

PONTIFÍCIA UNIVERSIDADE CATÓLICA DO RIO DE JANEIRO  
DEPARTAMENTO DE ECONOMIA

MONOGRAFIA DE FINAL DE CURSO

**ANÁLISE DAS ODDS: UM ESTUDO DA RELAÇÃO RISCO-RETORNO DO  
MERCADO DE APOSTAS ESPORTIVAS**

ROBERTO TOURIÑO YAZEJI CARDOSO  
NÚMERO DE MATRÍCULA: 1312867

ORIENTADOR: LEONARDO REZENDE

JUNHO DE 2017

PONTIFÍCIA UNIVERSIDADE CATÓLICA DO RIO DE JANEIRO  
DEPARTAMENTO DE ECONOMIA

MONOGRAFIA DE FINAL DE CURSO

**ANÁLISE DAS ODDS: UM ESTUDO DA RELAÇÃO RISCO-RETORNO DO  
MERCADO DE APOSTAS ESPORTIVAS**

ROBERTO TOURIÑO YAZEJI CARDOSO  
NÚMERO DE MATRÍCULA: 1312867

ORIENTADOR: LEONARDO REZENDE

JUNHO 2017

Declaro que o presente trabalho é de minha autoria e que não recorri para realizá-lo, a nenhuma forma de ajuda externa, exceto quando autorizado pelo professor tutor.

---

Roberto Touriño Yazeji Cardoso

As opiniões expressas neste trabalho são de responsabilidade única e exclusiva do autor.

## **Agradecimentos**

Ao meu orientador, professor Leonardo Rezende, por toda paciência e pronta disposição em ajudar.

A todos os meus amigos, em especial àqueles que conviveram comigo nesses últimos 4 anos e meio nos bons e maus momentos.

À minha namorada, Natalia Hollanda, maior presente que eu poderia ter recebido ao longo do curso, por todo apoio e cumplicidade.

Por fim, aos meus pais, Paulo Yazeji e Diane Touriño, por serem a base da minha existência e responsáveis por absolutamente todas as oportunidades que recebi.

## Sumário

1	Introdução.....	6
2	Revisão Bibliográfica.....	8
3	Base de Dados .....	13
4	Metodologia .....	15
5	Resultados .....	20
5.1	Resultado Geral.....	21
5.2	Controlando para o Spread.....	29
5.3	Caso específico: Análise de times gigantes .....	34
6	Conclusão .....	37
7	Referências Bibliográficas .....	38

## Índice de Tabelas

Tabela I.....	21
Tabela II.....	22
Tabela III .....	22
Tabela IV .....	23
Tabela V .....	23
Tabela VI.....	25
Tabela VII.....	25
Tabela VIII .....	26
Tabela IX .....	26
Tabela X .....	27
Tabela XI.....	30
Tabela XII.....	31
Tabela XIII .....	31
Tabela XIV .....	32
Tabela XV.....	32
Tabela XVI.....	35
Tabela XVII.....	35

# 1 Introdução

Poucas são as coisas genuinamente democráticas ao redor do mundo, pois basicamente tudo, em alguma instância, está relacionado ao poder aquisitivo do indivíduo. Certamente pode-se considerar que a paixão pelo esporte está contida nesse seleto grupo em que pessoas de diferentes realidades podem compartilhar um mesmo sentimento e que esta, portanto, possui uma abrangência fenomenal.

É essencial frisar que tal paixão não se traduz somente na prática esportiva, mas também em conversas, discussões, reuniões em casa de amigos para assistir jogos e inúmeras outras situações do cotidiano de uma pessoa comum.

Por possuir um alto teor de subjetividade, o esporte permite diversas opiniões acerca de um mesmo evento. Enquanto os argentinos acreditam que Maradona foi o maior jogador de futebol já visto, para os brasileiros este título pertence, indubitavelmente, ao Pelé. Talvez neste exemplo as opiniões estejam enviesadas pelas paixões nacionalistas, mas, em geral, quanto mais você sabe sobre um esporte, maiores são as chances de você estar correto em suas opiniões e previsões.

Existindo essa grande demanda a ser explorada e com o fácil acesso à internet sendo amplamente disseminado, não é de se estranhar que esteja havendo um crescimento exponencial do número de amantes de esportes que estão ingressando em sites de apostas esportivas.

Seja por diversão, por achar que entende muito sobre o assunto ou para ganhar dinheiro, o fato é que sites como Bet365, Bwin e Interwetten estão cada vez mais populares. Mas será que é possível realizar bons investimentos nesse mercado? Será que a remuneração oferecida costuma ser condizente com os riscos envolvidos nessas apostas?

Para poder responder tais questões, é necessário entender plenamente o funcionamento dos sites de apostas esportivas e, para que tal compreensão seja possível, é imprescindível realizar a conceituação do alicerce deste mundo: as odds.

Em termos informais, as odds representam o quanto a casa de aposta está disposta a remunerar cada real do apostador caso o evento em questão venha de fato a ocorrer.

Apesar de existirem outras notações, as odds são tipicamente expressas em forma de número decimal. Por exemplo, se a odd está cotada em 3,5, significa que caso se invista 1 real e acerte o palpite, o prêmio recebido será de 3,5 reais.

Vale ressaltar que o processo de formação das odds não é inteiramente de conhecimento público, pois o mesmo não é divulgado abertamente pelos sites. Entretanto, sabe-se que as odds são fortemente influenciadas por complexos algoritmos utilizados pelas casas de apostas e pela demanda dos apostadores.

Dessa maneira, a fim de prover uma análise completa e profunda da relação risco-retorno existente no mundo das apostas esportivas, a presente monografia se divide em sete capítulos. Além da Introdução, também serão desenvolvidos: Revisão Bibliográfica, Base de Dados, Metodologia, Resultados, Conclusão e Referências Bibliográficas, exatamente nesta ordem.

## 2 Revisão Bibliográfica

Não há dúvidas em relação à enorme expansão do mercado de apostas esportivas ao redor do mundo nos últimos anos. Segundo um relatório publicado, em 2014, pelo Centro Internacional para Segurança no Esporte, este mercado alcança o valor de 500 bilhões de euros por ano. Assim, de maneira não surpreendente, mais relevante se tornou o seu pleno entendimento e, conseqüentemente, diversos estudos têm sido produzidos a fim de destrinchar os mistérios por trás deste tema.

Para o prosseguimento do capítulo, é importante ressaltar duas grandes verdades. Primeiro, que apesar do crescimento e popularização terem acontecido em um passado mais recente, há muitos trabalhos de qualidade acerca das apostas esportivas desde a metade final do século passado. Segundo, que esta bibliografia possui um alto viés de heterogeneidade: enquanto muitos autores focam na análise das odds e dos comportamentos dos apostadores e a conseqüente eficiência deste mercado, alguns outros tentam construir modelos de previsão de resultados de jogos para, assim, poder auferir lucros neste mundo.

Tendo isto em mente, as ideias centrais de algumas obras desta literatura merecem ser percorridas nesta revisão bibliográfica. A profundidade que será dada em cada um destes trabalhos foi definida em função do quão relevante os seus métodos e conclusões parecem ser para o desenvolvimento da presente monografia. Dessa maneira, este capítulo será construído em um formato dicotômico: três principais trabalhos serão analisados de maneira mais detalhada, enquanto o restante será repaginado de maneira mais rasa, nesta ordem.

