~aldous/157/Papers/goddard.pdf>

Forrest, Goddard y Simmons (2005), "Odd-setters as forecasters: the case of English football"

Goddard (2005), "Regression models for forecasting goals and match results in association football"

Alves, Soares de Mello, Ramos y Sant'Anna (2010), "Logit models for the probability of winning football games".

Medina (2003), "*Modelos de elección discreta*" tomado de la página http://www.uam.es/personal_pdi/economicas/eva/pdf/logit.pdf

Wooldridge (2009) "*Introducción a la econometría: Un enfoque moderno*" Cuarta edición. Cengage Learning

Gujarati (2007) "*Econometría básica*" Cuarta edición. McGraw-Hill

Anexos

Anexo 1

Local equipo: A medida que el equipo juegue de local, la probabilidad de obtener la victoria es mayor.

Goles anotados partido anterior: Cuando el equipo obtiene goles en la fecha anterior, la probabilidad de que anote el próximo partido es mayor, dando como resultado una gran ventaja de obtener la victoria.

Equipo en torneo internacional: Si el equipo está participando en un torneo internacional, le da mayor probabilidad de obtener la victoria, ya que el plantel se concentrará en realizar un buen papel en el torneo internacional y eso hará que el equipo mejore su técnica de juego y consigan una actuación mejorable en el torneo local.

Goles acumulados a favor del contrincante: El coeficiente obtenido en el modelo, nos indica que a medida que el equipo contrario tiene más goles a favor la probabilidad de que el equipo obtenga la victoria es mucho menor.

Puntos acumulados del equipo: A mayores puntos acumulados, la probabilidad de obtener la victoria es mucho mayor.

Suma de resultados cinco fechas anteriores: A medida que el equipo tenga más fechas de victorias, genera una probabilidad de perder, ya que las condiciones emocionales afectan al plantel a medida que pasan las fechas.

Posición del equipo en la fecha anterior: Cuando el equipo tiene una posición favorable en la tabla de posiciones, la probabilidad de obtener la victoria en la siguiente fecha es mucho más alta.

Anexo 2

Odds Ratio Modelo pronóstico de Puntos:

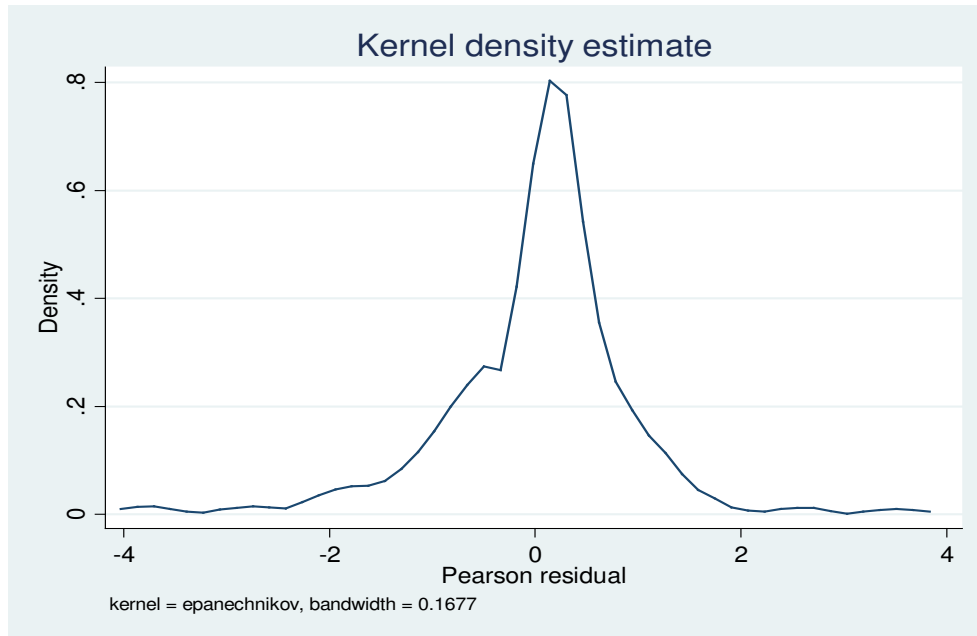
Numero de Observaciones	270
LR Chi2 (7)	119.13
Prob > Chi2	0.0000

