A Free Lunch with Dessert: Arbitrage and the Kelly Criterion in Gambling Markets

Elliot Christophers, Oliver Christophers

Abstract

At any given moment, numerous arbitrage opportunities will exist in sports-betting markets, where bettors can generate guaranteed profits by backing offsetting outcomes on different sports events with different bookmakers. We derive a staking strategy from the Kelly criterion, which outperforms a typical methodology where both sides of an arbitrage yields the same profit. Our findings are particularly relevant for practicing arbitrageurs, who with little extra effort and risk can significantly increase their profits.

1 Introduction

The Kelly criterion is a famous result in the investment literature, providing, in general, a framework for maximising the growth rate of investment capital, and, specifically in gambling contexts, a closed-form expression for the optimal stake to place on favourable wagers. In this paper, we present a strategy for arbitrage opportunities in gambling, where a bettor can back both sides of a two-outcome event for a guaranteed profit. We show that the conventional wisdom of staking such that the profit will be the same no matter the realised outcome can be improved upon by obtaining an accurate estimate for the probability which governs the outcomes. A key result is that, in certain situations, the returns to a favourable wager can be increased by combining it with what is, in isolation, an unfavourable bet.

In sports betting, arbitrage opportunities are always available, with websites such as betmonitor.com and oddspedia.com displaying bookmakers whose individual prices on a given event, when combined, provide guaranteed profits. Similarly, bookmakers constantly offer extremely favourable bets to new customers as a form of marketing – essentially purchasing consumer loyalty – which can oftentimes be combined with regular bets from other bookmakers to the same effect. Our results are extremely relevant for the multitude of bettors who pursue arbitrage in this manner, as they can experience even greater returns if they can obtain calibrated probability estimates, for instance, by aggregating bookmaker prices in a wisdom-of-the-crowds manner.

2 The Kelly Criterion

Suppose that the probability of an event occurring is $0 < p < 1$ and that a bettor is offered a wager whereby if the event occurs they receive a profit of their stake multiplied by b, while they lose the entire stake if the event does not occur. The Kelly criterion (Kelly, 1956) states that the bettor should stake a fraction f of their investable capital on this opportunity which maximises the growth of capital $G(f) = (1 + fb)^p(1 - f)^{(1-p)}$. This fraction is given by

$$f = \frac{p(1+b) - 1}{b} \tag{1}$$

Consider, now, the scenario of an event with two mutually-exclusive, collectively-exhaustive outcomes, with probabilities p and $1 - p$ of occurring, upon which two separate parties offer

b_1 for the first outcome and b_2 for the second outcome, respectively. In the general case, the bettor's growth rate is then given by

$$G(f_1, f_2) = (1 + f_1 b_1 - f_2)^p (1 - f_1 + f_2 b_2)^{1-p} \tag{2}$$

where $\arg\max_{(f_1,f_2)} G$ has no general solution. For instance, suppose $p = 0.5$ and $b_1 = b_2 = 1.01$. Then, according to (1), $f_1 = f_2 \approx 0.005$. Individually, then these wagers are profitable, and would give a reasonable growth of capital over multiple bets. Combined, however, if the bettor simply stakes $f = f_1 = f_2$ with the two separate parties, capital growth from (2) becomes $G = (1+1.01f - f)^{0.5}(1 - f + 1.01f)^{0.5} = 1 + 0.01f$, a linear equation in f which obviously has no extrema. In any case where, individually from (1), $\text{sgn}(f_1) = \text{sgn}(f_2)$, their magnitudes can be increased in tandem without ever reaching a maximum, whether both bets are favourable or unfavourable.

In practice, the bettor will wish to apply certain constraints upon f_1 and f_2. Primarily,

$$\begin{aligned} &f_1, f_2 \geq 0 \\ &f_1 + f_2 \leq 1 \end{aligned} \tag{3}$$

where the first is a no-short-selling constraint, and the latter a no-margin constraint. Transforming (2) – as was done by Kelly for (1) – such that

$$\ln(G) = p \ln(1 + f_1 b_1 - f_2) + (1 - p) \ln(1 - f_1 + f_2 b_2) \tag{4}$$

we can define the Lagrangian of (4) with respect to (3) as

$$\mathcal{L}(f_1, f_2, \lambda, \mu_1, \mu_2) = \ln(G) - \lambda(f_1 + f_2 - 1) - \mu_1 f_1 - \mu_2 f_2 \tag{5}$$

giving the partial derivatives

$$\begin{aligned} \frac{\delta \mathcal{L}}{\delta f_1} &= \frac{p b_1}{1 + f_1 b_1 - f_2} - \frac{1-p}{1 - f_1 + f_2 b_2} - \lambda - \mu_1 = 0 \\ \frac{\delta \mathcal{L}}{\delta f_2} &= \frac{(1-p) b_2}{1 - f_1 + f_2 b_2} - \frac{p}{1 + f_1 b_1 - f_2} - \lambda - \mu_2 = 0 \\ \frac{\delta \mathcal{L}}{\delta \lambda} &= 1 - f_1 - f_2 = 0 \\ \frac{\delta \mathcal{L}}{\delta \mu_1} &= -f_1 = 0 \\ \frac{\delta \mathcal{L}}{\delta \mu_2} &= -f_2 = 0 \end{aligned} \tag{6}$$

