Simulação de Estratégias de Apostas Desportivas com Cobertura de Perdas

por

Carlos Alexandre Silva Barros

Tese de Mestrado em Análise de Dados e Sistemas de Apoio à Decisão

Orientada por:

Prof. Rui Manuel Santos Rodrigues Leite

Universidade do Porto

Faculdade de Economia

2012

Nota Biográfica

Carlos Alexandre Silva Barros, natural do Porto, iniciou o seu percurso académico na Universidade do Minho onde obteve a licenciatura no curso de Gestão com um ano realizado na *University of Luton* no Reino Unido ao abrigo do Programa Erasmus. Em 2010 inscreve-se no curso de Análise de Dados e Sistema de Apoio à Decisão na Faculdade de Economia do Porto com o objectivo de obter um conjunto de ferramentas que permitissem estudar o tema que esta tese aborda.

Resumo

O mercado das apostas *online* é um sector que gera receitas na ordem dos 24.6 mil milhões de euros com o mercado das apostas desportivas a representar 43% desse valor. Neste contexto não é de estranhar que muitos jogadores tentem criar sistemas de apostas que lhes permitam vencer a casa de apostas e com isso conseguir um ganho continuo. Uma classe de sistemas de apostas baseia-se na previsão de resultados desportivos. Outros porém, designados sistemas de cobertura de perdas, baseiam-se na alteração dos montantes de apostas no sentido de obter ganhos finais. Este trabalho insere-se neste tipo de abordagem.

Um conhecido sistema de cobertura de perdas é o método Martingale que é define uma estratégia para um jogo repetido em que a probabilidade de ganhar em cada ronda é 50%. No entanto no contexto das apostas desportivas este sistema necessita de ser adaptado. Nesta tese apresentamos uma versão adaptada do método Martingale para apostas desportivas.

Através de simulações computacionais suportadas por um histórico de resultados desportivos (futebol) verificamos que o método adaptado não consegue obter ganhos globais positivos se não usarmos um critério de escolha de apostas. Realizámos uma análise do método segundo uma divisão das apostas em subconjuntos e verificámos que para alguns desses subconjuntos o ganho foi positivo.

Finalmente desenvolvemos uma metodologia que permitisse a um jogador usar o método adaptado num contexto realista. Experiencias preliminares indiciam que pode ser eficaz para obter ganhos reais, isto é, “conseguir vencer as casas de apostas”.

Abstract

The sports gambling market is a sector that generates revenues of 24.6 billion euros with sports betting market accounting for 43% of all online revenue. In this context it is not surprising that many players, try to create betting systems that allow them to beat the bookmaker and achieve a continuous gain. One class of betting systems is based on the prediction of sports results. On the other hand, systems designed to cover losses, are based on the alteration of the amounts of the bet in order to obtain final gains. This work fits into this approach.

A well-known betting system is the Martingale method. Is defined as a strategy for a repeated game in which the probability of winning each round is 50%. However in the context of sports betting, this system needs to be adapted. In this thesis we present an adapted version of the Martingale method for sports betting.

Through computer simulations supported by a record of sports results (football) we found that the adapted method can´t get a positive profit if we do not use a criterion for bet selection. We conducted an analysis of the method according to the betting partition into subsets and we found that for some of these subsets the gain was positive.

Finally we developed a methodology that allows the player to use the method, in a realistic context. Preliminary experiments indicate that it can be effective for real gains, ie "can beat the bookmakers."

Conteúdo

Resumo ii
Abstract iii
Índice de Tabelas v
Índice de Figuras v
1. Introdução 1
1.1. Estrutura da Tese 1
1.2. Alguns Conceitos sobre Apostas Desportivas 2
1.3. Métodos de Cobertura de Perdas 6
1.4. Método Martingale 7
1.5. Trabalhos Relacionados 8
1.6. Conceitos 8
2. Métodos Desenvolvidos 13
2.1. Método Martingale Adaptado 13
2.2. Metodologia Utilizada para Estudar a Simulação 16
3. Experiências 20
3.1. Escolha dos Dados que vamos Utilizar 20
3.2. Liga Seleccionada para as Experiências 20
3.3. Dimensão da Aposta Múltipla 21
3.4. Avaliação de Resultados 21
3.5. Simulação em Todo Conjunto de Dados 22
3.6. Simulação em Subconjuntos de Apostas 23
3.7. Simulação com Arranque na Jornada k 31
4. Conclusões 35

Índice de Tabelas

Tabela 1 – Simulação de resultados reais desportivos 3
Tabela 2 - Simulação do método Martingale original – uma sequência de perdas 7
Tabela 3 - Método Martingale original – Sequência com perdas e ganhos..... 8
Tabela 4 - Método Martingale Adaptado. Valor da Aposta 1€. 13
Tabela 5 - Método Martingale Adaptado - Sem saldo para realizar a aposta..... 15
Tabela 6 - Classificação dos clubes segundo a análise de agrupamentos..... 19
Tabela 7 – Principais ligas de futebol europeias e respectivo número de clubes. 20
Tabela 8 -Simulação em Todo Conjunto de Dados – Resultados Obtidos 22
Tabela 9 -Forma como se distribuem os ganhos..... 23
Tabela 10 - Tabela de partida para análise de agrupamentos..... 24
Tabela 11 - Vitórias Reais *vs* Esperadas 25
Tabela 12 - Clubes com resultado não esperado 26
Tabela 13 – Caracterização dos agrupamentos 27
Tabela 14 - Agrupamentos 27
Tabela 15 - Atribuição do grupo aos elementos da aposta recursiva 28
Tabela 16 - Universo de apostas recursivas pretendidas 28
Tabela 17 - Padrões com suporte superior a 10%. 29
Tabela 18 - Melhor resultado global para cada valor de aposta inicial 29
Tabela 19 - Resultado de cada múltipla com frequência superior a 10% 30
Tabela 20 - Tabela de partida para análise de agrupamentos com 6 jornadas 32
Tabela 21 - Caracterização dos agrupamentos 33
Tabela 22 - Escolha dos grupos correspondentes entre épocas 34
Tabela 23 - Resultados da avaliação 34
Tabela 24 - Ganhos de apostas segundo padrões obtidos usando a informação total 34

Índice de Figuras

Figura 1 - Exemplo de uma aposta simples. Fonte: www.bwin.com 3
Figura 2 – Odd´s de uma jornada. Fonte: www.bwin.com 4
Figura 3 - Odd com 50% de probabilidade. Fonte: www.bwin.com..... 5
Figura 4 - Dendrograma das diferenças 25
Figura 5 - Dendrograma com variável odd – 4 grupos..... 26
Figura 6 - Dendrograma com variável odd – 5 grupos..... 32

1. Introdução

A indústria das apostas *online* abarca vários jogos – lotarias, bingo, poker, casino e apostas desportivas. Esta indústria é um sector que gera receitas na ordem dos 24.6 mil milhões de euros, com a área das apostas desportivas a ser responsável por uma quota de 43% (H2 Gambling Capital, 2011). Neste contexto não é de estranhar que muitos jogadores tentem criar sistemas de apostas que lhes permitam vencer a casa de apostas e com isso conseguir um ganho continuo. Uma classe de sistemas de apostas baseia-se na previsão de resultados desportivos. Outros porém baseiam-se na alteração dos montantes de apostas no sentido de recuperar as perdas. Neste trabalho desenvolvemos algoritmos que permitem simular um destes sistemas e avaliar o seu desempenho.

O objectivo deste trabalho é responder à questão "É possível ganhar à casa de apostas?". Nesta perspectiva é abandonada a ideia de se o clube *a* ou *b* vai ganhar ou perder o próximo evento para uma perspectiva de longo prazo, em que um determinado universo de apostas apresenta um determinado ganho.

Concretamente:

- Pretendemos desenvolver um método de apostas desportivas que permita cobrir as perdas e gerar ganhos.
- Utilizar simulação baseado em resultados históricos para testar o método desenvolvido e averiguar situações em que o método possa funcionar.

1.1. Estrutura da Tese

O presente trabalho segue uma sequência que culmina no Capítulo 5 com as conclusões finais. No capítulo 2 apresenta-se a pertinência da área das apostas desportivas, conceitos inerentes às apostas desportivas, métodos de cobertura de perdas e trabalhos relacionados com o problema que estudámos. No Capítulo 3 apresentamos o método que propomos e a metodologia que utilizámos para correr as simulações. No Capítulo 4 apresentamos as experiências realizadas para verificar o desempenho dos métodos desenvolvidos. Em anexo colocámos informação que por motivo de facilitar a leitura optámos por retirar ou apresentar em resumo no texto principal. É o caso de algumas tabelas extensas e código fonte da implementação dos métodos.

Apostas Desportivas e Sistemas de Cobertura de perdas

As apostas em eventos desportivos constituem um mercado economicamente importante na Europa, sendo o Reino Unido o mercado mais importante. Inicialmente incidindo apenas nas corridas de cavalos, foi alargada ao futebol no início do séc. XX mas apenas comercializadas *over-the-counter* (OTC) – ao balcão (Gomber et al., 2008).

De acordo com Danduran (1999) a internet deu às casas de jogo tradicional, onde se incluem as apostas desportivas, a oportunidade de transformar e expandir as suas ofertas. Desde a primeira casa de jogo *online* em 1995, em 2006 já eram aproximadamente 2.500 sítios de jogo *online* com uma taxa de crescimento anual de 20% (Stewart 2006).

1.2. Alguns Conceitos sobre Apostas Desportivas

As apostas desportivas envolvem 2 tipo de agentes: a casa de aposta e o jogador. As casas de apostas aceitam apostas sobre resultados de vários eventos desportivos que são na sua natureza imprevisíveis. Em função da escolha do jogador oferecem um ganho, denominado multiplicador de ganho (*odd*), que é tanto maior quanto menos provável for esse o resultado. O jogador pode escolher o montante que quer apostar e dessa forma multiplicar o seu ganho no caso de acerto. Nas apostas *online* o jogador carrega um montante de saldo que poderá depois utilizar nas sucessivas apostas. A seguir detalhamos alguns conceitos relacionados com as apostas desportivas.

Tipos de Apostas

O jogador tem à sua disposição 2 tipos de apostas: Apostas simples e apostas múltiplas. Uma aposta simples incide sobre o resultado de um evento desportivo (por exemplo jogo de futebol). Basta acertar nesse resultado para a aposta ser bem sucedida e recolher o ganho. O ganho resulta da multiplicação do valor do multiplicador de ganhos pelo montante apostado.

A figura seguinte mostra um exemplo de aposta simples e respectivo ganho potencial (12.80-10=2.80 EUR). Neste caso a aposta é feita no clube que o jogador escolheu como vencedor.



Figura 1 - Exemplo de uma aposta simples. Fonte: www.bwin.com

Uma aposta múltipla é uma aposta efectuada no conjunto de vários eventos podendo o resultado escolhido ser o mesmo ou diferente para cada um dos eventos. A aposta é vencedora se se acertar em todos os resultados. Por exemplo o jogador pode apostar que no jogo 1 o resultado é vitória para equipa A e que o jogo 2 tem como resultado empate, só recolhendo ganho se os dois resultados se verificarem.

A seguinte tabela demonstra vários tipos de apostas múltiplas (o que o jogador escolhe), os resultados desportivos reais (o que aconteceu nos jogos) e o resultado da aposta (se o jogador ganha ou perde).

Tabela 1 – Simulação de resultados reais desportivos

Jornada	Aposta	Resultado Efectivo	Resultado da Aposta
1	{V,V,V,V}	{V,V,V,V}	Sucesso
2	{V,V,V,V}	{V,E,D,V}	Fracasso
3	{V,V,V,V}	{V,V,D,V}	Fracasso
4	{V,V,D,E}	{V,V,D,E}	Sucesso

O valor do multiplicador de ganhos, da aposta múltipla que é indicado pela casa de apostas, resulta do produto dos multiplicadores de ganho de toda as apostas simples.