Pseudo R2	0.3525
Log Likelihood	-109.3885

Variables	Odds Ratio	Std. Err.	Z	P> z
Variable Dependiente (Resultado Equipo)				
Variables Independientes				
Local Equipo	2.624665	.9166503	2.76	0.006
Goles anotados partido	3.862189	.8712647	5.99	0.000
Equipo en torneo Internacional	1.357116	.6914662	0.60	0.549
Goles acumulados a favor contrincante	.777898	.0345811	-5.65	0.000
Puntos acumulados	1.336198	.076842	5.04	0.000
Suma de resultados 5 fechas anteriores	.6391609	.1241442	-2.30	0.021
Posición equipo fecha anterior	1.13651	.0443741	3.28	0.001

Anexo 3

Distribución de los residuales del modelo pronóstico de puntos.



Aunque en el modelo Logit no es posible corregir normalidad y heterocedasticidad se observa como los datos poseen una distribución normal adecuada.

Anexo 4

Prueba Estat Classification para el modelo pronóstico de puntos:

		+	-	Total
Classified	+	163	29	192
	-	21	57	78
Total		184	86	270

Correctly classified	81.48%
----------------------	--------

Anexo 5

Prueba estat gof (goodness-of-fit test) para el modelo pronóstico de puntos:

Number of observations	270
Number of covariate patterns	270
Pearson chi2(311)	218.48
Prob > chi2	0.9922

Al ver el resultado nos muestra que el modelo se encuentra bien especificado ya que no se rechaza la hipótesis nula a un nivel de significancia del 95% y un error tipo I del 5%.

Anexo 6

Goles recibidos partido: A mayor cantidad de goles recibidos, la probabilidad de que el equipo anote goles es alta, ya que hace que los goles recibidos motivan a los jugadores a buscar un mejor resultado en el próximo encuentro.

Local equipo: Cuando el equipo es local, la probabilidad de hacer goles es alta, porque tendrá el apoyo de la afición.

Resultado del Equipo: Si el equipo en la fecha pasada obtuvo una victoria, la probabilidad de que consiga anotar goles en la próxima fecha es muy alta.

Equipo en torneo Internacional: Cuando el equipo se encuentra en un torneo internacional, la experiencia de estar disputando el torneo, hace que los delanteros mejore su efectividad y puedan tener una probabilidad alta de marcar goles en la próxima fecha del torneo local.

Ver anexo 7

Odds ratio modelo pronóstico goles anotados:

Numero de Observaciones	270
LR Chi2 (7)	149.24
Prob > Chi2	0.0000
Pseudo R2	0.4502
Log Likelihood	-85.965611

Variables	Odds Ratio	Std. Err.	Z	P> z
Variable Dependiente (Goles Anotados Partido)				
Variables Independientes				
Goles recibidos partido	6.4887	1.7414	6.97	0.000
Local equipo	1.7483	.6841	1.43	0.153
Resultado equipo	3289.414	3387.8	7.86	0.000
Equipo en torneo Internacional	1.6039	.9062	0.84	0.403

Anexo 8

Prueba estat classification para el modelo pronóstico goles anotados:

		+	-	Total
Classified	+	168	16	184
	-	20	66	86
Total		188	82	270

Correctly classified	86.67%
-----------------------------	---------------

Anexo 9

Prueba estat gof (goodness-of-fit test):

Number of observations	270
Number of covariate patterns	40
Pearson chi2(311)	129.69
Prob > chi2	0.0000

Anexo 10

Local Equipo: Cuando el equipo no es local, la probabilidad de recibir goles es alta, ya que el factor de la afición influirá mucho en el resultado del encuentro.

Goles anotados partido: A medida que el equipo haya hecho goles en la fecha anterior, la probabilidad de que la próxima fecha reciba goles es alta, porque el factor confianza puede jugar en contra al equipo

Goles acumulados a favor contrincante: Cuando el equipo contrincante tenga una gran cantidad de goles acumulados a favor, la probabilidad de que el equipo local reciba goles en la próxima fecha es alta.