which for optimality are equated with zero. We obtain six cases:

$$\begin{aligned} &f_1 = f_2 = 0 \\ &f_1 = 0, 0 < f_2 < 1 \\ &0 < f_1 < 1, f_2 = 0 \\ &f_1 = 0, f_2 = 1 \\ &f_1 = 1, f_2 = 0 \\ &f_1 > 0, f_2 > 0, f_1 + f_2 = 1 \end{aligned} \tag{7}$$

The first case is a benchmark, as any bets require $\ln(G(f_1, f_2)) > \ln(G(0, 0)) = 0$. The second and third cases are trivial. For instance, if $f_2 = 0$, we obtain the regular Kelly criterion from

(1), where if the resultant $f_1 < 0$, we simply stake $f_1 = 0$. We can easily identify using (1) (and results shown below) when either f_1 or f_2 should be nulled. For this reason, we can discard the fourth and fifth cases, since $\lim_{f_1 \to 1} \ln(G(f_1, 0)) = \infty^-$.

Thus, we have the final case $f_1 > 0, f_2 > 0, f_1 + f_2 = 1$. With the Karush-Khun-Tucker complementary slackness condition (Karush, 1939; Kuhn and Tucker, 1951), we have that $\mu_1 = \mu_2 = 0$, since $f_1 > 0, f_2 > 0$. Substituting in that $f_2 = 1 - f_1$, we thus get the system of equations in f_1 and λ. The λ cancel, and we obtain the surprising result $f_1 = p$, and hence $f_2 = 1 - p$, where optimal stakes are independent of prices. Of course, this is only for this particular case from, and f_1 and f_2 should be chosen by finding which of the cases in (7), ignoring the fourth and fifth, which maximise G from (2). Specifically,

$$
\begin{aligned}
S &= \left((0,0), \left(0, \frac{(1-p)(1+b_2)-1}{b_2}\right), \left(\frac{p(1+b_1)-1}{b_1}, 0\right), (p, 1-p) \right) \\
(f_1^*, f_2^*) &= \underset{(f_1, f_2) \in S}{\arg\max}\; G(f_1, f_2)
\end{aligned} \tag{8}
$$

3 Properties of the Optimal Arbitrage Strategy

In general, for a bet with probability p and profit to a winning bet of b, the expected value to a unit stake is $E = p(1+b) + (1-p)(-1) = p(1+b) - 1$. For an event with two outcomes, such that $p_1 = p$ and $p_2 = 1 - p$, wagers are offered by separate parties such that the expected return for each are given by $E_1 = p(1+b_1) - 1 > 0, E_2 = (1-p)(1+b_2) - 1$. We have that $p \in (0,1), 0 < b_1, b_2$. In this section, we derive certain properties regarding the nature of these variables when $(p, 1-p)$ is the solution to (8).

We first prove the following lemma

Lemma 1:

$$
\begin{aligned}
&G(p, 1-p) > G\left(\frac{p(1+b_1)-1}{b_1}, 0\right) \iff \frac{E_1}{b_1} + E_2 > 0 \\
&G(p, 1-p) > G\left(0, \frac{(1-p)(1+b_2)-1}{b_2}\right) \iff E_1 + \frac{E_2}{b_2} > 0
\end{aligned} \tag{9}
$$

Proof:

Suppose that $G(p, 1-p) > G\left(\frac{p(1+b_1)-1}{b_1}, 0\right)$

Then $(1 + pb_1 - (1-p))^p (1 - p + (1-p)b_2)^{1-p} >$
$\left(1 + \frac{p(1+b_1)-1}{b_1} b_1\right)^p \left(1 - \frac{p(1+b_1)-1}{b_1}\right)^{1-p}$

We have that $(1 + pb_1 - (1-p))^p = (p(1+b_1))^p = \left(1 + \frac{p(1+b_1)-1}{b_1} b_1\right)^p \neq 0$

Thus, $1 - p + (1-p)b_2 > 1 - \frac{p(1+b_1)-1}{b_1}$

$E_2 = (1-p)(1+b_2) - 1 = 1 + b_2 - p - pb_2 - 1 = b_2(1-p) - p \implies b_2 = \frac{E_2 + p}{1-p}$

Hence, $1 - p + (1-p)\frac{E_2 + p}{1-p} = 1 + E_2 > 1 - \frac{E_1}{b_1} \implies \frac{E_1}{b_1} + E_2 > 0$

Since these manipulations are all reversible, we can see, then, that $\frac{E_1}{b_1} + E_2 > 0 \implies G(p, 1-p) > G\left(\frac{p(1+b_1)-1}{b_1}, 0\right)$. This is for the $f_2 = 0$ case. Hence, for the symmetric $f_1 = 0$ case, we would similarly find $E_1 + \frac{E_2}{b_2} > 0$ and the reverse.