Multiplicador de Ganhos

O multiplicador de ganhos, habitualmente designado por *Odd*, é um elemento fundamental em todo este processo. O multiplicador de ganhos ou *Odd* reflecte a probabilidade do resultado. Desta forma o valor do multiplicador de ganhos é inversamente proporcional à probabilidade de sucesso. Isto significa que um multiplicador de valor de 2,00 apresenta uma probabilidade de sucesso 50%. Quanto mais alto for o valor do multiplicador maior o risco, mas também maior o retorno. A figura seguinte mostra os multiplicadores oferecidos por uma casa de apostas. Como vamos explicar a seguir os multiplicadores oferecidos são mais baixos do que o deveriam ser, pois incluem a margem de lucro da casa de apostas.

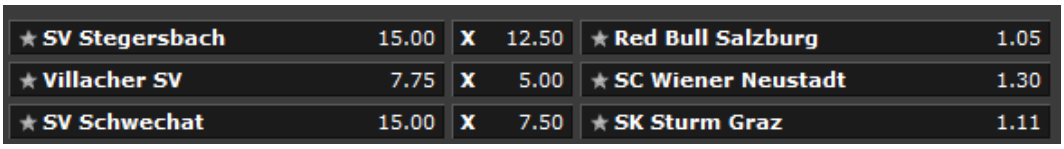


Figura 2 – Odd´s de uma jornada. Fonte: www.bwin.com

Margem de Lucro e relação com Odd

Outra importante característica do multiplicador de ganhos é que em teoria, caso o valor esteja correctamente calculado, em cada aposta da série o valor esperado do ganho é nulo como provamos a seguir:

$$Valor\ Esperado\ do\ Ganho = p * A\ (Odd - 1) + (1 - p)(-A),$$

$$em\ que$$

$$p\ \text{é}\ a\ probabilidade\ de\ a\ aposta\ ter\ sucesso\ e\ A\ o\ valor\ apostado$$

$$Se\ Odd\ for\ denifida\ \frac{1}{p}\ o\ valor\ esperado\ fica$$

$$p * A\left(\frac{1}{p} - 1\right) + (1 - p)(-A) = A - pA - A + pA = 0$$

Contudo atentemos à figura seguinte relativa ao mercado par ou impar de cantos numa partida de futebol, um evento com probabilidade de sucesso de 50%.



Figura 3 - Odd com 50% de probabilidade. Fonte: www.bwin.com

Apesar da probabilidade de sucesso ser de 50% o multiplicador oferecido tem apenas o valor de 1,85. Desta forma o valor esperado dos ganhos em cada uma das apostas não tende para 0, mas sim para um valor negativo (uma vantagem para a casa de apostas).

$$p * A\ (Odd - 1) + (1 - p)(-A) =$$

$$0{,}5 * A\ (1{,}85 - 1) + (1 - 0{,}5)(-A) =$$

$$(0{,}5 * 0{,}85 - 0.5)A \quad < \ 0, \quad pois\ A > 0$$

A diferença entre o multiplicador teórico e o multiplicador oferecido pela casa de apostas exprime a margem de lucro. Temos assim que a margem de lucro desta casa de apostas é de 7,50%.

$$Margem = 2 - 1{,}85 = 0{,}15$$

$$Margem\ (\%) = \frac{0{,}15}{2} = 7{,}5\%$$

1.3. Métodos de Cobertura de Perdas

Considerando um jogo de moeda ao ar repetido utilizando uma moeda não viciada é habitual distribuir uma recompensa no caso de ganhar igual ao prejuízo no caso de perder. Se por exemplo o jogador apostar X em que sai coroa o seu saldo aumenta X se sair coroa. Inversamente o saldo diminui o mesmo valor se sair cara.

Embora o jogo tenha um ganho esperado nulo, na prática isso não acontece, porque é possível ter um número de ganhos diferente do número de perdas. O jogador não pretende ter no fim do jogo um prejuízo nem tão pouco um ganho nulo. Se numa ronda ele perder vai adaptar a sua estratégia com o objectivo de obter um ganho que cubra o total das perdas e obter um ganho liquido. Os métodos de cobertura de perdas são usados neste contexto. Tipicamente, isto é alcançado aumentando o valor da aposta. A dinâmica do aumento da aposta, habitualmente designado por multiplicador de ganhos, define o tipo de método de cobertura de perdas. Existem vários métodos de cobertura de perdas: *Martingale, Fibonacci, Labouchere, Oscar, d'Alembert e Blundell* (Ethier, 2010).

Habitualmente apresentados como sistemas infalíveis os métodos de cobertura de perdas apresentam as seguintes fraquezas:

- O jogador não pode apostar o que não tem (note que no Martingale original não existe o conceito de saldo).

- Na prática existe um limite máximo para o valor da aposta a efectuar.

Nesta tese estudamos o método Martingale visto ser o mais amplamente divulgado. Estudamos o método Martingale na sua forma original bem como desenvolvemos adaptações para o aplicar a um cenário mais complexo – as apostas desportivas. Na próxima secção apresentamos em detalhe o método Martingale original.

1.4. Método Martingale

Neste sistema, na sua forma original, o jogador duplica a sua aposta após cada fracasso e pára depois do primeiro sucesso. O jogo pode continuar, voltando o montante da aposta ao seu valor inicial.

Se considerarmos que A_i representa genericamente o valor da aposta na ronda i, e N o indice da primeira ronda em que o jogador ganha, A_i pode ser definido por:

$$A_i = 2^{i-1}X \quad (i \leq N) \quad , \qquad em\ que\ X\ \text{é}\ o\ valor\ da\ Aposta\ inicial\ A_1$$

Ou definindo em função do valor da aposta anterior

$$A_i = 2A_{i-1} \quad (1 < i \leq N)$$

A razão para a duplicação da aposta é motivada pelo facto da probabilidade de sucesso em cada ronda ser ½. Se pretendermos adaptar este método num caso em a probabilidade de sucesso seja diferente, a dinâmica de modificação do valor da aposta terá que ser também adaptado.

A próxima tabela ilustra o funcionamento do método de Martingale com multiplicador de ganhos (*Odd*) igual a 2.

Tabela 2 - Simulação do método Martingale original – uma sequência de perdas

Ronda	Aposta (A_i)	Resultado	Lucro
1	1	Perdeu	-1
2	2	Perdeu	-3
3	4	Perdeu	-7
4	8	Perdeu	-15
5	16	Perdeu	-31
6	32	Ganhou	1

No exemplo apresentado o valor inicial da aposta, A_1, é 1€. O jogador perde nas 5 primeiras rondas e na sexta recupera as perdas $(-31€)$ e ganha o valor da aposta $(1€)$.

Num jogo real a aplicação do método Martingale é feita apenas em sequências de perdas consecutivas, voltando o jogador a apostar o valor inicial após a ronda ganhadora. A seguinte tabela apresenta um exemplo de jogo que envolve estas duas fases.

Tabela 3 - Método Martingale original – Sequência com perdas e ganhos

Ronda	Aposta (A_i)	Resultado	Lucro
1	1	Perdeu	-1
2	2	Perdeu	-3
3	4	Ganhou	1
4	1	Perdeu	-1
5	1	Ganhou	1
6	1	Ganhou	2

1.5. Trabalhos Relacionados

Não conseguimos encontrar trabalhos relacionados com métodos de cobertura de perdas aplicados a apostas desportivas. Há contudo muitos trabalhos (Goddard 2004), (Boulier 2003), (Harville 1980), (Stefani 1980) na área de apostas desportivas mas com o foco na previsão de resultados desportivos. Noutra vertente deste contexto recentemente Moya (2012) publicou um trabalho sobre sistemas de apostas em que aborda alguns métodos de cobertura de perdas, como o *D'Alembert*, mas como indicação para trabalhos futuros.

1.6. Conceitos

Nesta secção são apresentados alguns conceitos que são importantes para a compreensão dos métodos apresentados no próximo capítulo.

Simulação

A simulação refere-se a um conjunto de processos e aplicações que replicam o comportamento de sistemas reais. Na verdade o conceito de simulação pode ser um conceito muito abrangente podendo incluir diversas áreas. Actualmente devido ao desenvolvimento dos computadores a simulação tornou-se uma ferramenta poderosa e largamente aceite.

A simulação de processos é uma ferramenta eficaz no auxílio à tomada de decisões. Ela permite avaliar o desempenho de um sistema, melhorar operações ou desenhar um sistema caso ele ainda não exista (Kelton, 2002).

A simulação pode ser feita através de mudanças no próprio sistema real. As vantagens são muitas, contudo em muitos casos não é de todo possível. Ou pelos custos que representa, ou porque o modelo físico não existe, ou até mesmo porque pode acarretar perigo. Desta forma a solução passar pelo recurso a um modelo. O modelo pode ser um modelo físico, também designado por modelo icónico, modelos lógicos ou matemáticos, ou simulação em computador (Kelton, 2002).

O objectivo proposto enquadra-se no último caso. Pretende-se criar um modelo computacional de simulação de apostas desportivas que através de dados reais incorpore o âmago de todo o processo – Risco e Rendibilidade.

Risco e Rendibilidade

Por risco entende-se uma situação futura de resultado desconhecido. O risco distingue-se da incerteza por ser uma situação à qual se pode atribuir probabilidade (Matias, 2005).

De acordo com Neves (2002) a análise do risco desempenha dois papeis fundamentais:

- Fornece uma perspectiva mais abrangente para estruturar o processo de tomada de decisão;

- Fornece um conjunto de técnicas para avaliar o interesse das alternativas em jogo com vista a uma decisão final.

Da relação entre risco e rendibilidade, podemos retirar alguns princípios: (Matias 2005, pag.152).

- Quanto maior a incerteza maior remuneração é exigida pelo investidor.

- Activos com níveis de risco idênticos devem ter a mesma rendibilidade esperada. Se tal não se verificar existirá uma clara tendência do mercado para corrigir essa situação.

No âmbito das apostas desportivas, é nesta dicotomia risco e rendibilidade que todo o processo de formação de preços se forma. Ou seja o apostador, para um evento com resultado de elevada incerteza, vai exigir uma rendibilidade também elevada. Neste caso o risco é traduzido na probabilidade de vitória de um determinado clube. A rendabilidade por sua vez é o inverso do risco e assume a forma de multiplicador de ganhos (*Odd*).

Abordada a relação entre risco e rendibilidade importa conhecer o instrumento que permite identificar quais os clubes com risco e rendibilidade semelhantes. Para tal usamos a análise de agrupamentos (*clustering hierárquico*).

Análise de Agrupamentos

A análise de agrupamentos tem como objectivo encontrar semelhanças entre objectos e por conseguinte agrupar esses mesmos objectos em grupos o mais homogéneos possíveis e ao mesmo tempo o mais diferentes entre si.

Encontrados esses grupos podem-lhes ser atribuídos nomes e os seus atributos definidos. Uma das possibilidades para definir os atributos de um grupo é através do valor médio dos seus objectos ou através de um intervalo de valor. Este procedimento permite uma descrição simples dos dados (Alpaydin, 2004).

Caso a análise de agrupamentos tenha por objectivo definir sobre um conjunto de dados, comparáveis, partições ou hierarquias de partições que respeitem o melhor possível as semelhanças entre esses objectos sem qualquer outro requisito, então estamos perante uma análise de agrupamentos hierárquica.