Situando-se no segundo grupo da heterogeneidade descrita no segundo parágrafo, Airol di (2014) é, sem dúvidas, uma das obras mais interessantes no ramo. Apesar de ter como principal enfoque a Liga Italiana de Futebol (“Série A League”), o seu grande objetivo foi implementar um método de previsão de resultados que pudesse ser aplicado a partidas de qualquer outro campeonato de futebol.

Atualmente, nos sites de apostas esportivas, é possível apostar na ocorrência de inúmeros eventos, tais como o número de escanteios que acontecerão no primeiro tempo, se um time será campeão no final do ano ou até mesmo se um jogador irá fazer um gol na partida seguinte. Devido ao óbvio excesso de complexidade que enfrentaria

para prever tais variáveis, o autor restringe o seu estudo à previsão do resultado simples do jogo, isto é, às probabilidades de um time vencer, perder ou empatar.

Na prática, a estratégia adotada foi a de apostar de modo cumulativo, isto é, o eventual ganho de uma aposta deveria ser utilizado para a aposta seguinte e assim por diante. Tal técnica é bastante arriscada, mas, se aplicada corretamente, pode levar a ganhos extraordinários. Como o método cumulativo culmina em sucessivas apostas, uma análise profunda de cada uma das partidas não parecia ser razoável.

Assim, assume-se que as odds oferecidas pelas casas de apostas já incorporam a maioria das variáveis das partidas, devendo, portanto, o trabalho em questão atentar-se somente a algumas consideradas essenciais, como as performances históricas dos times envolvidos.

Segundo o autor, as odds são calculadas através de estatísticas passadas e ajustadas pela demanda dos apostadores. Além disso, ao mostrar que as odds oferecidas refletem uma soma de probabilidades de vitória, derrota ou empate (as três únicas ocorrências possíveis dentro de uma partida) maiores do que 100%, conclui-se que as casas de apostas inferem nas suas cotações uma espécie de “spread” a seu favor, o que dificultaria ainda mais o seu objetivo final de obter lucro neste mercado.

Dessa maneira, o estudo tenta desenvolver seu programa de previsão, batizado como “Sviluppo”, a fim de eliminar o efeito desses spreads nas odds. A ideia básica do Sviluppo é, utilizando como ponto de partida as odds oferecidas pelo mercado, determinar a probabilidade correspondente da ocorrência do evento e a comparar com as probabilidades históricas (geradas através de base de dados) e, finalmente, avaliar o potencial de resultado correto. O Sviluppo utiliza permutações dos resultados esperados para tornar a previsão mais precisa, reduzindo a exposição a erros, o que é essencial, uma vez que a estratégia de aposta cumulativa é extremamente arriscada.

Apesar de ter se mostrado eficiente na amostra utilizada no estudo, o autor faz questão de reiterar que o sistema deve ser testado muito mais vezes para que se possa ter uma real conclusão sobre a sua eficiência.

A segunda obra que também merece destaque é Snoeberg & Wolfers (2010). Neste estudo, os autores têm como pretensão procurar as verdadeiras razões para a recorrente

e curiosa ocorrência de um evento no mercado de apostas: a superestimação, por parte dos apostadores, dos “azarões” ou “zebras”.

Essa má atribuição de probabilidades foi notada pela primeira vez em 1949 e, desde então, existem duas correntes de pensamento que se propõem a explicar tal fenômeno: a Neoclássica e a embasada em Teorias comportamentais.

A explicação Neoclássica se pauta, basicamente, no estudo das funções de utilidade dos indivíduos. Segundo este grupo, a disposição a pagar (neste caso, apostar) pode ser utilizada para se obter a função de utilidade do apostador e, como a evidência empírica torna nítido o fato de que se aposta nos “azarões” com uma frequência maior do que se realmente deveria, postula-se que o motivo para isto é que os indivíduos possuem uma função de utilidade que remete a uma “paixão pelo risco”.

Alternativamente, a segunda corrente busca explicações através do campo psicológico. Diversos estudos evidenciam que os indivíduos costumam ter dificuldades em avaliar corretamente a probabilidade de ocorrência de eventos poucos prováveis, não sendo capazes, muitas vezes, de distinguir probabilidades pequenas de probabilidades ínfimas, precificando-as de maneira semelhante.

Dessa maneira, utilizando uma extensa e confiável (fonte: Jockey oficial) base de dados de 6,4 milhões corridas de cavalo realizadas nos EUA entre 1992 e 2001, os autores buscam avaliar a aplicação de ambas as correntes. Os predicados da base de dados são de extrema importância, pois muitos outros estudos e conclusões já haviam sido realizados neste ramo, porém com amostras significativamente menores.

Por fim, sob a hipótese de que todos os apostadores possuem as mesmas preferências, os autores construíram dois modelos de utilidade: um para a classe “Risk Love” e outro para a “Misperceptions”. Como conclusão, assim como já era esperado, constatou-se que o viés de superestimação dos “azarões” (odds altas) de fato existia. Além disso, os resultados empíricos foram mais consistentes com o modelo que pregava a ideia de que o fato gerador deste viés eram as percepções imprecisas por parte dos apostadores quando os eventos possuem probabilidades extremamente baixas de ocorrer.

O terceiro e último trabalho que será discorrido com certa profundidade é DiFilippo et al.(2014). Nesta obra, os autores buscam averiguar se o mercado de apostas

esportivas, mais especificamente da NFL (National Football League), pode ser considerado eficiente.

A motivação primordial para o desenvolvimento da pesquisa se deu por fatores de extrema simplicidade. Segundo opiniões e afirmações de analistas e jornalistas esportivos, durante o início de temporada os times costumam apresentar, em geral, uma característica bastante específica: o melhor desempenho da defesa em comparação com o ataque.

Este pensamento é balizado pela ideia de que para o poder ofensivo de um time se tornar verdadeiramente eficiente é necessário mais prática e treinamento do que o nível necessário para que se obtenha um bom desempenho defensivo, o qual é mais influenciado pelo instinto e condições físicas dos atletas. Dessa maneira, os pesquisadores foram às estatísticas, constatando que isso de fato ocorria e levantando a seguinte questão: será que o mercado precifica corretamente esta “anomalia sazonal”?

Utilizando como base de dados todas as partidas ocorridas entre as temporadas dos anos 2000 e 2010, os autores buscaram examinar uma possível existência de ineficiência neste mercado. Em caso positivo, significaria que as casas de apostas não incorporam toda informação disponível na formação das suas odds, assim, como tipicamente os ataques são superados pelas defesas em inícios de temporada, menos pontos, em geral, são feitos por partida, abrindo-se uma oportunidade de se ter um lucro sistemático ao se apostar que a pontuação total dos jogos será menor do que determinado número de pontos.

Com a concretização do exercício empírico para testar a hipótese de eficiência, foi possível concluir que realizar apostas seguindo a estratégia descrita no parágrafo acima levaria, em média, a um retorno estatisticamente significativo de 13,6% por jogo.

Portanto, se torna nítido o fato de que as casas de apostas não incorporam da maneira correta toda a informação disponível acerca dos jogos, precificando suas cotações de maneira errônea e, então, gerando uma ineficiência neste mercado, da qual os apostadores podem se aproveitar para adquirir lucros de maneira consistente.

Adentrando agora na segunda e última parte da dicotomia a qual constrói o presente capítulo, a revisão bibliográfica continuará sendo desenvolvida, porém de maneira mais

sucinta, dando uma ênfase maior somente na ideia central e conclusões de cada trabalho.