Puntos acumulados: Mayor puntos acumulados del equipo, la probabilidad de que reciban goles es muy baja.

Posición del equipo fecha anterior: Cuando el equipo tiene una posición privilegiada, la probabilidad de que le hagan goles es mucho menor.

Resultado del Equipo: Si en la fecha anterior el equipo logro un resultado favorable, la probabilidad de que el próximo partido reciba goles es mínima.

Anexo 11

Odds ratio modelo pronóstico goles recibidos:

Numero de Observaciones	270
LR Chi2 (7)	164.59
Prob > Chi2	0.0000
Pseudo R2	0.5016
Log Likelihood	-81.782341

Variables	Odds Ratio	Std. Err.	Z	P> z
Variable Dependiente (Goles Recibidos Partido)				
Variables Independientes				
Local equipo	.5349	.2277	-1.47	0.142
Goles anotados partido	6.580	1.9063	6.50	0.000
Goles	1.177	.0669	2.86	0.004

internacional																		
Goles acumulados a favor contrincante	21	- 5.85	17	- 4.73	26	- 7.24	24	- 6.68	15	- 4.17	24	- 6.68	12	- 3.34	21	- 5.85	27	- 7.52
Puntos acumulados	16	4.71	21	6.18	17	5.00	18	5.30	26	7.66	22	6.48	17	5.00	24	7.07	24	7.07
Suma de resultados 5 fechas anteriores	2	- 0.87	1.5	- 0.65	2	- 0.87	3	- 1.31	3	- 1.31	1.5	- 0.65	2.5	- 1.09	3.5	- 1.53	3	- 1.31
Posición equipo fecha anterior	14	1.75	10	1.25	13	1.62	11	1.37	2	0.25	7	0.87	12	1.50	4	0.50	5	0.62
Resultado		5.56		4.52		4.34		2.83		6.57		7.51		4.55		0.99		3.01
Predicción		E		E		E		E		G		G		E		P		E

Anexo 14

Pronostico goles anotados

	1		2		3		4		5		6		7		8		9	
Local Equipo	1	1.94	1	1.94	1	1.94	1	1.94	1	1.94	1	1.94	1	1.94	1	1.94	1	1.94
Goles anotados partido	0	0	2	1.28	0	0	1	0.64	2	1.28	0	0	0	0	0	0	1	0.64
Resultado equipo	0.5	3.88	1	7.77	0.5	3.88	0	0	1	7.77	0.5	3.88	1	7.77	0	0	0	0
Equipo en torneo internacional	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Resultado		5.83		11.0		5.83		2.58		11.0		6.47		9.72		1.94		2.58
Predicción Goles Anotados		1		3		1		0		3		1		2		0		0

Pronostico goles recibidos

	1		2		3		4		5		6		7		8		9	
--	---	--	---	--	---	--	---	--	---	--	---	--	---	--	---	--	---	--

Local Equipo	1	-	1	-	1	-	1	-	1	-	1	-	1	-	1	-	1	-
		0.68		0.68		0.68		0.68		0.68		0.68		0.68		0.68		0.68
Goles anotados partido	0	2.30	0	2.30	0	2.30	2	2.30	1	2.30	0	2.30	1	2.30	0	2.30	1	
Goles acumulados a favor contrincante	17	0.15	21	0.15	21	0.15	19	0.15	23	0.15	16	0.15	16	0.15	8	0.15	24	0.15
Puntos acumulados	28	-	19	-	18	-	23	-	12	-	22	-	21	-	18	-	20	-
		0.13		0.13		0.13		0.13		0.13		0.13		0.13		0.13		0.13
Posición equipo fecha anterior	2	-	14	-	13	-	6	-	17	-	8	-	10	-	12	-	0	-
		0.05		0.05		0.05		0.05		0.05		0.05		0.05		0.05		0.05
Resultado equipo	0	-	0	-	0	-	1	-	0.5	-	0.5	-	0.5	-	0	-	0.5	-
		8.37		8.37		8.37		8.37		8.37		8.37		8.37		8.37		8.37
Resultado		-		-		-		-		-		-		-		-		-
		1.86		0.75		0.56		4.88		1.57		5.76		3.44		2.52		2.00
Predicción Goles Recibidos		0		2		1		0		1		0		0		0		1