Os agrupamentos hierárquicos são representados por uma árvore de classificação ou dendrograma. Um dendrograma é uma árvore onde as folhas correspondem aos objectos que aparecem segundo a ordem como foram sendo agrupados.

Esta abordagem implica a utilização de uma similaridade, ou um modo de medida equivalente a uma distância, definida entre os objectos. Geralmente é usada a distância Euclidiana na abordagem de agrupamentos hierárquica.

Existem dois tipos de métodos de agrupamentos hierárquicos: O método aglomerativo ou ascendentes e o método divisível ou descendente (Alpaydin, 2004).

O método aglomerativo começa com N grupos, cada um contendo um elemento que junta sucessivamente a elementos similares criando grupos maiores até obter apenas um grupo (Alpaydin, 2004).

O método divisível começa com um grupo único e divide-o até obter grupos que contenham um objecto único (Alpaydin, 2004).

Ao longo do processo interactivo de formação de grupos é necessário calcular a distância entre as classes que se vão formando. Os métodos mais comuns são o método de ligação simples e o método de ligação completa.

O método de ligação simples define distância como a menor distância entre todos os pares de elementos possíveis. No método de ligação completo, a distância é definida como a maior distância entre todos os pares de elementos possíveis (Alpaydin, 2004).

Outra possibilidade de formação e grupos é através do método *Ward*. De uma maneira geral este método é bastante eficiente. Este método difere dos anteriores porque usa uma abordagem baseada na variância para calcular a distância entre os agrupamentos. Cada objecto é avaliado através do cálculo da variância. É alocado ao agrupamento o objecto que minimiza a variação da inércia interclasse (Burns et al., 2008).

No âmbito desta tese a técnica que foi usada como ponto de partida foi a análise de agrupamentos com distâncias euclidianas e formação de grupos através do método *Ward*.

Padrões Frequentes

A procura de padrões frequentes é uma das técnica que iremos utilizar para analisar o método Martingale. O objectivo é identificar os padrões que ocorrem mais vezes. Nesta dissertação foi utilizado o pacote do R *arules* (Hahsler et al,2008) que tem uma implementação do algoritmo *apriori (*Agrawal 1994).

Um dos parâmetros necessário no método *apriori* é o *Suporte mínimo* (%), que condiciona a selecção aos padrões que ocorrem pelo menos nessa percentagem.

2. Métodos Desenvolvidos

2.1. Método Martingale Adaptado

Nesta adaptação do Método Martingale Original o valor do multiplicador passa a ser dinâmico em vez de constante e igual a $2{,}00$. Ou seja em cada jornada o valor da aposta tem em conta o valor do multiplicador dessa mesma jornada como presente na tabela seguinte.

Tabela 4 - Método Martingale Adaptado. Valor da Aposta 1€.

Ronda	Aposta	Odd	Resultado	Lucro
1	1,00	2,00	Perdeu	-1,00
2	1,50	3,00	Perdeu	-2,50
3	3,50	2,00	Perdeu	-6,00
4	4,00	3,00	Perdeu	-10,00
5	3,50	5,00	Perdeu	-13,50
6	5,50	4,00	Ganhou	3,00

Tal como no caso do Método Martingale Original a cobertura de perdas é realizada através do aumento do valor da aposta. Se considerarmos um prejuízo acumulado P que decorre duma série de apostas mal sucedidas, o jogador vai apostar um valor que lhe permita anular o prejuízo e obter um ganho adicional. Esse ganho adicional é o que resultaria por recompensa da *Odd* a uma aposta de valor igual ao inicialmente definido.

A seguinte equação define o valor da aposta utilizado pelo método, num contexto em que existem prejuízos acumulados P.

$$A_i = \frac{P}{Odd - 1} + A_1 \quad , considerando\ que\ P > 0$$

Por exemplo para a linha 4 da tabela onde o valor da aposta é de 4 temos:

$$A_4 = \frac{6}{3-1} + 1$$

Um sistema que simula as apostas usando esta estratégia de cobertura de perdas inclui também as rondas em que as apostas são bem sucedidas. Tal como no Método Martingale Original neste caso o valor da aposta passa para o valor inicialmente definido. Apresentamos a seguir o algoritmo de simulação definido utilizando pseudo-código.

```
Simula_Aposta(apostaEscolhida, saldoInicial, valorApostaInicial, )   {

        saldo = saldoInicial
        prejuizo = 0
        vAposta = valApostaInicial


        Repete
            Se sucesso(apostaEscolhida)
                prejuizo = 0
                saldo = saldo + vAposta *(odd(apostaEscolhida) - 1)
              (a)  vAposta= valApostaInicial
            Senão
                prejuizo = prejuizo + vAposta
                saldo = saldo - vAposta
              (b)  vAposta= prejuizo / (odd(apostaEscolhida) - 1) + valorApostaInicial
            Fim Se

        Até acabarem rondas ou saldo <aposta

}
```

A simulação compreende um ciclo que é executado até que não existam mais jornadas ou que seja insuficiente o saldo para realizar a aposta. Se o resultado desportivo for favorável o valor da aposta é recolocado no seu valor inicial (passo a), se pelo contrário a aposta for mal sucedida o valor da aposta é alterado (passo b). Os valores do prejuízo e saldo são actualizados de acordo com os ganhos ou perdas que decorrem da aposta. A seguinte tabela mostra a aplicação do algoritmo a um conjunto de 4 jornadas em que na

última jornada o apostador não dispõe de saldo suficiente para efectuar a aposta. O saldo inicial foi de 10€ e o valor da aposta de 1€.

Tabela 5 - Método Martingale Adaptado - Sem saldo para realizar a aposta.

Ronda	Odd	Saldo	Aposta	Resultado	Lucro
1	1.70	10	1,00	Ganhou	0,70
2	1.80	10,70	1,00	Perdeu	-1,00
3	1.25	9,70	5,00	Perdeu	-5,00
4	1.12	4,70	51,00	-	-

Aposta Recursiva

Além de múltipla a aposta a usar na simulação a testar é também recursiva. Isto quer dizer que fixamos um grupo de clubes e apostamos sempre na vitória desses. O resultado escolhido será sempre a vitória. No caso de jornadas que contemplem jogos entre clubes escolhidos, essa jornada será eliminada.

Saldo Inicial

À partida é necessário determinar qual o valor do saldo inicial. Ou seja qual o valor que vamos investir no início do campeonato para uma determinada aposta múltipla. Uma vez que as casas de apostas consideram 5.000€ como valor máximo de transferência inicial no momento da abertura de uma conta, considera-se esse valor como referência. O saldo inicial sofre alterações com os ganhos conseguidos durante o campeonato. O valor do saldo inicial tem que ser maior ou igual que o valor da aposta em cada jornada.

Valor da Aposta

O valor da aposta em cada jornada é de crucial importância. Um valor demasiado elevado leva a que o sistema rapidamente colapse, ou seja, que o valor da aposta ultrapasse o saldo disponível numa determinada jornada do campeonato. Um valor demasiado baixo não permite uma rentabilidade interessante. Na nossa abordagem o

valor da aposta tem um valor inicial fixado (é um parâmetro) e é alterado de acordo com o método.

2.2. Metodologia Utilizada para Estudar a Simulação

Nesta secção vamos apresentar as metodologias utilizadas para estudar a simulação. Uma simulação utiliza um histórico constituído pelos resultados reais de determinado evento. A metodologia apresentada incide sobre os jogos de futebol. Estes jogos pertencem a uma liga e época específica e aos quais correspondem os respectivos multiplicadores de ganhos (fixados pela casa de apostas). Por esta razão temos que primeiro escolher a liga e época onde pretendemos correr a simulação.

Em cada aposta temos vários resultados possíveis e o resultado a escolher é sempre o mesmo. Como o objectivo de qualquer equipa é a vitória foi esse o resultado escolhido. Ou seja para todas as equipas constituintes da aposta múltipla o resultado escolhido é a vitória. Dito de outra forma a aposta múltipla é recursiva, ou seja a aposta é feita sempre na mesma múltipla de clubes.

De seguida temos que fixar a dimensão da aposta múltipla. Uma aposta múltipla com muitos clubes leva, reduz bastante o número de jornadas disponíveis. Uma jornada que envolva jogos entre equipas da aposta múltipla considerada têm que ser ignorada pois não é possível às 2 equipas ganharem. Uma múltipla com poucos clubes, ou até mesmo uma aposta simples, conduz a um valor de multiplicador conjunto muito baixo. Por consequência o valor da cobertura de perdas torna-se muito elevado o que leva a que facilmente o valor da aposta exceda o valor do saldo disponível.

Uma abordagem exaustiva tem ganhos positivos?

O objectivo é saber se o método funcionaria se o jogador tivesse recursos para suportar todas as apostas possíveis. Embora tenhamos a suspeita que é uma hipótese condenada ao fracasso pretendemos correr experiências para comprovar esse facto.

Nesta secção apresentamos a metodologia que utilizamos no capítulo seguinte para testar esta hipótese. Esta consiste nos passos indicados a seguir.

```
Enumerar todas apostas possíveis
Para cada vAposta em {V1, V2 , ...}
        Correr a simulação
        para as apostas enumeradas (considerando vAposta)

        Contabilizar total ganhos
                guardar em ganhos[vAposta]
```

No fim o total de ganhos vai revelar se o sistema funciona numa situação em que apostamos em todas as alternativas. Esta análise é feita para vários valores de apostas iniciais considerados.

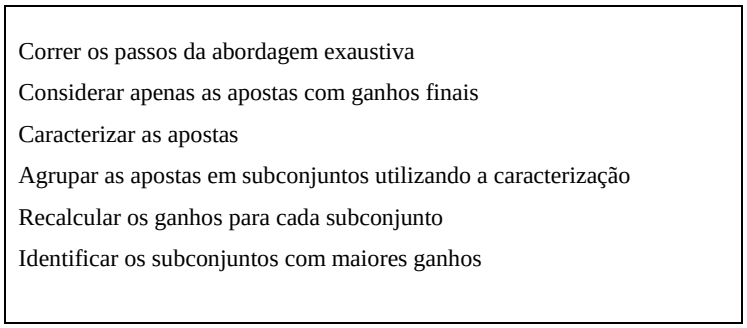
O número de apostas recursivas é determinado pelo numero de clubes N e a dimensão da aposta múltipla n. A tabela seguinte mostra essa relação considerando uma liga constituída por 20 clubes.

Dimensão da aposta múltipla	**Combinações possíveis**
1	20
2	190
3	1140
4	4845
…	…
n	C_n^{20}

Podemos encontrar subconjuntos de apostas com ganho?

Na secção anterior apresentamos a metodologia para testar se é possível ter resultados positivos usando o método Martingale Adaptado num contexto em que as apostas não são escolhidas mas sim exaustivamente consideradas. Se a resposta a esta hipótese for negativa isto não quer dizer que o método não possa ser usado num contexto em que restringimos as nossas apostas a um subconjunto. O subconjunto terá que ser definido por um critério em que sejam consideradas características das apostas. O que pretendemos é definir formas de caracterizar as apostas de modo a identificar subconjuntos em que o método funcione.

Nesta secção apresentamos a metodologia que utilizamos no capítulo seguinte para testar esta hipótese. Esta metodologia consiste nos passos indicados a seguir.



Apresentamos a seguir detalhes sobre estes passos.

Caracterizar as apostas

Para caracterizar as apostas, representadas por apostas recursivas, que como sabemos correspondem a subconjuntos de apostas, temos que primeiro realizar uma análise de agrupamentos de clubes baseada na *Odd* correspondente à vitória. Utilizamos esta variável pois é a que melhor incorpora a relação entre rendibilidade e risco, os dois conceitos fundamentais em todo este processo.