Analisando o mercado de apostas de corrida de cavalo, McGlothlin (1956) e Asch, Malkiel & Quandt (1982) demonstram que o nível de retorno ao apostar no cavalo azarão cai na última corrida do dia. Tal resultado foi interpretado da seguinte maneira: como boa parte dos apostadores chega ao final do dia acumulando prejuízos, encara-se este momento como a última oportunidade de recuperação, levando a uma superestimação dos eventos que poderiam propiciar essa reabilitação financeira, isto é, dos eventos com odds altas (vitória dos cavalos menos favoritos). Anos depois, ainda explorando este mercado, Thaler & Ziemba (1988) conclui que há oportunidades de lucro em estratégias simples, como sempre apostar no favorito.

Imergindo no mercado de apostas futebolísticas, Forrest & Simons (2008) identifica que as odds praticadas em jogos da primeira divisão da Liga Espanhola são, em alguns momentos, enviesadas. Segundo os autores, este viés é gerado, em alguma instância, pela existência de fatores “não-financeiros” (“sentimentais”) nas preferências dos apostadores. O fato constatado é que o tamanho das torcidas parece influenciar nos níveis das odds praticadas, pois um time com muitos torcedores acaba sendo alvo de apostas excessivas em seu favor, levando a uma diminuição das suas odds e uma conseqüente maior expectativa de retorno relativa ao se apostar em clubes de menor expressão popular.

Por último, vale mencionar Wever & Aadland (2012). No intuito de averiguar a existência de estratégias de apostas consistentemente rentáveis, os autores focam seu estudo no mercado de apostas da NFL (National League Soccer) e chegam à conclusão que ao se apostar nos times largamente menos favoritos, gera-se, em média, um lucro estatisticamente significativo.

Uma explicação comum para este fenômeno é o conceito de “comportamento em manada”. Segundo essa ideia, os agentes não baseiam seus investimentos somente nas suas previsões de retorno futuro (o que seria essencialmente racional), mas também em ações e opiniões de terceiros. Dessa maneira, como os maiores times da liga naturalmente recebem mais holofotes por parte da mídia esportiva, os apostadores tendem a superestimá-los, o que abre espaço para a possibilidade de realização consistente de lucro ao se apostar nos times largamente “azarões”.

### 3 Base de Dados

A única fonte de dados utilizada para a confecção da monografia foi provida pelo site “Kaggle”. Esta plataforma foi fundada no ano de 2010 com o intuito de auxiliar e promover competições entre diferentes modelos de previsão através da disponibilização de bases de dados para o público. Tal disponibilização ocorre, na maioria das vezes, por meio da postagem de companhias e pesquisadores, formando, portanto, uma espécie de “Crowdsourcing”.

Dessa maneira, foi possível, sem maiores dificuldades, baixar do site (e posteriormente exportar para o excel) uma base de dados descrita como "The ultimate Soccer database for data analysis and machine learning", na qual estão contidas as odds históricas praticadas por três das maiores casas de apostas esportivas do mundo – Bet365, Bwin e Interwetten.

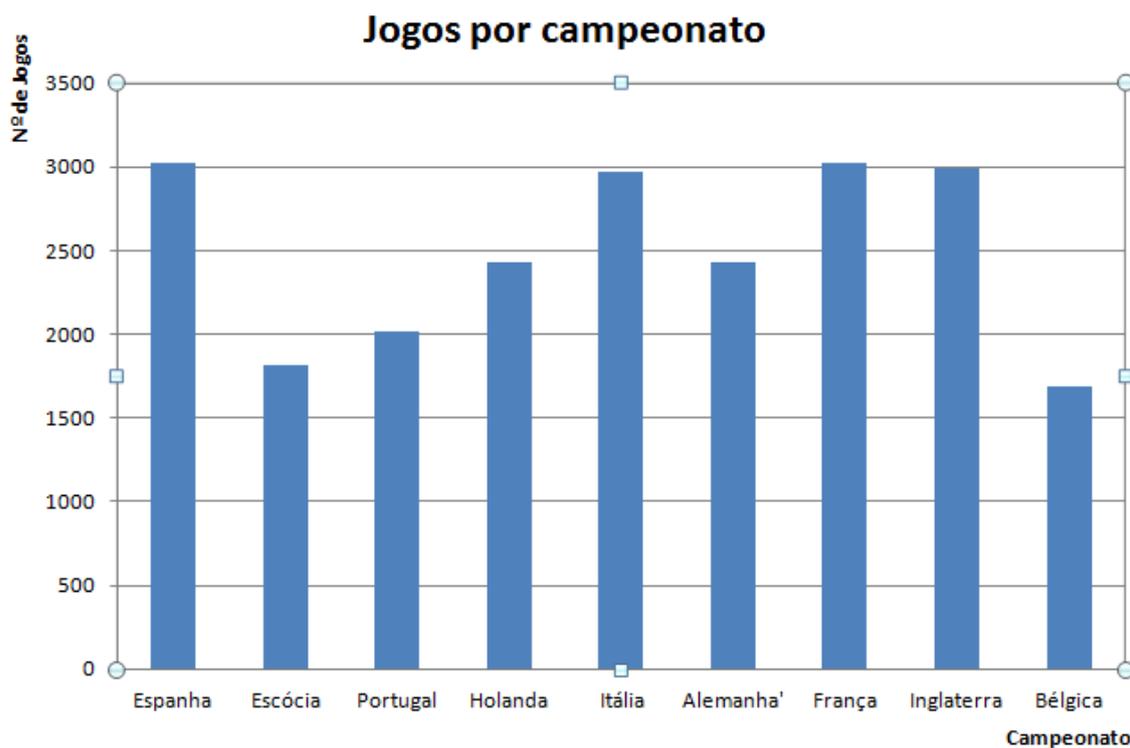
Mais especificamente, o arquivo continha as odds dos jogos realizados em nove das principais ligas nacionais europeias entre as temporadas de 2008 e 2016. Vale lembrar que a presente monografia tem como foco o estudo das odds apenas dos três eventos mais importantes dentro de uma partida de futebol: vitória do time mandante, empate ou vitória do time visitante.

De maneira geral, as ligas são disputadas em formato de “pontos corridos”, havendo turno, retorno e, aproximadamente, vinte times. Com isso, a princípio, a base de dados possuía mais de 25.000 jogos, porém uma pequena parte destes não foi oferecida pelas três casas de apostas simultaneamente (em sua maioria, partidas de campeonatos de menos destaque, tais como o Belga e o Escocês).

Assim, a fim de manter a simetria do exercício e partindo do pressuposto de que não haveria nenhuma perda amostral relevante, optou-se por excluir estes jogos da base de dados, alcançando um total de 22.397 partidas, todas inteiramente cobertas pelos três sites de apostas esportivas.

Como cada jogo tem três resultados possíveis e o presente trabalho abrangerá a cobertura de três casas de apostas, chega-se, por fim, à impactante amostra de 201.573 odds a serem estudadas.

Dessa maneira, após realizar os procedimentos de baixar, exportar e “limpar” a base de dados, esta acabou se tornando nada mais do que uma planilha no Excel que contém os jogos e seus resultados, além das três odds praticadas para cada evento possível dentro do jogo por cada uma das três casas de apostas. Assim estruturados, os dados estavam prontos para serem submetidos ao exercício principal da monografia, o qual será desbravado em mais detalhes nos capítulos posteriores.



## 4 Metodologia

Todo trabalho que possui como objetivo principal realizar um exercício prático necessita de uma boa metodologia e uma eficiente estratégia empírica. Na presente monografia, o cerne da metodologia utilizada se pauta inteiramente em um assunto bastante recorrente no curso de economia: a Teoria da Utilidade Esperada (ou Teoria do Valor Esperado).