Anexo 15

Tabla de posiciones Fecha 16

Fecha 16								
Equipos	PJ	PG PRONOSTICO	PE PRONOSTICO	PP PRONOSTICO	GF PRONOSTICO	GC PRONOSTICO	GD PRONOSTICO	Puntos PRONOSTICO
Tolima	16	8	5	3	31	19	12	29
Santa Fe	16	8	5	3	27	21	6	29
Itagui	16	7	6	3	22	15	7	27
Nacional	16	6	8	2	24	19	5	26
Millonarios	16	7	4	5	21	15	6	25
Pasto	16	6	7	3	19	22	-3	25
Cali	16	5	9	2	27	20	7	24
Once Caldas	16	6	6	4	19	15	4	24
Cúcuta	16	6	4	6	22	18	4	22
Junior	16	6	3	7	24	22	2	21
Chicó F.C.	16	4	7	5	24	24	0	19
Huila	16	4	7	5	20	23	-3	19
Envigado FC	16	5	4	7	14	21	-7	19
Medellín	16	4	6	6	17	17	0	18
La Equidad	16	4	6	6	15	15	0	18
Alianza Petrolera	16	3	5	8	15	28	-13	14
Patriotas FC	16	0	11	5	14	23	-9	11
Quindío	16	1	5	10	5	23	-18	8

Anexo 15

Tabla de posiciones Fecha 17 y Fecha 18

Fecha 17								
	PJ	PG PRONOSTICO	PE PRONOSTIC O	PP PRONOSTIC O	GF PRONOSTIC O	GC PRONOSTIC O	GD PRONOSTIC O	Puntos PRONOSTICO
Tolima	17	9	5	3	33	20	13	32
Santa Fe	17	8	6	3	28	22	6	30
Itagui	17	7	6	3	24	17	7	28
Nacional	17	6	9	2	25	20	5	27
Once Caldas	17	7	6	4	20	15	5	27
Pasto	17	6	8	3	20	23	-3	26
Cali	17	5	10	2	28	21	7	25
Millonarios	17	7	4	6	22	17	5	25
Cúcuta	17	6	5	6	23	19	4	23
Junior	17	6	4	7	24	22	2	22
Huila	17	5	7	5	21	23	-2	22
Chicó F.C.	17	4	8	5	26	26	0	20
Medellín	17	4	7	6	18	18	0	19
La Equidad	17	4	7	6	15	15	0	19
Envigado FC	17	5	4	8	14	22	-8	19
Patriotas FC	17	1	11	5	15	23	-8	14
Alianza Petrolera	17	3	5	9	15	29	-14	14
Quindío	17	1	5	11	5	24	-19	8

Fecha 18									
Posicion	Equipos	PJ	PG PRONOSTIC O	PE PRONOSTIC O	PP PRONOSTIC O	GF PRONOSTIC O	GC PRONOSTIC O	GD PRONOSTIC O	Puntos PRONOSTICO
1	Tolima	18	10	5	3	35	21	14	35
2	Itagui	18	8	6	3	26	18	8	31
3	Santa Fe	18	8	6	4	28	25	3	30
4	Cali	18	6	10	2	29	21	8	28
5	Millonarios	18	8	4	6	23	17	6	28
6	Once Caldas	18	7	7	4	21	16	5	28
7	Nacional	18	6	9	3	25	22	3	27
8	Cúcuta	18	7	5	6	25	19	6	26
9	Pasto	18	6	8	4	21	25	-4	26
10	Huila	18	6	7	5	24	23	1	25
11	Junior	18	6	5	7	25	23	2	23
12	Chicó F.C.	18	5	8	5	27	26	1	23
13	Medellín	18	4	8	6	19	19	0	20
14	La Equidad	18	4	7	7	16	17	-1	19
15	Envigado FC	18	5	4	9	14	23	-9	19
16	Patriotas FC	18	1	12	5	16	24	-8	15
17	Alianza Petrolera	18	3	5	10	15	30	-15	14
18	Quindío	18	1	5	12	5	25	-20	8