Para uma aposta recursiva $\{clubeA, clubeB, clubeC\}$ a caracterização é $\{X, Y, Z\}$, em que $\{X, Y, Z\}$ são os grupos respectivos dos clubes. Esta caracterização, no entanto, por si só, pouca informação contém. Ou seja um resultado que devolva um grupo constituídos pelos clubes $\{X, Y, Z\}$ nada diz. Além disso a necessidade de utilizar essa informação em épocas futuras não seria possível. Primeiro porque o desempenho dos clubes é dinâmico e segundo porque os clubes de uma época de determinada liga não são os mesmos da época seguinte devido à promoção e despromoção dos mesmos. Assim importa caracterizar os grupos, e por conseguinte as apostas, através de informação mais facilmente comparável.

Agrupar as Apostas

Após realizados os agrupamentos, todas os clubes constituintes da aposta múltipla foram classificados segundo a letra correspondente ao grupo a que pertencem.

Tabela 6 - Classificação dos clubes segundo a análise de agrupamentos

Clube			Grupo		
Arsenal	Aston Villa	Fulham	A	B	C
Arsenal	Bolton	Burnley	A	C	D
Arsenal	Aston Villa	Blackburn	A	B	X
Arsenal	Aston Villa	Tottenham	A	B	X
…	…	…	…	…	…

A simulação foi executada sendo determinados os ganhos totais de cada aposta múltipla. Posteriormente o universo total foi limitado apenas às apostas com maior ganho, sobre o qual foi executado o algoritmo *apriori*. Esta execução permitiu determinar quais os padrões frequentes das apostas geradoras de ganho. Determinados os padrões foi depois calculado para cada um desses padrões o total de ganho. Ou seja qual o ganho total de todas as apostas constituintes desse padrão, quer as geradoras de ganhos quer as não geradoras de ganho.

3. Experiências

3.1. Escolha dos Dados que vamos Utilizar

De forma pública estão disponíveis no sitio http://www.football-data.co.uk os seguintes dados relativos às principais ligas de futebol europeias. As ligas e respectivo número de clubes estão representados na tabela seguinte.

Tabela 7 – Principais ligas de futebol europeias e respectivo número de clubes.

Liga	Número de clubes
Portuguesa	16
Inglesa	20
Espanhola	20
Italiana	20
Alemã	18
Francesa	20
Holandesa	18

3.2. Liga Seleccionada para as Experiências

Das várias ligas candidatas tem mais interesse escolher as ligas com mais clubes pelo facto de que o funcionamento do algoritmo pressupõe que os jogos entre clubes escolhidos sejam eliminados. Desta forma uma liga com poucos clubes fica com poucas jornadas disponíveis para executar a simulação. A análise será efectuada apenas com os dados da Liga Inglesa – *Premier League* – época 2009-2010.

Estes dados relativos à liga Inglesa tiveram de ser tratados de modo a identificar a que jornada pertence determinado jogo, uma vez que os dados surgem ordenados cronologicamente. Nesta liga, devido a impeditivos de ordem climatérica, é comum as partidas serem realizadas em data posterior à prevista. Por este motivo a data de realização nem sempre é idêntica à data em que decorreram os restantes jogos dessa jornada. Foram ainda criados vários scripts em R (R Development Core Team, 2012) de modo a extrair conhecimento dos dados.

3.3. Dimensão da Aposta Múltipla

Como já foi referido uma aposta múltipla é uma aposta efectuada no conjunto de vários eventos, neste caso jogos de futebol, sendo o resultado escolhido a vitória. Importa agora saber qual a dimensão da aposta múltipla a utilizar. Além do aumento do risco inerente a uma aposta múltipla com muitos clubes, a própria execução da simulação que elimina os jogos entre os clubes escolhidos e por conseguinte essa jornada, reduziria bastante as jornadas disponíveis. Por outro lado uma aposta múltipla com poucos clubes, ou até mesmo uma aposta simples, apesar de conduzir a uma diminuição do risco, aumentaria o valor da cobertura de perdas. Este facto levaria a que facilmente o valor da aposta excedesse o valor do saldo disponível. Tendo em conta estes 2 aspectos optamos por uma aposta múltipla constituída por 3 clubes

3.4. Avaliação de Resultados

Os resultados foram avaliados no fim do campeonato. Para o conjunto de dados utilizado foi calculado a soma dos ganhos (eventualmente negativos) de todas as apostas múltiplas pertencentes a esse mesmo conjunto de dados. Para os casos em que o valor da aposta excede o saldo disponível foi considerado uma perda igual ao valor total do saldo inicial.

3.5. Simulação em Todo Conjunto de Dados

A simulação foi executada com todos os clubes da liga e época escolhida. O algoritmo foi executado em todas as combinações possíveis de 3 clubes, originado 1140 execuções. Forma simulados vários valores de aposta sempre com saldo inicial de 5.000€. Note-se que o saldo inicial é colocado em cada aposta múltipla. Os "Sucessos" são o total de apostas múltiplas cujo resultado se traduz num ganho.

A análise dos resultados da simulação permitiu verificar que para todos os valores de aposta inicial considerados, o resultado é negativo, confirmando as nossas expectativas - se assim não fosse a casa de apostas facilmente perderia.

A seguinte tabela mostra os resultados obtidos considerando apostas iniciais de 200€, 400€, 600€, 800€. A linha Σ Ganho apresenta o total dos resultados líquidos. A linha #Sucessos apresenta o número de apostas recursivas com resultado liquido positivo no final do campeonato. O Ganho Médio resulta da divisão do Σ Ganho pelo número total de apostas recursivas, neste caso $1140 = C_3^{20}$. Podemos observar que o aumento do valor inicial de aposta implica uma redução do número de sucessos que se explica pela maior facilidade em esgotar o saldo. No entanto esse facto não afecta muito Σ Ganho, pois nos casos de sucesso a recompensa é maior.

Tabela 8 -Simulação em Todo Conjunto de Dados – Resultados Obtidos

	V. Aposta			
	200€	**400€**	**600€**	**800€**
Σ Ganho	-2.695.028€	-2.656.683€	-2.346.946€	-2.536.386€
Universo	1140	1140	1140	1140
Ganho Médio	-2.364 €	-2.330 €	-2.059 €	-2.225 €
# Sucessos	225	124	94	72

A forma como se distribuem os ganhos das apostas recursivas é apresentada na seguinte tabela.

Tabela 9 -Forma como se distribuem os ganhos

	Valor de Aposta			
	200€	**400€**	**600€**	**800€**
Min.	-5.000€	-5.000€	-5.000€	-5.000€
1st Qu.	-5.000€	-5.000€	-5.000€	-5.000€
Median	-5.000€	-5.000€	-5.000€	-5.000€
Mean	-2364€	-2.330€	-2.059€	-2.225€
3rd Qu.	-4.006€	-5.000€	-5.000€	-5.000€
Max.	46.360€	92.720€	135.600€	164.600€

É de notar que para mais de metade das apostas recursivas, considerando todos os valores da aposta inicial, o jogador perde todo o saldo inicial (5.000€). Evidentemente as apostas recursivas a que correspondem maiores ganhos são aquelas que envolvem um valor inicial de aposta maior.

Vamos de seguida apresentar uma análise da simulação em que estudamos os ganhos em subconjuntos de apostas.

3.6. Simulação em Subconjuntos de Apostas

Como visto na secção anterior a execução do algoritmo em todo o conjunto de dados resulta em perda, para qualquer valor de aposta testado. Vamos agora analisar a simulação agrupando as apostas em vários subconjuntos. Para criar esses subconjuntos de dados a técnica utilizada foi a *análise de agrupamentos*.

Criação de Agrupamentos

Para executar a criação dos agrupamentos, várias são as variáveis candidatas: número de pontos, número de vitórias, número de empates, número de derrotas, valor do multiplicador. Estas estatísticas são apresentadas na tabela seguinte.

Tabela 10 - Tabela de partida para análise de agrupamentos

Team	Points	Win	Draw	Loss	Goal Dif	Odd
Chelsea	86	27	5	6	71	1,478
Man United	85	27	4	7	58	1,587
Arsenal	75	23	6	9	42	1,751
Tottenham	70	21	7	10	26	2,397
Man City	67	18	13	7	28	2,237
Aston Villa	64	17	13	8	13	2,719
Liverpool	63	18	9	11	26	1,915
Everton	61	16	13	9	11	3,035
Birmingham	50	13	11	14	-9	4,371
Blackburn	50	13	11	14	-14	4,300
Stoke	47	11	14	13	-14	4,867
Fulham	46	12	10	16	-7	4,055
Sunderland	44	11	11	16	-8	4,225
Bolton	39	10	9	19	-25	4,777
Wolves	38	9	11	18	-24	5,388
Wigan	36	9	9	20	-42	4,772
West Ham	35	8	11	19	-19	4,429
Burnley	30	8	6	24	-40	5,639
Hull	30	6	12	20	-41	5,591
Portsmouth	28	7	7	24	-32	5,909

Pela observação da tabela verificamos que alguns clubes têm um valor de multiplicador demasiado alto para o número de vitórias conseguidas. O *West Ham* é um claro exemplo. Com apenas 8 vitórias tem um valor de multiplicador que o situa nas 13 vitórias. No sentido inverso o Tottenham é um bom exemplo. Com 21 vitórias tem um valor de multiplicador superior ao do Man City, que apenas conseguiu 18 vitórias.

Relembramos que este valor do multiplicador é apenas referente à vitória, ou seja, exclui empate e derrota. É importante que estas *surpresas positivas e negativas*, sejam ignoradas na criação da aposta. Caso contrário qualquer múltipla que inclua uma "surpresa positiva", ou seja um clube com um valor do multiplicador acima do número de vitórias esperadas, vai desde logo resultar num resultado positivo. Ou dito de outra forma são clubes que, ao violarem o princípio da rendibilidade *vs* risco, têm uma rendibilidade superior ao risco atribuído ou o inverso.

Por este motivo executamos uma análise de agrupamentos para verificar quais os clubes que se enquadram nesta descrição. A tabela seguinte indica a diferença entre o número de vitórias conseguidas e o número de vitórias esperadas que se obtém dividindo o número de jogos (38) pelo valor médio do multiplicador de ganhos. Por exemplo para o caso do Chelsea o número de vitórias esperadas é de 38/1,478 = 25,7

Tabela 11 - Vitórias Reais *vs* Esperadas

Equipa	V. Conseguidas	Odd Media	V. Esperadas	Diferença
Chelsea	27	1,478	26	1
Man United	27	1,587	24	3
Arsenal	23	1,751	22	1
Tottenham	21	2,397	16	5
…	…	…	…	…

Efectuando uma análise de agrupamentos sobre a variável *diferença*, obtemos o seguinte dendrograma:

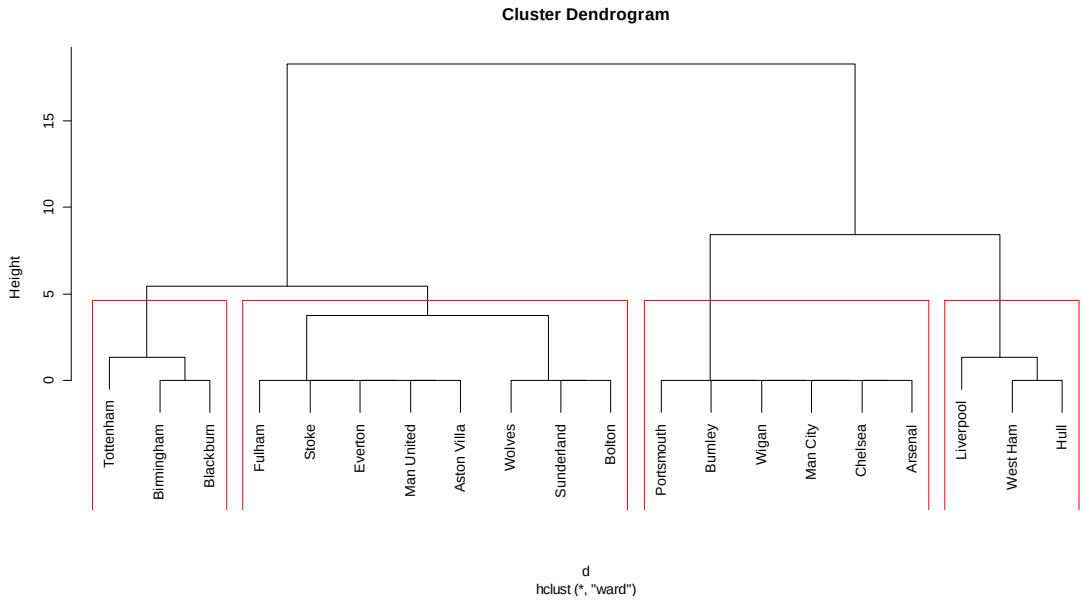


Figura 4 - Dendrograma das diferenças

O dendrograma permite-nos identificar as equipas que se revelaram ser surpresas. Esses clubes estão representados na tabela seguinte.