A Teoria da Utilidade Esperada é muito utilizada em diferentes segmentos da ciência econômica, tais como microeconomia, estatística e finanças. Dessa maneira, pouco sentido faria explicar a metodologia que será posta em prática sem antes destrinchar, mesmo que de forma não tão profunda, os conceitos e implicações desta Teoria.

Como o mundo é extremamente complexo e, muitas vezes, imprevisível, não é de se estranhar que com uma frequência razoavelmente grande os indivíduos se deparem com situações de incerteza. A Teoria do Valor Esperado se predispõe a quantificar a relação risco-retorno dos cenários possíveis gerados por essa obscuridade e, com isso, definir uma tomada de decisão ótima.

Mas como mensurar algo tão subjetivo como uma relação risco-retorno? Em termos informais, a Teoria utiliza como medida de risco a probabilidade do evento ocorrer e como medida de retorno um conceito denominado “payoff” (ou utilidade), o qual pode ser compreendido como o nível de satisfação obtido pelo indivíduo caso aquele cenário (tipicamente nomeado como “estado da natureza”) venha de fato a ocorrer. Na prática, o que acontece é uma soma ponderada dos payoffs, na qual os pesos são dados pelas probabilidades de ocorrência dos eventos.

Apesar de possuir um extenso arcabouço teórico, a Teoria da Utilidade Esperada é bastante intuitiva e parece descrever razoavelmente bem a realidade. A verdade é que grande parte dos indivíduos utiliza cotidianamente, mesmo que de maneira inconsciente, os conceitos da teoria para tomar decisões em situações de incerteza.

Um exemplo que ilustra de maneira brilhante esta ideia é decidir entre levar ou não o guarda-chuva ao sair de casa. Na cidade do Rio de Janeiro faz Sol a maior parte do ano, portanto, os cidadãos cariocas não costumam sair com seus guarda-chuvas na bolsa todos os dias. Entretanto, existem alguns fatores que ajudam a prever as chances de

chover no dia seguinte, tais como a estação do ano e, principalmente, a previsão do tempo segundo os meteorologistas.

Não sair com o guarda-chuva por assumir que não deve chover ou, caso contrário, sair com guarda-chuva por acreditar que vai chover são, claramente, tomadas de decisão que condizem com a Teoria do Valor Esperado. Em termos práticos, o que está ocorrendo é, simplesmente, uma avaliação de probabilidades entre chover ou não e tudo isso só é relevante porque o indivíduo atribui diferentes níveis de satisfação para cada cenário descrito no parágrafo abaixo.

Se chover e não estiver com o guarda-chuva, ficará encharcado. Se chover e estiver com o guarda-chuva, ficará bem. Se não chover e não estiver com o guarda-chuva, ficará ainda melhor, pois além de não se molhar, não carregou o peso do guarda-chuva na mochila. Por último, se não chover e ele estiver com o guarda chuva, ele não se molhará, porém carregou o peso à toa. O que a Teoria da Utilidade Esperada faz é sistematizar tudo em termos de payoffs e, em seguida, comparar a utilidade esperada de cada cenário. Feito isso, o cenário que levar a uma maior esperança de nível de satisfação deve ser a escolha ótima do indivíduo.

Antes de adentrar no foco principal deste capítulo, isto é, na metodologia utilizada no trabalho, cabe fazer somente mais um exemplo, dessa vez ainda mais simples, no intuito de realizar uma conexão com o mercado de apostas esportivas. Suponha que uma pessoa faça a seguinte proposta a outra qualquer: “vou jogar essa moeda para cima, caso dê cara, você me paga 10 reais, caso dê coroa, eu te pago 5 reais”. Certamente a oferta será rejeitada. Entretanto, caso faça-se uma nova oferta: “vou jogar essa moeda para cima, caso dê cara, você me paga 10 reais, caso dê coroa, eu te pago 15 reais”. A não ser que a pessoa seja extremamente avessa ao risco, parece razoável que a oferta seja aceita. Mas por quê? Novamente, por causa da Teoria da Utilidade Esperada.

Assumindo que uma moeda tem 50% de chance de cair virada para cada lado, de fato não faz sentido aceitar a primeira oferta, pois, por exemplo, caso a moeda seja jogada para cima 100 vezes, a tendência é que o indivíduo tenha que pagar 10 reais em 50 desses lançamentos e receber 5 reais nos outros 50. Assim, seu saldo líquido será de  $5 \cdot 50 + (-10) \cdot 50$ , ou seja, espera-se que ele tenha um prejuízo de 250 reais. Sob o mesmo raciocínio, na segunda oferta espera-se que ele tenha um lucro neste mesmo montante.

Mas e caso fosse feita uma terceira oferta da seguinte maneira: “vou jogar essa moeda para cima, caso dê cara, você me paga 10 reais, caso dê coroa, eu te pago 10 reais”. A princípio, não parece haver vantagem para nenhum dos lados, pois a tendência é que se perca 10 reais em metade dos lançamentos e ganhe-se o mesmo valor na mesma quantidade de vezes. Dessa maneira, o lucro esperado de ambos os envolvidos convergirá para zero no longo prazo, sendo uma aposta “justa”.

Uma vez feita toda essa longa introdução, pode-se, enfim, chegar a grande conclusão que servirá como base para a metodologia utilizada no desenvolvimento do exercício principal da monografia: segundo a Teoria da Utilidade Esperada, existe uma probabilidade implícita “justa” condizente com cada payoff. Tal probabilidade é, justamente, aquela que remete a um lucro médio igual a zero no longo prazo.

O mercado de apostas esportivas possui um funcionamento análogo ao exemplo do lançamento de moedas. Para cada evento possível dentro do jogo, a casa de apostas oferece uma remuneração aos apostadores. Esta remuneração é ilustrada através das “odds”, que, como explicado no capítulo de introdução, podem ser entendidas como o quanto o site está disposto a pagar por cada real investido caso o apostador acerte seu palpite. Ou seja, a odd é, simplesmente, o payoff em caso de acerto. Vale lembrar que caso o apostador erre seu palpite, ele perderá o total do dinheiro investido, isto é, terá um payoff igual a zero.

Uma vez entendidos o funcionamento dos sites e a ideia de probabilidade e payoffs justos à luz da Teoria da Utilidade Esperada, o exercício central da monografia pode ser facilmente compreendido. Tal exercício consiste, basicamente, em analisar se as casas de apostas oferecem, em média, odds superestimadas (remuneram mais do que deveriam), subestimadas (remuneram menos do que deveriam) ou justas (remuneram exatamente o que deveriam). Tudo isso está embutido nas seguintes equações:

$$odd \times P_1 + 0 \times (1 - P_1) = 1 \quad (1)$$

$$P_1 = \frac{1}{odd} \quad (2)$$

Conforme se pode observar, as equações acima são extremamente simples. A variável  $P1$  é a probabilidade de o apostador acertar seu palpite; a odd é o payoff (o quanto ele receberá) em caso de acerto;  $(1-P1)$  é a probabilidade do apostador errar seu palpite e zero é o seu payoff neste caso. Do outro lado da equação, o número 1 representa o seu real investido.

Segundo o manual da Teoria da Utilidade Esperada, o apostador só deve estar disposto a investir seu dinheiro caso o lado esquerdo da equação seja maior do que o lado direito, ou seja, caso a sua expectativa de retorno supere o seu desembolso de 1 real. Em caso de igualdade entre os lados, o seu lucro esperado é igual a zero e, portanto, ele se deve manter indiferente entre entrar ou não na aposta.

Por algebrismo básico, chega-se na relação mais importante desta monografia, a qual está explicitada pela equação (2). Para que o lucro esperado seja igual a zero, a probabilidade de acerto deve ser exatamente o inverso da odd oferecida e vice-versa. A partir desta paridade é possível, enfim, colocar em prática o exercício empírico.

Uma vez coletadas as odds praticadas pelos sites em jogos passados, estas serão subdivididas em diferentes amostras de acordo com os seus valores. Dado que, obviamente, os jogos presentes na base de dados já têm o seu resultado definido, será feita a comparação entre a real frequência de ocorrência dos eventos com a frequência que seria condizente com a probabilidade justa de sua respectiva odd.

Para finalizar, apenas a título de ilustração, serão, por exemplo, selecionados diversos jogos onde foram ofertadas odds iguais a 2. Pela equação (2), sabe-se que a probabilidade justa é o inverso da odd, sendo, então, 0,5. Ou seja, dado que a odd oferecida é 2, a probabilidade que leva o lucro esperado do apostador para zero é 50%.

Assim, como os jogos já ocorreram, basta verificar com qual frequência os eventos de fato aconteceram. Se a frequência foi maior do que 50%, a casa de apostas superestimou a odd, ou seja, ofereceu uma remuneração maior do que deveria. Se foi igual a 50%, a odd oferecida foi justa e, por último, se foi menor do que 50%, houve uma subestimação da odd, isto é, a recompensa ofertada foi menor do que deveria.

Por último, utilizando a distribuição de probabilidade binomial, a relevância estatística dos resultados será estudada através de testes de hipótese, nos quais a hipótese nula colocada em prova é de que as odds ofertadas são estatisticamente iguais

as odds justas. Caso a hipótese nula seja rejeitada, será possível concluir que o lucro esperado no longo prazo é diferente de zero, existindo, portanto, um padrão sistemático de lucro ou prejuízo neste mercado.

## 5 Resultados

Uma vez concretizado o exercício empírico, chega-se, enfim, na parte mais esperada e instigante de todo o trabalho: os resultados. Além dos predicados citados, o capítulo de resultados também é o mais extenso de toda a monografia e, por isso, será subdividido em três seções.

A primeira seção tem como objetivo responder a pergunta principal da monografia de maneira geral, sendo, então, a mais importante do capítulo. A segunda seção refaz o exercício realizado na primeira, só que, dessa vez, “controlando” para o “spread” implícito que as casas de apostas embutem em suas cotações.

A terceira e última seção é uma espécie de exercício extra e trata, também, de estudar a relação risco-retorno expressa pelas odds, porém com um olhar mais específico. A motivação para tal é a ideia de que, em virtude da demanda dos apostadores e da atuação da própria casa de aposta, as odds tendem a ser desconexas com a real probabilidade de ocorrência do evento em certas situações. Assim, será analisada uma situação que, intuitivamente, apresenta esta potencial fonte de viés.

A situação em si é uma investigação específica nos jogos de times considerados gigantes. A premissa neste caso é que times extremamente grandes possuem uma cobertura muito intensa da mídia, o que pode levar a uma superdimensão de suas notícias (sendo positivas ou negativas) perante aos agentes e, conseqüentemente, a decisões equivocadas por parte da casa e dos investidores.

## 5.1 Resultado Geral

Para obter uma análise geral da relação risco-retorno expressa pelas odds, foi utilizada a base de dados como um todo, sendo feita apenas uma separação entre faixas de odds no intuito de averiguar se há diferenças na análise entre eventos menos ou mais prováveis. As odds foram divididas de maneira arbitrária em 5 intervalos:

. 1 - 1,2

. 1,2 - 2

. 2 – 5

. 5 – 10

. 10 em diante

A fim de facilitar a clara visualização e entendimento dos resultados, os mesmos serão ilustrados por meio de tabelas. As 3 colunas representam cada site de aposta esportiva: Bet365, Bwin e Interwetten, respectivamente. Em seguida, as tabelas e seus parâmetros serão discutidos intuitivamente.

Antes de prosseguir, vale ressaltar que há uma dupla direção na metodologia. Nas tabelas I a V, o caminho foi inferir as odds justas através da probabilidade dos eventos (utilizando como variável proxy a real frequência dos mesmos); já para as tabelas VI a X, partiu-se das odds praticadas para estimar a probabilidade que seria condizente com o que foi pago.

**Tabela I**

<b>ODD 1 – 1,2</b>	<b>B365</b>	<b>BWA</b>	<b>IWA</b>
<b>Nº Sucesso</b>	807	807	807
<b>Nº Jogos</b>	923	923	923
<b>Frequência</b>	87,43%	87,43%	87,43%
<b>Odd Justa</b>	1,1437	1,1437	1,1437
<b>Odd Praticada</b>	1,1552	1,1539	1,1623

**Tabela II**

<b>ODD 1,2 – 2</b>	<b>B365</b>	<b>BWA</b>	<b>IWA</b>
<b>Nº Sucesso</b>	7160	7160	7160
<b>Nº Jogos</b>	11976	11976	11976
<b>Frequência</b>	59,78%	59,78%	59,78%
<b>Odd Justa</b>	1,6726	1,6726	1,6726
<b>Odd Praticada</b>	1,6498	1,6416	1,6493

**Tabela III**

<b>ODD 2 -5</b>	<b>B365</b>	<b>BWA</b>	<b>IWA</b>
<b>Nº Sucesso</b>	13372	13372	13372
<b>Nº Jogos</b>	44965	44965	44965
<b>Frequência</b>	29,73%	29,73%	29,73%
<b>Odd Justa</b>	3,3626	3,3626	3,3626
<b>Odd Praticada</b>	3,3096	3,2368	3,1451

**Tabela IV**

<b>ODD 5 – 10</b>	<b>B365</b>	<b>BWA</b>	<b>IWA</b>
<b>Nº Sucesso</b>	960	960	960
<b>Nº Jogos</b>	7294	7294	7294
<b>Frequência</b>	13,16%	13,16%	13,16%
<b>Odd Justa</b>	7,5979	7,5979	7,5979
<b>Odd Praticada</b>	6,8875	6,4833	5,9411

**Tabela V**

<b>ODD 10 +</b>	<b>B365</b>	<b>BWA</b>	<b>IWA</b>
<b>Nº Sucesso</b>	101	101	101
<b>Nº Jogos</b>	2033	2033	2033
<b>Frequência</b>	4,96%	4,96%	4,96%
<b>Odd Justa</b>	20,1287	20,1287	20,1287
<b>Odd Praticada</b>	14,6740	12,8490	11,4916

Para a Tabela I, foram coletados apenas jogos onde estavam sendo ofertados eventos com odds dentro do intervalo de 1 a 1,2, chegando ao total de 923 observações. Dentro destas, 807 vezes o evento em questão veio, de fato, a ocorrer. Assim, a frequência de ocorrência foi de 87,43% e, portanto, a odd justa seria, em média, de 1,1437.

Surpreendentemente, a última linha da tabela evidencia que as 3 casas de apostas remuneraram, em média, mais do que esse valor, indicando que houve uma superestimação das odds e, portanto, uma boa oportunidade de lucro para os apostadores.

De maneira curiosa, esta foi a única amostra na qual isso ocorreu. Uma possível explicação para tal fenômeno é um tema que já foi abordado por trabalhos citados no capítulo de revisão bibliográfica: indivíduos costumam ter dificuldades de internalizar corretamente eventos com probabilidades muito extremas (nesse caso, os eventos tinham probabilidades muito altas de ocorrerem).