Tabela 12 - Clubes com resultado não esperado

Clubes com mais vitórias que o previsto	Clubes com menos vitórias que o previsto
Tottenham – [+5]	Liverpool – [-2]
Birmingham – [+4]	West Ham – [-1]
Blackburn – [+4]	Hull – [-1]

Sobre os novos dados realizamos uma análise hierárquica através do método *ward* usando distâncias *euclidianas*. O número de agrupamentos não foi fixado previamente. Contudo a intuição indica que se situem entre os três e os cinco agrupamentos. Mais que cinco conduz a grupos unitários. Menos que três agrupamentos traduzem pouca heterogeneidade entre grupos. A variável escolhida para realizar a análise de agrupamentos foi o valor do multiplicador de ganhos – *odd* – uma vez que este representa a relação entre risco e rendibilidade. A análise do dendrograma da Figura 5 permite distinguir 4 grupos. Importa agora caracterizar esses grupos.

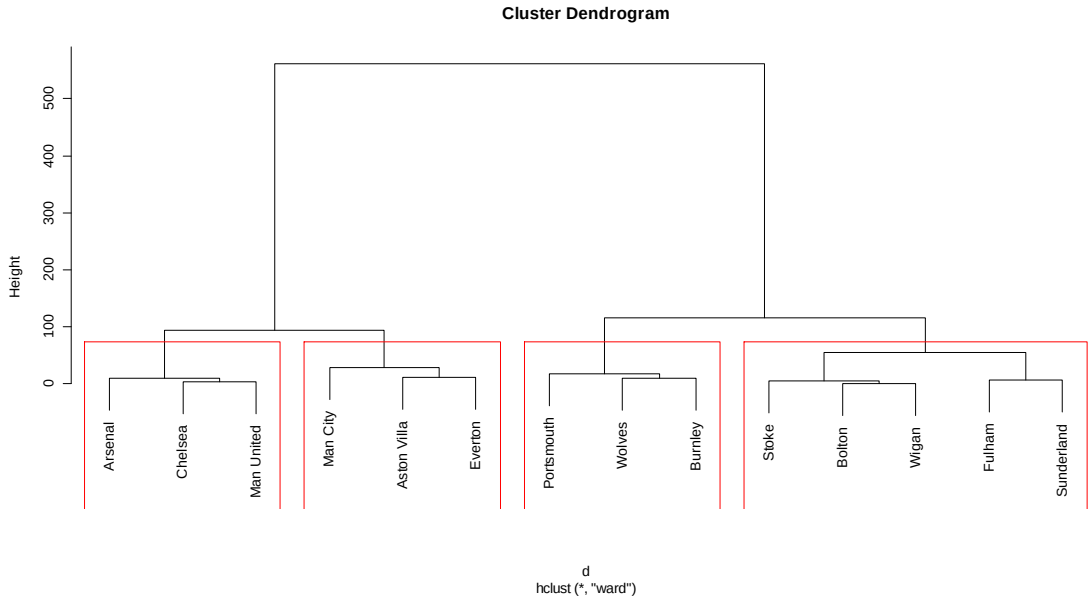


Figura 5 - Dendrograma com variável odd – 4 grupos

Caracterização dos Agrupamentos

Após a análise do dendrograma da Figura 5 efectuamos a respectiva caracterização dos grupos:

Tabela 13 – Caracterização dos agrupamentos

Grupo A	**Grupo B**	**Grupo C**	**Grupo D**
Arsenal	Man City	Stoke	Burnley
Chelsea	Aston Villa	Bolton	Portsmouth
Man United	Everton	Wigan	Wolves
		Fulham	
		Sunderland	
Intervalo do Multiplicador			
[1,48;1,75]	[2,24;3,04]	[4,06;4,87]	[5,39;5,91]
Multiplicador Médio			
1,61	2,67	4,54	5,65

Cada grupo fica identificado pelo valor do multiplicador médio. O valor do multiplicado médio permite calcular o valor do risco de cada subconjunto de dados.

Padrões Frequentes

Como vimos em cima cada agrupamento é classificado por uma letra, a qual corresponde um clube.

Tabela 14 - Agrupamentos

Grupo A	**Grupo B**	**Grupo C**	**Grupo D**
Arsenal	Man City	Stoke	Burnley
Chelsea	Aston Villa	Bolton	Portsmouth
Man United	Everton	Wigan	Wolves
		Fulham	
		Sunderland	

Assim, após a realização da execução exaustiva de todas as combinações possíveis de múltiplas compostas por 3 clubes, foi atribuída a letra correspondente do agrupamento ao qual pertence o clube. O resultado foi a tabela seguinte que por razões óbvias apresentamos de forma abreviada:

Tabela 15 - Atribuição do grupo aos elementos da aposta recursiva

Aposta recursiva			Ganho	Excede Saldo	Grupo		
Arsenal	Aston Villa	Fulham	5358,69	Não	A	B	C
Arsenal	Bolton	Burnley	3903,64	Sim	A	C	D
Arsenal	Aston Villa	Blackburn	-250,57	Sim	A	B	X
Arsenal	Aston Villa	Tottenham	3721,98	Não	A	B	X
…	…	…	…	…	…	…	…

Da tabela foram retiradas as apostas recursivas que incluíssem clubes pertencentes ao grupo X, ou seja os clubes "surpresa" e também as apostas com resultado negativo. O objectivo era identificar, nas restantes, os padrões mais frequentes.

Tabela 16 - Universo de apostas recursivas pretendidas

Aposta Recursiva			Ganho	Excede Saldo	Padrão da Aposta Recursiva		
Arsenal	Aston Villa	Fulham	5358,69	Não	A	B	C
Arsenal	Bolton	Burnley	3903,64	Não	A	C	D
Arsenal	Aston Villa	Blackburn	250,57	Não	A	B	B
Arsenal	Aston Villa	Tottenham	3721,98	Não	A	B	A
…	…	…	…	…	…	…	…

Para encontrar os padrões mais frequentes utilizamos o algoritmo *apriori*. Foi fixado o parâmetro de suporte mínimo em 10%. Ou de outra forma, foram identificados os padrões mais propensos para o resultado pretendido. A titulo de exemplo os padrões

mais frequentes encontrados foram os que apresentamos na tabela seguinte, para um determinado valor de aposta.

Tabela 17 - Padrões com suporte superior a 10%.

Padrão	Suporte	Número de Apostas Recursivas
{B,C,C}	14%	30
{A,B,C}	11%	45
{C,C,C}	11%	10

Análise dos Resultados da Simulação

A simulação foi analisada segundo os subconjuntos definidos pelos padrões identificados para o respectivo valor da aposta inicial.

É necessário recalcular o ganho global associado a cada padrão. Para a contabilização do ganho global foram incluídos também os ganhos negativos que não poderão ser inferiores a -5.000€, ou seja o valor do saldo inicial. Este processo foi realizado para vários valores de aposta. O melhor resultado global para cada valor de aposta inicial é apresentado na tabela seguinte (ver mais detalhe no anexo).

Tabela 18 - Melhor resultado global para cada valor de aposta inicial

Valor Aposta	# Sucessos	Melhor Padrão	Número de Apostas	Ganho Global
100€	15	{B,C,C}	30	-5.832€
200€	13	{A,C,C}	30	39.451€
300€	10	{A,C,C}	30	59.205€
400€	4	{C,C,C}	10	86.121€
500€	4	{C,C,C}	10	115.150€
600€	5	{A,C,C}	30	97.009€
700€	5	{A,C,C}	30	134.010€
800€	5	{A,C,C}	30	171.010€
900€	5	{A,C,C}	30	208.012€
1000€	4	{A,C,C}	30	178.349€

Na tabela seguinte apresentamos os resultados relativos a todos os padrões extraídos para o caso do valor de aposta 900€ (o que tem maior ganho total).

Tabela 19 - Resultado de cada múltipla com frequência superior a 10%

Frequent Items	Número de apostas	Número de Sucessos	Ganho Global
{A,C,C}	30	5	208.012€
{B,C,C}	30	4	12.110€
{C,C,C}	10	3	83.608€

Os resultados mostram que a combinação que gera maiores ganhos é a combinação {A,C,C}. A combinação {A,C,C} é a combinação de multiplicadores de ganhos {1,61; 4,54; 4,54}.

Com esta experiência baseada em resultados reais conseguimos identificar um grupo de apostas recursivas que nos permitiria obter um ganho total positivo. O que demonstrámos não ser possível sem usar uma forma de escolha das apostas. Contudo não é possível usar esta metodologia na realidade. A identificação dos subgrupos utiliza dados só disponíveis no fim da época e portanto inexistentes no momento da aposta. Embora não tenha sido o objectivo desta tese desenvolver um sistema que pudesse ser usado nesse contexto, mas apenas aferir se é possível identificar subconjuntos de apostas recursivas vencedoras, apresentamos uma possibilidade de como poderíamos estender o método à realidade. O princípio desta extensão é simular a partir de uma dada jornada utilizando os agrupamentos definidos pelos dados das jornadas anteriores.

3.7. Simulação com Arranque na Jornada k

A simulação mais aplicável à realidade consiste nos seguintes passos:

1. Calcular *Odds* médias para cada equipa (restrito às k primeiras jornadas).
2. Efectuar uma análise de agrupamentos baseado nas *Odds* médias obtidas no passo 1.
3. Identificar o conjunto de padrões e calcular a *Odd* média que o representa.
4. Recuperar de estudos anteriores a *Odd* média dos melhores padrões.
5. Calcular o padrão mais próximo segundo a *Odd* média.
6. Simular o processo de apostas considerando as apostas recursivas do padrão mais próximo.

É importante referir que após o passo 3 dispomos do conjunto de padrões possíveis mas não sabemos qual o melhor, pois não dispomos de nenhuma informação sobre o ganho total. A estratégia aqui terá que ser diferente. Decidimos procurar um padrão que tivesse funcionado no passado. Uma vez que a identificação usando letras (ex. {A,C,C}) não tem informação (o padrão {A,A,C} numa época pode corresponder a {C,C,D} noutra), tornou-se necessário usar as *Odds* médias do agrupamento e dessa forma calcular as diferenças para identificar, no novo caso, o padrão mais próximo daquele que se revelou melhor no passado.