Na faixa de odds de 1,2 a 2 (Tabela II), a remuneração oferecida girou sempre em torno de 1,64, enquanto que a odd justa seria de, aproximadamente, 1,67. Utilizando o mesmo raciocínio, observa-se que no intervalo de odds de 2 a 5 (Tabela III) as casas de apostas pagaram de 3,14 (IWA) a 3,30 (B365), e o justo seria 3,36. Nas odds de 5 a 10 (Tabela IV), foi pago de 5,94 (B365) até 6,88 (IWA), e a odd compatível com a real frequência de acontecimento seria de 7,59. Por último, na análise de odds maiores do que 10 (Tabela V), foi remunerado de 11,49 (IWA) a 14,67 (B365), enquanto a odd justa seria de 20,12.

Dessa maneira, todos os sites, mesmo que em escalas diferentes, tendem a pagar menos do que o valor justo para todas as faixas de odds acima de 1,2. Além disso, parece ter sido constatado um interessante padrão: quanto mais improvável é o evento (ou seja, quanto maior é a faixa de odd em questão), mais distante a odd praticada tende a ser da odd justa. Isto é, há uma subestimação bastante intensa dos eventos muito improváveis, como, por exemplo, a vitória dos times considerados “zebras” dentro do confronto.

Apesar de ser uma conclusão em direção oposta à feita em relação à Tabela I, uma possível explicação para tal fenômeno é, novamente, o fato de pessoas costumarem ter dificuldades de internalizar corretamente eventos com probabilidades muito extremas (nesse caso, os eventos tinham probabilidades muito baixas de ocorrerem). Mas até que ponto todos estes resultados são, de fato, confiáveis e relevantes?

Como mencionado no final do capítulo de Metodologia, a significância estatística dos resultados será estudada por meio de testes de hipóteses nos quais os parâmetros advêm da distribuição de probabilidade binomial. As hipóteses nulas que serão testadas são, em última instância, que as odds praticadas são estatisticamente iguais as odds justas.

Mais uma vez por uma questão estética, os resultados dos testes serão exibidos em tabelas, e estas podem ser entendidas como uma extensão das tabelas já ilustradas há

pouco, pois cada uma das já ilustradas terá uma nova tabela correspondente que analisa suas informações em termos estatísticos. Assim, a tabela VI é uma extensão da tabela I, a VII é uma extensão da II e assim por diante.

**Tabela VI**

<b>ODD 1 – 1,2</b>	<b>B365</b>	<b>BWA</b>	<b>IWA</b>
<b>Nº Sucesso</b>	807	807	807
<b>Prob. Condizente</b>	86,56%	86,66%	86,03%
<b>com o pago</b>			
<b>Valor Esperado</b>	798,99	799,88	794,09
<b>Desvio Padrão</b>	10,36	10,32	10,53
<b>Teste</b>	0,77	0,68	1,22

**Tabela VII**

<b>ODD 1,2 - 2</b>	<b>B365</b>	<b>BWA</b>	<b>IWA</b>
<b>Nº Sucesso</b>	7160	7160	7160
<b>Prob. Condizente</b>	60,61%	60,91%	60,62%
<b>com o pago</b>			
<b>Valor Esperado</b>	7258,94	7295,01	7260,84
<b>Desvio Padrão</b>	53,47	53,39	53,46
<b>Teste</b>	-1,85	-2,52	-1,88

**Tabela VIII**

<b>ODD 2 - 5</b>	<b>B365</b>	<b>BWA</b>	<b>IWA</b>
<b>Nº Sucesso</b>	13372	13372	13372
<b>Prob. Condizente com o pago</b>	30,21%	30,89%	31,79%
<b>Valor Esperado</b>	13586,16	13891,37	14296,77
<b>Desvio Padrão</b>	97,37	97,97	98,74
<b>Teste</b>	-2,19	-5,30	-9,36

**Tabela IX**

<b>ODD 5 - 10</b>	<b>B365</b>	<b>BWA</b>	<b>IWA</b>
<b>Nº Sucesso</b>	960	960	960
<b>Prob. Condizente com o pago</b>	14,51%	15,42%	16,83%
<b>Valor Esperado</b>	1059	1125,03	1227,71
<b>Desvio Padrão</b>	30,08	30,84	31,95
<b>Teste</b>	-3,29	-5,35	-8,37

**Tabela X**

<b>ODD 10 +</b>	<b>B365</b>	<b>BWA</b>	<b>IWA</b>
<b>Nº Sucesso</b>	101	101	101
<b>Prob. Condizente com o pago</b>	6,81%	7,78%	8,70%
<b>Valor Esperado</b>	138,54	158,22	176,91
<b>Desvio Padrão</b>	11,36	12,07	12,70
<b>Teste</b>	-3,30	-4,73	-5,97

Intuitivamente, o teste de hipótese verifica se o valor esperado (que nada mais é do que o número de sucessos esperado) é estatisticamente igual ao número de sucessos que de fato ocorreu. O resultado do teste exhibe a diferença entre esses dois parâmetros em unidades de desvio padrão.

Analisando a tabela VI, é possível observar que o número de sucessos que de fato ocorreram é sempre maior do que o esperado. Vale lembrar que tal fato é harmônico com a constatação obtida previamente pela Tabela I de que as casas de apostas praticaram odds acima do nível justo. Porém, observando os resultados dos testes, em nenhum site essa diferença é estatisticamente significativa aos níveis de confiança convencionais (90% e 95%). Dessa maneira, a hipótese nula não pode ser rejeitada, ou seja, não se pode concluir que o que foi pago é estatisticamente maior do que deveria ter sido e, conseqüentemente, a oportunidade de lucro não parece ser tão consistente.

Acerca da tabela VII, nota-se que o número de sucessos ocorridos foi menor em 1,85, 2,52 e 1,88 desvios padrão do que o seu valor esperado para a Bet365, Bwin e Interwetten, respectivamente. Dessa maneira, a hipótese nula é rejeitada ao nível de confiança de 90% em todos os casos e, especificamente para a Bwin, também ao nível de 95%. Assim, pode-se concluir que o que foi pago é menor do que o valor justo de

maneira estatisticamente significativa e que, certamente, a maioria dos apostadores incorreu em prejuízos ao apostar sistematicamente nestas odds.

Nas demais faixas de odds, os valores dos testes passam a ser ainda mais negativos, permitindo a rejeição da hipótese nula com um nível de confiança ainda maior (muitas vezes tendendo a 100%) e tornando a conclusão do parágrafo anterior ainda mais válida e intensa.

Por último, vale ressaltar que, à exceção da Bet365, o padrão constatado no início desta seção (quanto mais improvável é o evento, mais subestimada é sua odd) não aparece de maneira nítida ao observarmos as tabelas VIII, IX e X. Para estarem de acordo com o padrão, os valores dos testes deveriam ser mais negativos à medida que as odds se tornam maiores, e não é isso que acontece.

Porém, uma explicação possível são os diferentes tamanhos de amostra utilizados. Como o número de jogos analisados é bem maior na tabela VIII, a estimação da odd justa pode ser considerada bem mais precisa e uma pequena diferença em relação a odd praticada será refletida mais fortemente no resultado do teste.

Portanto, é possível imaginar que caso as amostras possuíssem tamanhos não tão discrepantes, o padrão de subestimação de eventos improváveis tenderia a se manter claro não só para a Bet365, mas também para os demais sites.

## 5.2 Controlando para o Spread

Apesar de possuírem fontes de renda ligadas a propaganda e marketing, a maioria dos sites de apostas esportivas não cobra nenhum tipo de taxa dos apostadores pela utilização de seus serviços. Desde a abertura e manutenção de conta até a efetuação das apostas em si, normalmente não há nenhum custo explícito. Entretanto, ao realizar uma análise mais minuciosa do funcionamento dos sites e de suas odds, é possível afirmar que as casas cobram, mesmo que de maneira pouco visível, uma espécie de comissão (também conhecida como “spread”) pelo seu funcionamento.

Para chegar a tal conclusão, basta ter em mente a principal relação explorada durante a presente monografia, ou seja, de que a odd justa do evento é o inverso de sua probabilidade de ocorrência. Partindo dessa concepção e aliando-a ao simples fato de que a soma das probabilidades de vitória do time da casa, empate e vitória do time visitante deve ser igual a 1 (ou 100%, pois são os únicos resultados possíveis dentro da partida), a percepção da existência do spread se torna básica.

Na prática, o que ocorre é que a soma dos inversos das odds dentro de um mesmo jogo quase nunca é igual a 1, mas sim a um número ligeiramente superior. Matematicamente, se a soma dos inversos está dando muito alta, significa que seus denominadores (odds) estão menores do que deveriam. E é assim, oferecendo uma remuneração menor do que a justa, que os sites monetizam seus serviços.

Uma vez introduzido o mecanismo e conceito do spread, esta seção tem como objetivo repetir o exercício geral realizado na seção 5.1, só que, dessa vez, obtendo uma análise (fictícia) do que aconteceria caso as comissões não existissem. Daí nasce o desafio: como isolar o spread da cotação das odds já praticadas?

Conforme explicado, o spread é a diferença entre a soma dos inversos das odds e 1. Para checar quais seriam as odds oferecidas expurgando a comissão, basta multiplicar cada uma das odds pela soma acima citada.

Apenas a título de ilustração, segue o exemplo: se o site está oferecendo uma remuneração de 1,04, 13 e 34 para os três eventos possíveis da partida, pode-se inferir que a soma de seus inversos é aproximadamente igual a 1,067 e, portanto, que o spread nesse caso é de 0,067. Assim, as odds oferecidas, caso a comissão não existisse, seriam

1,1, 13,8 e 36,2. Para confirmar que a estratégia é válida, basta verificar que a soma dos inversos das novas odds tende a ser igual a 1.

Dessa maneira, foi possível calcular o spread médio embutido pelas três casas de apostas em suas cotações em todos os jogos da base de dados e, posteriormente, realizar a análise com as odds devidamente controladas para este efeito. O spread médio foi de, aproximadamente, 5,89%, 8,04% e 9,91% para Bet365, Bwin e Interwetten, respectivamente.

Antes de partir para o foco da seção, vale ressaltar a importância do conteúdo exposto no parágrafo acima: as odds podem variar bastante entre os sites e, em média, a Bet365, por embutir um menor spread em suas cotações, parece ser o site mais vantajoso para os apostadores.

Assim como feito na seção anterior, os resultados serão exibidos por meio de tabelas que, em seguida, serão discutidas intuitivamente. Como a lógica metodológica é a mesma e é interessante que a análise não se prolongue em excesso, acredita-se não haver problemas em, dessa vez, partir diretamente para as tabelas de cunho estatístico.

**Tabela XI**

<b>ODD 1 – 1,2</b>	<b>B365</b>	<b>BWA</b>	<b>IWA</b>
<b>Nº Sucesso</b>	807	807	807
<b>Prob. Condizente</b>	81,74%	80,21%	78,27%
<b>com o pago</b>			
<b>Valor Esperado</b>	754,55	740,35	722,49
<b>Desvio Padrão</b>	11,73	12,10	12,52
<b>Teste</b>	4,46	5,50	6,74

**Tabela XII**

<b>ODD 1,2 - 2</b>	<b>B365</b>	<b>BWA</b>	<b>IWA</b>
<b>Nº Sucesso</b>	7160	7160	7160
<b>Prob. Condizente com o pago</b>	57,24%	56,38%	55,16%
<b>Valor Esperado</b>	6855,17	6752,14	6606,17
<b>Desvio Padrão</b>	54,14	54,27	54,42
<b>Teste</b>	5,63	7,51	10,17

**Tabela XIII**

<b>ODD 2 - 5</b>	<b>B365</b>	<b>BWA</b>	<b>IWA</b>
<b>Nº Sucesso</b>	13372	13372	13372
<b>Prob. Condizente com o pago</b>	28,53%	28,59%	28,91%
<b>Valor Esperado</b>	12830,44	12857,62	13000,61
<b>Desvio Padrão</b>	95,75	95,81	96,13
<b>Teste</b>	5,65	5,36	3,86

**Tabela XIV**

<b>ODD 5 - 10</b>	<b>B365</b>	<b>BWA</b>	<b>IWA</b>
<b>Nº Sucesso</b>	960	960	960
<b>Prob. Condizente com o pago</b>	13,71%	14,27%	15,31%
<b>Valor Esperado</b>	1000,1	1041,32	1117,01
<b>Desvio Padrão</b>	29,37	29,87	30,75
<b>Teste</b>	-1,36	-2,72	-5,10

**Tabela XV**

<b>ODD 10 +</b>	<b>B365</b>	<b>BWA</b>	<b>IWA</b>
<b>Nº Sucesso</b>	101	101	101
<b>Prob. Condizente com o pago</b>	6,43%	7,20%	7,91%
<b>Valor Esperado</b>	130,83	146,44	160,95
<b>Desvio Padrão</b>	11,06	11,65	12,17
<b>Teste</b>	-2,69	-3,89	-4,92

Como era de se esperar, ao controlar para o spread, a análise passa a ser feita com odds maiores. Dessa maneira, a probabilidade condizente com o valor pago diminui e, conseqüentemente, o mesmo acontece com o valor esperado de sucessos. Por fim, ao se comparar o número real de sucessos (que, obviamente, se mantém fixo) com um menor valor esperado dessa mesma variável, chega-se a resultados maiores para todos os testes de hipótese.

Assim, todo o efeito da exclusão do spread pode ser sentido ao se comparar as tabelas VI a X com as XI a XV. Em relação ao intervalo de odds de 1 a 1,2 (Tabela XI), passa-se a rejeitar a hipótese nula a níveis de confiança que tendem a 100%, transformando o que parecia ser apenas uma oportunidade de lucro não muito concisa em uma que pode ser totalmente confiável à luz da estatística.

De maneira ainda mais intensa, o mesmo ocorre para os dois intervalos seguintes, isto é, de 1,2 até 2 e 2 até 5 (representadas pelas tabelas XII e XIII). Sem expurgar o spread da análise, nas duas amostras era possível concluir que um apostador certamente teria prejuízo no longo prazo. Ao controlar para o spread, essa conclusão se inverte totalmente: continua sendo possível rejeitar a hipótese nula, porém, desta vez, pelo fato do número de sucessos estar à extrema direita da curva normal padrão. Ou seja, é possível afirmar que o que foi pago é de maneira estatisticamente significativa maior do que odd justa.

Por último, essa reversão de tendência parece não ocorrer nos últimos dois intervalos (de 5 a 10 e de 10 em diante, representados pelas tabelas XIV e XV). Apesar da contabilização do spread atenuar o resultado dos testes, basicamente todas as hipóteses nulas continuam sendo rejeitadas ao nível de confiança de 95%, o que significa que mesmo isolando a comissão implícita do site, os apostadores teriam, em média, prejuízos consistentes ao apostar nestes jogos. Tal constatação ilustra, dentre outros fatores, a força do padrão de subestimação de odds altas, previamente constatado na seção anterior.