Para simular este processo decidimos considerar as 6 primeiras jornadas para recolha de informação. Convém que o número de jornadas seja par de modo a que o factor casa *vs* fora seja compensado. Ou seja que a maioria das equipas tenha um número de jogos em casa igual ao número de jogos fora. Uma vez que se tratam de poucas jornadas esse desequilíbrio influencia bastante o valor médio do multiplicador de ganhos, uma vez que a *Odd* se altera pelo facto de uma equipa jogar em casa ou fora.

Criação de Agrupamentos

Esta análise é realizada segundo o mesmo método usado anteriormente. Os dados para os agrupamentos obtidos estão apresentados na tabela seguinte (versão reduzida. Completa em anexo).

Tabela 20 - Tabela de partida para análise de agrupamentos com 6 jornadas

Team	Points	Win	Draw	Loss	Goal Dif	Odd
Chelsea	15	5	0	1	19	1.31
ManUnited	12	3	3	0	7	1.46
Arsenal	11	3	2	1	9	1.57
ManCity	11	3	2	1	5	2.24
…	…	…	…	…	…	…

Tal como na execução anterior sobre novos dados foi realizada uma análise hierárquica através do método *ward* com distâncias euclidianas. O número de agrupamentos foi fixado em 5. A Figura 6 mostra o dendrograma obtido.

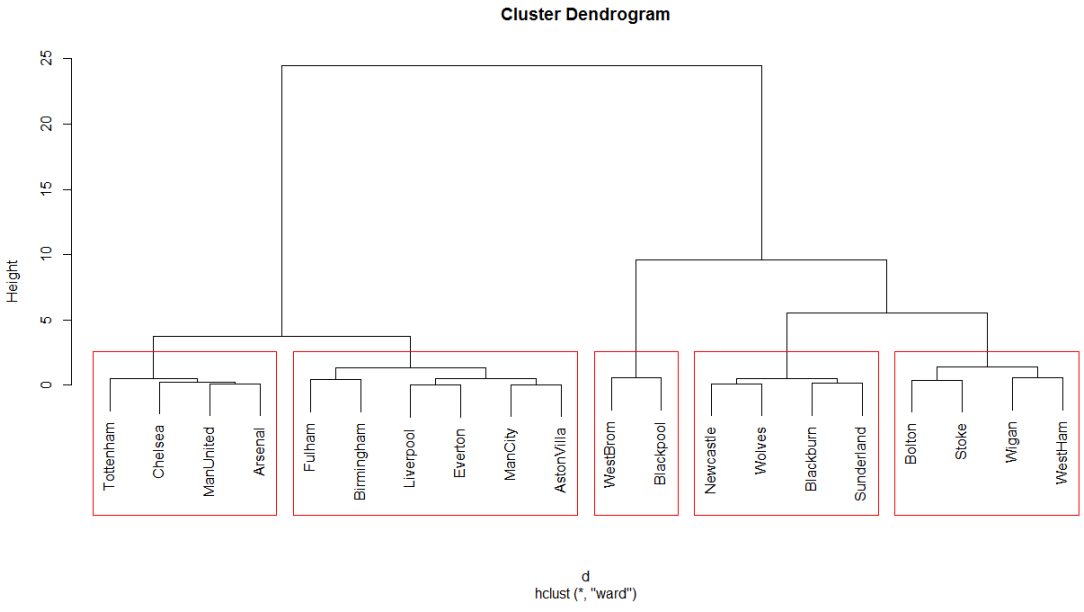


Figura 6 - Dendrograma com variável odd – 5 grupos

Podemos distinguir 5 grupos e efectuar a respectiva caracterização. Cada grupo foi caracterizado pelo valor do multiplicador de ganhos médios. Podemos ver essa informação na tabela seguinte.

Tabela 21 - Caracterização dos agrupamentos

Grupo A	**Grupo D**	**Grupo E**	**Grupo B**	**Grupo C**
Chelsea	Blackburn	Bolton	ManCity	WestBrom
ManUnited	Newcastle	Stoke	AstonVilla	Blackpool
Arsenal	Sunderland	Wigan	Fulham	
Tottenham	Wolves	WestHam	Birmingham	
			Liverpool	
			Everton	
Intervalo do Multiplicador				
[1,31;1,84]	[4,05;4,52]	[5,41;6,83]	[2,24;3,27]	[8,60;9,15]
Multiplicador Médio				
1,54	4,29	6,08	2,62	8,87

No passado (época 2009-2010) o padrão com melhores resultados foi {A,C,C} a que corresponde a combinação de multiplicadores {1,61; 4,54; 4,54}.

Antes de efectuar a simulação na época 2010-2011 foi necessário verificar qual o grupo da época 2010-2011 que corresponde aos multiplicadores {4,54} e {1,61} por forma a criar a combinação de multiplicadores de valor {1,61; 4,54; 4,54}. Assim para cada um dos grupos da época 2010-2011 foi calcula a diferença em módulo entre o valor do seu multiplicador médio e o multiplicador {4,54} e {1,61} (ver tabela 22) e escolhido o grupo com a diferença mínima.

Tabela 22 - Escolha dos grupos correspondentes entre épocas

Odds 2009-2010	Grupos e *Odds* correspondentes (6 jornadas de 2010-2011)				
	A	**D**	**E**	**B**	**C**
	1,54	4,29	6,08	2,62	8,87
4,54	3	0,25	1,54	1,92	4,33
1,61	0,07	2,68	4,47	1,01	7,26

O resultado da análise de agrupamentos para a Premier League época 2010-2011 mostra que a combinação mais próxima de {1,61; 4,54; 4,54}. é a combinação {A,D,D}. Foi então simulada a combinação {A,D,D} com um valor de aposta de 1.000€ e um saldo inicial para cada pasta recursiva de 5.000€. Os resultados foram os seguintes:

Tabela 23 - Resultados da avaliação

Frequent Items	Número de apostas	Número de Sucessos	Ganho Global
{A,D,D}	24	4	77.074€

O resultado é deveras interessante originado uma rendibilidade de 64%. Se tivéssemos acesso às *Odds* finais (usando toda a época e após 6 jornada) – injusta comparação com este método - os padrões mais frequentes identificados e respectivos ganhos totais seriam os seguintes:

Tabela 24 - Ganhos de apostas segundo padrões obtidos usando a informação total

Padrão Frequente	Ganho Total
{B,D,D}	166.816€
{A,B,D}	-229.087€

É curioso constatar que embora na melhor das hipóteses o ganho fosse maior um dos padrões frequentes não consegue ter um ganho global positivo.

No próximo capítulo apresentamos um resumo das conclusões desta tese.

4. Conclusões

Os objectivos propostos foram atingidos. Concretamente:

- Desenvolvemos uma adaptação ao método de Martingale para apostas desportivas
- Conduzimos experiências que demonstraram que o método não pode ser aplicado com sucesso se usarmos todas as apostas sem critério de escolha
- Uma análise dos resultados das experiências por subconjuntos de apostas com determinadas características, mostraram que alguns grupos apresentam ganhos compensadores, ficando por isso a aplicação prática dependente do desenvolvimento de um critério de escolha do subconjunto
- Desenvolvemos uma última versão do método que determinava um subconjunto promissor baseado em informação sobre dados do passado. Corremos experiências que mostraram que o subconjunto de apostas identificado corresponde a um ganho total positivo.

O trabalho pode ser melhorado segundo os seguinte aspectos:

- Realizar experiências de simulação envolvendo outras ligas/épocas ou mesmo apostas múltiplas baseadas em jogos de várias ligas.
- Estender o trabalho adaptando outros sistemas de cobertura de perdas (Ethier, 2010) e comparar os seus desempenhos.

Referências

Alpaydin, E. (2004), "**Introduction to Machine Learning",** MIT Press.

Boulier, B.L. and H. O. Stekler (2003), **"Predicting the outcomes of national footballleague games"**, International Journal of Forecasting, 19(2), 257–270.

Burns, Robert P., Richard Burns (2009), **Business Research Methods and Statistics Using SPSS** SAGE Publications Ltd.

Dandurand, L. (1999), "**A global market analysis of casino gaming on the Internet. In The business of gaming: Economic and management issues**", ed. W. Eadington and J. Cornelius, Reno: Institute for the Study of Gambling and Commercial Gaming, College of Business Administration, University of Nevada, 91–112.

Esperança, José Paulo; Fernanda Matias (2005), "**Finanças Empresariais"**, Publicações Dom Quixote.

Ethier, S.N. (2010), "**The Doctrine of Chances, Probability and its Applications**" Springer-Verlag Berlin Heidelberg.

Goddard, J. and I. Asimakopoulos (2004), "**Forecasting football results and the efficiency of fixed-odds betting"**, Journal of Forecasting, 23 (1), 51–66.

Gomber, Peter *et al* (2008), "**Sports betting as a new asset class—current market organization and options for development",** Financial Markets and Portfolio Management 22: 169–192.

Hahsler, Michael, Bettina Gruen and Kurt Hornik (2008), "**arules: Mining Association Rules and Frequent Itemsets",** R package version 1.0-10.

Harville, D. (1980), **"Predictions for national football league games via linear-model methodology"**, Journal of the American Statistical Association, 75 (371), 516–524.

Kelton, W. David, Randall P. Sadowski, Deborah A. Sadowski (2002), "**Simulation with Arena",** McGraw-Hill.

Neves, João Carvalho das (2002), "**Avaliação de Empresas e de Negócios"**, McGraw Hill.

R Development Core Team (2012), "**R: A Language and Environment for Statistical Computing"**, R Foundation for Statistical Computing.

Stefani, R.T. (1980), "**Improved least squares football, basketball, and soccer predictions"**, Systems, Man and Cybernetics, IEEE Transactions on, 10(2), 116–123.

Stewart, D. O. (2006),"**An analysis of Internet gambling and its policy implication".**

Bwin**, Bwin - Interactive Entertainment AG –** http://www.bwin.com, (acedido em 2012)

Football-Data, **Football-Data.co.uk - Results / Odds / Transfers**, http://www.football-data.co.uk, (acedido em 2012)

H2 Gambling Capital, **Preliminary Results 2011 H2 eGaming Dataset Now Available,** http://www.h2gc.com/article/New-H2-eGaming-Dataset-Now-Available, (acedido em 2012)

Anexos

Tabela de partida para análise de agrupamentos em subconjuntos de dados

Team	Points	Win	Draw	Loss	Goal Dif	Odd
Chelsea	86	27	5	6	71	1,478
Man United	85	27	4	7	58	1,587
Arsenal	75	23	6	9	42	1,751
Tottenham	70	21	7	10	26	2,397
Man City	67	18	13	7	28	2,237
Aston Villa	64	17	13	8	13	2,719
Liverpool	63	18	9	11	26	1,915
Everton	61	16	13	9	11	3,035
Birmingham	50	13	11	14	-9	4,371
Blackburn	50	13	11	14	-14	4,300
Stoke	47	11	14	13	-14	4,867
Fulham	46	12	10	16	-7	4,055
Sunderland	44	11	11	16	-8	4,225
Bolton	39	10	9	19	-25	4,777
Wolves	38	9	11	18	-24	5,388
Wigan	36	9	9	20	-42	4,772
West Ham	35	8	11	19	-19	4,429
Burnley	30	8	6	24	-40	5,639
Hull	30	6	12	20	-41	5,591
Portsmouth	28	7	7	24	-32	5,909

Diferença entre o número de vitórias conseguidas e o número de vitórias esperadas