### 5.3 Caso específico: Análise de times gigantes

Conforme explicitado no início do capítulo, a terceira e última seção se mantém sob o prisma de estudar a relação risco-retorno expressa pelas odds, porém, dessa vez, com uma ótica mais pontual. Visando investigar uma situação na qual intuitivamente as odds têm maior probabilidade de serem desajustadas em relação a real probabilidade de ocorrência do evento, serão explorados apenas jogos que envolvem times reconhecidamente gigantes.

Os pontos de partida para acreditar que essa situação específica apresenta uma potencial fonte de viés são bem simples. Primeiro: sabe-se que no mundo futebolístico, especialmente no continente europeu, a cobertura da mídia ocorre de maneira muito desigual. Segundo: sabe-se, também, que a mídia é, naturalmente, uma grande formadora de opinião.

Assim, a premissa é que times extremamente grandes possuem uma cobertura midiática muito intensa, o que pode levar a uma superdimensão de suas notícias (positiva ou negativamente) perante aos agentes e, conseqüentemente, a decisões equivocadas por parte das casas de apostas e dos investidores.

Dessa maneira, para a realização do exercício foram selecionados arbitrariamente três dos maiores times do mundo: Real Madrid, Barcelona e Bayern de Munique. Por reduzir drasticamente a base de dados, diferentemente do que foi feito anteriormente, nesta seção não haverá uma divisão entre níveis de odds. Toda a subamostra será sintetizada de maneira única.

Como de praxe, os resultados e a sua relevância estatística serão exibidos em forma de tabela. As intuições serão discutidas na sequência.

**Tabela XVI**

<b>Times Gigantes</b>	<b>B365</b>	<b>BWA</b>	<b>IWA</b>
<b>Nº Sucesso</b>	643	643	643
<b>Nº Jogos</b>	858	858	858
<b>Frequência</b>	74,94%	74,94%	74,94%
<b>Odd Justa</b>	1,3343	1,3343	1,3343
<b>Odd Praticada</b>	1,3954	1,3773	1,3950

**Tabela XVII**

<b>Times Grandes</b>	<b>B365</b>	<b>BWA</b>	<b>IWA</b>
<b>Nº Sucesso</b>	643	643	643
<b>Prob. Condizente com o pago</b>	71,66%	72,60%	71,68%
<b>Valor Esperado</b>	614,85	622,93	615,04
<b>Desvio Padrão</b>	13,20	13,06	13,19
<b>Teste</b>	2,13	1,53	2,11

Como é possível observar pela tabela XVI, Real Madrid, Barcelona e Bayern de Munique entraram em campo 858 vezes e saíram vitoriosos em 643 oportunidades, ou seja, em 74,94% dos jogos. Assim, pode-se inferir que a odd justa seria, em média, de 1,3343 e, se atentando para a última linha da tabela, nota-se que todos os sites ofereceram remunerações mais elevadas do que esse patamar. Ou seja, assim como ocorreu na primeira faixa de odd (de 1 a 1,2) durante o exercício geral (seção 5.1), houve uma superestimação das odds.

Entretanto, a tabela XVII ilustra uma imprescindível diferença: dessa vez (e pela primeira vez em toda a monografia) essa superestimação parece ter ocorrido de maneira estatisticamente significativa. Pelos resultados dos testes, é possível afirmar com 95% de confiança que a Bet365 e a Interwetten praticaram odds substancialmente maiores do que a justa.

Dessa maneira, pode-se concluir que existiu uma oportunidade concisa e consistente de realização de lucros. Em média, os investidores que apostaram recorrentemente nas vitórias dos três times em questão obtiveram um retorno financeiro positivo.

Uma possível explicação para essa ocorrência pode se pautar justamente na premissa intuitiva da seção: os investidores e as casas de apostas podem, devido à superdimensão da cobertura da mídia, dar um peso excessivo a notícias negativas relacionadas aos três clubes, levando a uma subestimação das suas probabilidades de vitória e, conseqüentemente, a uma superestimação da cotação das odds de tais eventos.

## 6 Conclusão

Pautando-se inteiramente em uma metodologia que deriva da Teoria da Utilidade Esperada e utilizando uma base de dados com mais de 200.000 odds, acredita-se que o presente trabalho foi capaz de prover uma análise profunda da relação risco-retorno dos sites de apostas esportivas.

Uma vez realizados os exercícios empíricos, ficou evidente que não é nada fácil auferir lucros de maneira consistente neste mercado. De modo geral, foi possível concluir que, em média, as casas de apostas costumam praticar níveis de odds inferiores àqueles considerados justos de maneira estatisticamente significativa.

A principal causa dessa evidência é a existência da cobrança de spreads pelos sites. Como consequência desse custo implícito nas cotações, a expectativa de retorno dos apostadores no longo prazo é, na grande maioria das vezes, menor do que zero. Assim, naturalmente espera-se uma recorrência bem maior na ocorrência de prejuízos do que de ganhos.

Além disso, essa recorrência deve ser observada com maior intensidade nos investidores que buscam apostar em eventos com pequenas probabilidades de ocorrer, pois foi apurado que, à medida que as odds crescem, mais subestimadas elas tendem a ser. Em outras palavras, esse padrão de comportamento significa que quanto mais improvável é a ocorrência do evento, mais subvalorizadas são as odds praticadas em relação às odds que levariam a expectativa de retorno para zero no longo prazo.

Entretanto, foi possível concluir também que, apesar de todo esse aparente cenário negativo para os apostadores, existem situações específicas em que a possibilidade de lucro é plenamente sólida e estatisticamente significativa. Como verificado na análise dos jogos de times gigantes, caso o investidor possua boas intuições acerca do esporte, é possível delimitar estratégias para se apostar sempre em eventos nos quais as odds costumam ser superestimadas e, dessa maneira, realizar bons investimentos neste mercado.

## 7 Referências Bibliográficas

AIROLDI, Luca. **Statistical Data Analytics Of Football For Beating The Odds**. Tese de Doutorado. Dublin, National College of Ireland, 2014.

SNOWBERG, Erik; WOLFERS, Justin. **Explaining the favorite-longshot bias: Is it risk-love or misperceptions?** National Bureau of Economic Research. Working Paper 15923, 2010.

DIFILIPPO, Michael et al. **Early Season NFL Over/Under Bias**. Journal of Sports Economics, p. 1527002512454544, 2012.

ASCH, Peter; MALKIEL, Burton G.; QUANDT, Richard E. **Racetrack betting and informed behavior**. Journal of financial economics, v. 10, n. 2, p. 187-194, 1982.

MCGLOTHLIN, William. **Stability of choices among uncertain alternatives**. The American journal of psychology, v. 69, n. 4, p. 604-615, 1956.

THALER, Richard; ZIEMBA, William. Anomalies: **Parimutuel betting markets: Racetracks and lotteries**. The Journal of Economic Perspectives, v. 2, n. 2, p. 161-174, 1988.

FORREST, David; SIMMONS, Robert. **Sentiment in the betting market on Spanish football**. Applied Economics, v. 40, n. 1, p. 119-126, 2008.

WEVER, Sean; AADLAND, David. **Herd behaviour and underdogs in the NFL**. Applied Economics Letters, v. 19, n. 1, p. 93-97, 2012.