Equipa	V. Conseguidas	Odd Media	V. Esperadas	Diferença
Chelsea	27	1,478	26	1
Man United	27	1,587	24	3
Arsenal	23	1,751	22	1
Tottenham	21	2,397	16	5
Man City	18	2,237	17	1
Aston Villa	17	2,719	14	3
Liverpool	18	1,915	20	-2
Everton	16	3,035	13	3
Birmingham	13	4,371	9	4
Blackburn	13	4,300	9	4
Stoke	11	4,867	8	3
Fulham	12	4,055	9	3
Sunderland	11	4,225	9	2
Bolton	10	4,777	8	2
Wolves	9	5,388	7	2
Wigan	9	4,772	8	1
West Ham	8	4,429	9	-1
Burnley	8	5,639	7	1
Hull	6	5,591	7	-1
Portsmouth	7	5,909	6	1

Nova tabela para análise expurgada de clubes "surpresa"

Team	Points	Win	Draw	Loss	Goal Dif	Odd
Chelsea	86	27	5	6	71	1.48
Man United	85	27	4	7	58	1.59
Arsenal	75	23	6	9	42	1.75
Man City	67	18	13	7	28	2.24
Aston Villa	64	17	13	8	13	2.72
Everton	61	16	13	9	11	3.04
Stoke	47	11	14	13	-14	4.87
Fulham	46	12	10	16	-7	4.06
Sunderland	44	11	11	16	-8	4.23
Bolton	39	10	9	19	-25	4.78
Wolves	38	9	11	18	-24	5.39
Wigan	36	9	9	20	-42	4.77
Burnley	30	8	6	24	-40	5.64
Portsmouth	28	7	7	24	-32	5.91

Execução para um valor de aposta de 1000€

Foram seleccionados as combinações cujo valor da aposta não excede o saldo disponível e cujo resultado final resulte num ganho. Para um valor de aposta de 1.000€ os resultados estão presentes na tabela

Frequent Items	Suporte
{A,C,C}	17%
{C,C,C}	13%
{B,C,C}	13%

Os resultados globais para um valor de aposta de 1.000€ estão presentes na tabela

Frequent Items	Número de apostas	Número de Sucessos	Ganho Global
{A,C,C}	30	4	178.349€
{C,C,C}	10	3	96.786€
{B,C,C}	30	3	12.775€

Execução para um valor de aposta de 900€

Foram seleccionados as combinações cujo valor da aposta não excede o saldo disponível e cujo resultado final resulte num ganho. Para um valor de aposta de 900€ os resultados estão presentes na tabela

Frequent Items	Suporte
{A,C,C}	20%
{B,C,C}	16%
{C,C,C}	12%

Os resultados globais para um valor de aposta de 900€ estão presentes na tabela

Frequent Items	Número de apostas	Número de Sucessos	Ganho Global
{A,C,C}	30	5	208.012€
{B,C,C}	30	4	12.110€
{C,C,C}	10	3	83.608€

Execução para um valor de aposta de 800€

Foram seleccionados as combinações cujo valor da aposta não excede o saldo disponível e cujo resultado final resulte num ganho. Para um valor de aposta de 800€ os resultados estão presentes na tabela

Frequent Items	Suporte
{B,C,C}	14%
{A,C,C}	18%
{A,B,C}	11%
{C,C,C}	11%

Os resultados globais para um valor de aposta de 800€ estão presentes na tabela

Frequent Items	Número de apostas	Número de Sucessos	Ganho Global
{B,C,C}	30	4	-3.681€
{A,C,C}	30	5	171.010€
{A,B,C}	45	3	-91.245€
{C,C,C}	10	3	70.429€

Execução para um valor de aposta de 700€

Foram seleccionados as combinações cujo valor da aposta não excede o saldo disponível e cujo resultado final resulte num ganho. Para um valor de aposta de 700€ os resultados estão presentes na tabela

Frequent Items	Suporte
{B,C,C}	21%
{A,C,C}	17%
{A,B,C}	10%
{C,C,C}	10%

Os resultados globais para um valor de aposta de 700€ estão presentes na tabela

Frequent Items	Número de apostas	Número de Sucessos	Ganho Global
{B,C,C}	30	6	130.802€
{A,C,C}	30	5	134.010€
{A,B,C}	45	3	-106.089€
{C,C,C}	10	3	57.250€

Execução para um valor de aposta de 600€

Foram seleccionados as combinações cujo valor da aposta não excede o saldo disponível e cujo resultado final resulte num ganho. Para um valor de aposta de 600€ os resultados estão presentes na tabela

Frequent Items	Suporte
{B,C,C}	20%
{A,C,C}	17%
{A,B,C}	10%
{C,C,C}	10%

Os resultados globais para um valor de aposta de 600€ estão presentes na tabela

Frequent Items	Número de apostas	Número de Sucessos	Ganho Global
{B,C,C}	30	6	94.973€
{A,C,C}	30	5	97.009€
{A,B,C}	45	3	-120.934€
{C,C,C}	10	3	44.072€

Execução para um valor de aposta de 500€

Foram seleccionados as combinações cujo valor da aposta não excede o saldo disponível e cujo resultado final resulte num ganho. Para um valor de aposta de 500€ os resultados estão presentes na tabela

Frequent Items	Suporte
{B,C,C}	23%
{A,C,C}	17%
{C,C,C}	11%

Os resultados globais para um valor de aposta de 500€ estão presentes na tabela

Frequent Items	Número de Apostas	Número de Sucessos	Ganho Global
{B,C,C}	30	8	103.365€
{A,C,C}	30	6	87.753€
{C,C,C}	10	4	115.151€

Execução para um valor de aposta de 400€

Foram seleccionados as combinações cujo valor da aposta não excede o saldo disponível e cujo resultado final resulte num ganho. Para um valor de aposta de 400€ os resultados estão presentes na tabela

Frequent Items	Suporte
{B,C,C}	21%
{A,C,C}	16%
{A,B,C}	13%
{C,C,C}	11%

Os resultados globais para um valor de aposta de 400€ estão presentes na tabela

Frequent Items	Número de apostas	Número de Sucessos	Ganho Global
{B,C,C}	30	8	60.692€
{A,C,C}	30	6	46.202€
{A,B,C}	45	5	-118.073€
{C,C,C}	10	4	86.121€

Execução para um valor de aposta de 300€

Foram seleccionados as combinações cujo valor da aposta não excede o saldo disponível e cujo resultado final resulte num ganho. Para um valor de aposta de 300€ os resultados estão presentes na tabela

Frequent Items	Suporte
{B,C,C}	21%
{A,C,C}	16%
{A,B,C}	13%
{C,C,C}	11%

Os resultados globais para um valor de aposta de 300€ estão presentes na tabela

Frequent Items	Número de apostas	Número de Sucessos	Ganho Global
{B,C,C}	30	8	18.019€
{A,C,C}	30	10	59.205€
{A,B,C}	45	7	-111.677€
{C,C,C}	10	4	57.090€

Execução para um valor de aposta de 200€

Foram seleccionados as combinações cujo valor da aposta não excede o saldo disponível e cujo resultado final resulte num ganho. Para um valor de aposta de 200€ os resultados estão presentes na tabela

Frequent Items	Suporte
{A,C,C}	18%
{B,C,C}	11%
{A,B,C}	11%

Os resultados globais para um valor de aposta de 200€ estão presentes na tabela

Frequent Items	Número de apostas	Número de Sucessos	Ganho Global
{A,C,C}	30	13	39.452€
{B,C,C}	30	8	-21.372€
{A,B,C}	45	8	-129.244€

Execução para um valor de aposta de 100€

Foram seleccionados as combinações cujo valor da aposta não excede o saldo disponível e cujo resultado final resulte num ganho. Para um valor de aposta de 100€ os resultados estão presentes na tabela

Frequent Items	Suporte
{A,B,C}	16%
{B,C,C}	14%
{A,C,C}	13%

Os resultados globais para um valor de aposta de 100€ estão presentes na tabela

Frequent Items	Número de apostas	Número de Sucessos	Ganho Global
{A,B,C}	45	18	-64.656€
{B,C,C}	30	15	-5.382€
{A,C,C}	30	14	-11.696€

Tabela de partida onde apenas foram consideradas 6 jornadas

Team	Points	Win	Draw	Loss	Goal Dif	Odd
Chelsea	15	5	0	1	19	1.31
ManUnited	12	3	3	0	7	1.46
Arsenal	11	3	2	1	9	1.57
ManCity	11	3	2	1	5	2.24
AstonVilla	10	3	1	2	-2	2.28
WestBrom	10	3	1	2	-3	9.15
Fulham	8	1	5	0	1	3.27
Tottenham	8	2	2	2	1	1.84
Blackburn	8	2	2	2	0	4.21
Newcastle	7	2	1	3	1	4.52
Sunderland	7	1	4	1	0	4.05
Birmingham	7	1	4	1	-1	2.86
Bolton	7	1	4	1	-1	5.41
Stoke	7	2	1	3	-2	5.80
Blackpool	7	2	1	3	-5	8.60
Liverpool	6	1	3	2	-3	2.55
Wolves	5	1	2	3	-3	4.40
Wigan	5	1	2	3	-11	6.28
WestHam	4	1	1	4	-9	6.83
Everton	3	0	3	3	-3	2.53

Código R

```
Setup.Table= function(datafile,epoca,liga,teams){

Equipa<-NULL; Setup.Table<-NULL

for(i in seq(along=teams)) {

Equipa<-teams[i]

tmpResults = read.csv(datafile, sep = ";",
stringsAsFactors=FALSE)[,c("Div","Season","Round","Date","HomeTeam","A
wayTeam","FTR","FTHG","FTAG", "BWH", "BWD", "BWA")]

#Criar tabela

tmpTable = data.frame(tmpResults, ValorAposta=0, WinTeam=0, Odd=0 ,
Result=0)

#Colocar equipa vencedora

tmpTable$WinTeam=

ifelse(tmpResults$FTR=="D","Draw",ifelse(tmpTable$FTR=="H",tmpResults$
HomeTeam,tmpResults$AwayTeam))

#Alimentar coluna Odd com o valor da equipa escolhida

tmpTable$Odd= ifelse(tmpTable$HomeTeam==Equipa
,tmpTable$BWH,tmpTable$BWA)

#Reduzir tabela

tmpTable<-subset(tmpTable,(Season==epoca & Div==liga &
(HomeTeam==Equipa | AwayTeam==Equipa)))

Setup.Table<-(rbind(Setup.Table, tmpTable) )

Setup.Table<-Setup.Table[order(Setup.Table$Round),]

}

return(Setup.Table)

}

Resultado.Multipla<-function(Setup.Table,teams){

result<-sapply(1:length(unique(Setup.Table$Round)), function(jorn) {
```

```
    ganhou<-T
    for(idgame in which(Setup.Table$Round==jorn)) {
        game<-Setup.Table[idgame,]
        if (! game$WinTeam %in% teams | game$WinTeam=="Draw") {
            ganhou<-F
        } else {
            if (game$AwayTeam %in% teams && game$HomeTeam  %in% teams
&& game$WinTeam==game$AwayTeam) {
                ganhou<-F
            }
        }
    }
    return(ganhou)
          })
Tabela.Com.Resultado.Multipla<-as.data.frame
(cbind(Round=unique(Setup.Table$Round),
Result=result,Odd=by(Setup.Table$Odd,Setup.Table$Round,prod)))
 return(Tabela.Com.Resultado.Multipla)
}
Elimina.Equipas.Escolhidas<-
function(Setup.Table,teams,Tabela.Com.Resultado.Multipla){
Tabela.Com.Equipas.Escolhidas<-subset(Setup.Table,(HomeTeam %in% teams
& AwayTeam %in% teams))
                Jornada.Equipas.Escolhidas<-
(unique(Tabela.Com.Equipas.Escolhidas$Round))
                Tabela.Sem.Equipas.Escolhidas<-
subset(Tabela.Com.Resultado.Multipla,!          (Round          %in%
Jornada.Equipas.Escolhidas))
return(Tabela.Sem.Equipas.Escolhidas)
```

```
}
Valor.Aposta<-function(Tabela.Sem.Equipas.Escolhidas,bet){
valor.aposta<-NULL; primeira.derrota<-NULL; acumula.perdas<-NULL
for(i in seq(along= Tabela.Sem.Equipas.Escolhidas$Round)) {
    if (i==1)
{
                              valor.aposta[i]=(bet)
                              primeira.derrota<-(i)
                         }
      else{
if (Tabela.Sem.Equipas.Escolhidas$Result[i-1]==1)
{
valor.aposta[i] <- bet
                  primeira.derrota<-(i)
                  } else
                        {
                  acumula.perdas                                    <-
sum(c(valor.aposta[primeira.derrota:(i-1)]))
                  valor.aposta[i]                                   <-
(acumula.perdas/(Tabela.Sem.Equipas.Escolhidas$Odd[i]-1)+bet)
                         }
           }
                         }
Bet<-round(valor.aposta,2)
Tabela.Com.Valor.Aposta<-as.data.frame(Bet)
return(Tabela.Com.Valor.Aposta)
```

```
      }
Ganho.Perda<-
function(Tabela.Sem.Equipas.Escolhidas,Tabela.Com.Valor.Aposta){
Win.Loss=
ifelse(Tabela.Sem.Equipas.Escolhidas$Result=1,
Tabela.Com.Valor.Aposta$Bet*Tabela.Sem.Equipas.Escolhidas$Odd-
Tabela.Com.Valor.Aposta$Bet, Tabela.Com.Valor.Aposta$Bet*-1)
Win.Loss<-round(Win.Loss,2)
Tabela.Ganho.Perda<-as.data.frame(Win.Loss)
return(Tabela.Ganho.Perda)
}
Ganho.Perda.Acumulado<-function(Tabela.Ganho.Perda){
Cumulated<-NULL
for(i in seq(along= Tabela.Ganho.Perda$Win.Loss)) {
    if (i==1)
                        {
      Cumulated[i]=Tabela.Ganho.Perda$Win.Loss[i]
                        }
      else{
            Cumulated[i]=Tabela.Ganho.Perda$Win.Loss[i]+Cumulated[i-1]
            }
}
Ganho.Perda.Acumulado<-as.data.frame(Cumulated)
return(Ganho.Perda.Acumulado)
            }
Balance<-function(Tabela.Ganho.Perda){
Saldo<-NULL ; Saldo.Inicial<-5000
```

```
for(i in seq(along=Tabela.Ganho.Perda$Win.Loss)) {

    if (i==1)


                        {

                              Saldo[i]=Saldo.Inicial


                        }

      else{


            Saldo[i]==Saldo[i-1]+Tabela.Ganho.Perda$Win.Loss[i-1]

            }

}

Balance<-as.data.frame(Saldo)

return(Balance)

                  }

#Saldo.Disponivel.Para.Aposta<-
function(Tabela.Balance,Tabela.Com.Valor.Aposta,teams){

#

#    if
      (any(Tabela.Balance$Saldo<Tabela.Com.Valor.Aposta$Bet))

#                       {

# cat(teams,"- O valor da aposta excede o saldo disponivel.

# ")


                        # }

      # else{

# cat(teams,"- Sempre com saldo disponivel para executar a aposta.
```

```
# ")

            # }

# return(Saldo.Disponivel.Para.Aposta)

                                    # }

Introduz.Aposta=function(){

aposta<-0

cat("Introduza valor da aposta = ")

t<-readLines (n=1)

aposta<-t

bet<-as.numeric(aposta)

return(bet)

}

Lista.Epocas=function(datafile){

Lista.Epocas<-read.csv(datafile,          sep          =          ";",
stringsAsFactors=FALSE)[,c("Season")]

Lista.Epocas<-sort(unique(c(as.character(Lista.Epocas))))

print(Lista.Epocas)

}

# Introduz.Epocas<-function(){

# epocas<-c()

# repeat{

# cat("Introduza Epocas (0 para sair) = ")

# t<-readLines (n=1)

# if(t==0) break

# epocas<-c(epocas,t)

# }
```

```
# return(epocas)
# }
Introduz.Epoca<-function(){
epoca<-NULL
cat("Introduza Epoca = ")
t<-readLines (n=1)
epoca<-t
return(epoca)
}
Lista.Ligas=function(datafile){
Lista.Ligas=read.csv(datafile, sep = ";",
stringsAsFactors=FALSE)[,c("Div","Season","HomeTeam")]
Ligas=sort(unique(c(as.character(Lista.Ligas$Div))))
print(Ligas)
}
# Introduz.Ligas<-function(){
# ligas<-c()
# repeat{
# cat("Introduza Ligas (0 para sair) = ")
# t<-readLines (n=1)
# if(t==0) break
# ligas<-c(ligas,t)
# }
# return(ligas)
# }
Introduz.Liga<-function(){
```

```
liga<-NULL

cat("Introduza Liga = ")

t<-readLines (n=1)

liga<-t

return(liga)

}

Lista.Equipas=function(datafile,ligas,epoca){

 Lista.Equipas      <-      read.csv(datafile,      sep      =      ";",
stringsAsFactors=FALSE)[,c("HomeTeam","Div","Season")]

 Lista.Equipas<-data.frame(Lista.Equipas)

 Lista.Equipas<-subset(Lista.Equipas,Lista.Equipas$Div==ligas         &
Lista.Equipas$Season==epoca)

 Lista.Equipas<- sort(unique(c(as.character(Lista.Equipas$HomeTeam))))

 print(Lista.Equipas)

 }

Introduz.Equipas<-function(){

teams<-c()

repeat{

cat("Introduza Equipas (0 para sair) = ")

t<-readLines (n=1)

if(t==0) break

teams<-c(teams,t)

}

return(teams)

}

Valida.Codigo<-
function(Tabela.Sem.Equipas.Escolhidas,Tabela.Com.Valor.Aposta,Tabela.
Ganho.Perda,Tabela.Ganho.Perda.Acumulado,Tabela.Balance){
```

```
Tabela.De.Validacao<-
cbind(Tabela.Sem.Equipas.Escolhidas,Tabela.Balance,
Tabela.Com.Valor.Aposta,Tabela.Ganho.Perda,
Tabela.Ganho.Perda.Acumulado)

print(Tabela.De.Validacao)

}

Sumario<-
function(Tabela.Com.Valor.Aposta,Tabela.Ganho.Perda,ligas,epoca,teams,
Tabela.Balance){

Aposta.Maxima<-as.vector(Tabela.Com.Valor.Aposta$Bet)

Aposta.Maxima<-max(Aposta.Maxima)

Ganho.Total<-as.vector(Tabela.Ganho.Perda$Win.Loss)

Ganho.Total<-sum(Ganho.Total)

Saldo.Minimo<-as.vector(Tabela.Balance$Saldo)

Saldo.Minimo<-min(Saldo.Minimo)

Aposta.Quando.Saldo.Minimo<-
subset(Tabela.Com.Valor.Aposta$Bet,Tabela.Balance$Saldo==Saldo.Minimo)

Saldo.Quando.Aposta.Maxima<-
subset(Tabela.Balance$Saldo,Tabela.Com.Valor.Aposta$Bet==Aposta.Maxima
)

if (any(Tabela.Balance$Saldo < Tabela.Com.Valor.Aposta$Bet)) {

#    cat(teams, "- O valor da aposta excede o saldo disponivel.\n\n")

      excede<-"Sim"

}

else {

 #     cat(teams, "- Sempre com saldo disponivel para executar a
aposta.\n\n\n")

      excede<-"Não"

}
```

```
Summary = list(Liga=ligas, Epoca=epoca, Equipa = teams, Aposta.Maxima=Aposta.Maxima,Saldo.Quando.Aposta.Maxima=Saldo.Quando.Aposta.Maxima,Saldo.Minimo=Saldo.Minimo,Aposta.Quando.Saldo.Minimo=Aposta.Quando.Saldo.Minimo, Ganho.Total=Ganho.Total,Excede.Aposta=excede)

print(Summary)

}

Lista.Equipas.Combinadas<-function(datafile,ligas,epoca){

 Lista.Equipas <- read.csv(datafile, sep = ";", stringsAsFactors=FALSE)[,c("HomeTeam","Div","Season")]

 Lista.Equipas <-data.frame(Lista.Equipas)

 Lista.Equipas<-subset(Lista.Equipas,Lista.Equipas$Div==ligas & Lista.Equipas$Season==epoca)

 Lista.Equipas<- sort(unique(c(as.character(Lista.Equipas$HomeTeam))))

 Combina.Equipas<- combn(Lista.Equipas,3 )

 Tabela.Combina.Equipas<-as.matrix(Combina.Equipas)

return(Tabela.Combina.Equipas[,,drop=FALSE])

 }

Simula.Betting=function(datafile){

Lista.Epocas(datafile)

epoca <-Introduz.Epoca()

Lista.Ligas(datafile)

ligas <-Introduz.Liga()

#Lista.Equipas(datafile,ligas,epoca)

#teams <-Introduz.Equipas()

teams <-Lista.Equipas.Combinadas(datafile,ligas,epoca)

bet <-Introduz.Aposta()

Executa.Funcoes<-function(bet,datafile,teams,ligas,epoca){
```

```
cat("Liga;Epoca;Equipa;Aposta.Max;Saldo.Quando.Aposta.Maxima;Saldo.Min
imo;Aposta.Quando.Saldo.Minimo;Ganho.Total;Excede.Aposta\n",file="expe
riencia.csv")

Simula.Betting.P<-function(teams){

Setup.Table                                    <-
Setup.Table(datafile,epoca,ligas,teams)

Tabela.Com.Resultado.Multipla                                          <-
Resultado.Multipla(Setup.Table,teams)

Tabela.Sem.Equipas.Escolhidas                                          <-
Elimina.Equipas.Escolhidas(Setup.Table,teams,Tabela.Com.Resultado.Mult
ipla)

Tabela.Com.Valor.Aposta                                                <-
Valor.Aposta(Tabela.Sem.Equipas.Escolhidas,bet)

Tabela.Ganho.Perda                                                     <-
Ganho.Perda(Tabela.Sem.Equipas.Escolhidas,Tabela.Com.Valor.Aposta)

Tabela.Ganho.Perda.Acumulado                                           <-
Ganho.Perda.Acumulado(Tabela.Ganho.Perda)

Tabela.Balance                          <-Balance(Tabela.Ganho.Perda)

#Tabela.Para.Validar.Codigo                                            <-
Valida.Codigo(Tabela.Sem.Equipas.Escolhidas,Tabela.Com.Valor.Aposta,Ta
bela.Ganho.Perda,Tabela.Ganho.Perda.Acumulado, Tabela.Balance)

resumo<-
Sumario(Tabela.Com.Valor.Aposta,Tabela.Ganho.Perda,ligas,epoca,teams,T
abela.Balance)

#Tabela.Saldo.Disponivel.Para.Aposta                                   <-
Saldo.Disponivel.Para.Aposta(Tabela.Balance,Tabela.Com.Valor.Aposta,te
ams)

cat(resumo[[1]],";",resumo[[2]],";",paste(resumo[[3]],collapse="
"),";",resumo[[4]],";",resumo[[5]],";",resumo[[6]],";",resumo[[7]],";"
,resumo[[8]],";",resumo[[9]],"\n",file="experiencia.csv",append=T)

return(resumo)

}
```

```
apply(teams, 2, function(x) {Simula.Betting.P(x)})

}

Tabela.Executa.Funcoes<-
Executa.Funcoes(bet,datafile,teams,ligas,epoca)

return(Tabela.Executa.Funcoes)

}

Simula.Betting("T1.csv")
```