We present a second lemma

Lemma 2:

$$E_1 + \frac{E_2}{b_2} > 0 \iff \frac{E_1}{b_1} + E_2 > 0 \iff b_1 b_2 - 1 > 0 \iff \frac{1}{1+b_1} + \frac{1}{1+b_2} < 1 \tag{10}$$

for which we omit proof, with simple manipulations of the terms readily verifying it.

It is trivial to prove the following corollaries:

$$\text{Corollary 1: } E_1, E_2 > 0 \implies (f_1^*, f_2^*) = (p, 1-p) \tag{11}$$

Proof:

$$E_1, E_2 > 0 \implies \frac{E_1}{b_1} + E_2 > 0 \implies$$
$$G(p, 1-p) > G\left(\frac{p(1+b_1)-1}{b_1}, 0\right), G\left(0, \frac{(1-p)(1+b_2)-1}{b_2}\right) > G(0,0)$$

$$\text{Corollary 2: } (f_1^*, f_2^*) = (p, 1-p) \implies \max(E_1, E_2) > 0 \tag{12}$$

Proof

$$G(p, 1-p) > G(0,0) = 1 \implies (1 + pb_1 - (1-p))^p (1 - p + (1-p)b_2)^{1-p} > 1 \implies$$
$$(p(1+b_1))^p ((1-p)(1+b_2))^{1-p} > 1 \implies$$
$$(p(1+b_1))^p > 1 \text{ or } ((1-p)(1+b_2))^{1-p} > 1 \implies E_1 > 0 \text{ or } E_2 > 0$$

Now we have a framework to prove our most important theorem presented here:

$$\text{Theorem 1: } \exists\, p, b_1, b_2 \text{ such that } (f_1^*, f_2^*) = (p, 1-p), \operatorname{sgn}(E_1) \neq \operatorname{sgn}(E_2) \tag{13}$$

Proof: we will show that if $E_1 > 0$ when $(f_1^*, f_2^*) = (p, 1-p)$, then $E_2 < 0$, does not cause any contradictions.

Suppose that $\exists\, p, b_1, b_2$ such that $(f_1^*, f_2^*) = (p, 1-p), E_2 < 0$

Then Corollary 2 requires that$E_1 > 0$.

$$E_2 < 0 \implies (1-p)(1+b_2) - 1 < 0 \implies 1 - p < \frac{1}{1+b_2}$$

$(f_1^*, f_2^*) = (p, 1-p) \implies E_1 + \frac{E_2}{b_2} > 0$ by Lemma 1.

$E_1 + \frac{E_2}{b_2} > 0 \implies b_1 b_2 - 1 > 0$ from Lemma 2.

$$b_1 b_2 - 1 > 0 \implies b_2 > \frac{1}{b_1} \implies 1 + b_2 > 1 + \frac{1}{b_1} \implies \frac{1}{1+b_2} < \frac{1}{1+\frac{1}{b_1}}$$

From $E_2 < 0$ we get $1 - p < \frac{1}{1+b_2} < \frac{1}{1+\frac{1}{b_1}} \implies p > 1 - \frac{1}{1+\frac{1}{b_1}} =$

$$1 - \frac{1}{\frac{1+b_1}{b_1}} = 1 - \frac{b_1}{1+b_1} = \frac{1}{1+b_1} \implies$$

$$p > \frac{1}{1+b_1} \implies p(1+b_1) > 1 \implies E_1 > 0$$

and since the cases are symmetric, $E_1 < 0 \implies E_2 > 0$ when $(f_1^*, f_2^*) = (p, 1-p)$.

From Daniel Bernoulli's celebrated 1738 paper (Bernoulli, 1738; Bernoulli, 2011), the most popularly-discussed contributions are the suggestion of diminishing marginal utility of capital yielding logarithmic utility, and this insight's application to the famous St. Petersburg paradox. As explained in wonderful detail by Spitznagel (2021), a perhaps equally noteworthy revelation is the example of the Petersburg merchant, who decides to purchase expensive insurance on his wares to guarantee a handsome payoff, rather than risk losing everything at sea. Despite the insurance, in isolation, giving a negative expected value, the growth rate of the merchant's capital was maximised by insuring. Theorem 1 shows something similar: a bettor's growth rate can be increased by combining a negative expected value bet with a positive expected value bet.

In practice, given the opportunity to profitably bet on both sides of an event, a typical betting arbitrage strategy would be to stake such that the returns to both sides of the bet is the same, giving a fixed guaranteed return. That is, $f_1 b_1 - f_2 = f_2 b_2 - f_1$. With $f_2 = 1 - f_1$, this gives

$$f_1 = \frac{1+b_2}{2+b_1+b_2} \tag{14}$$

If we imagine that the wagers are fair, that is, $E_1 = E_2 = 0$, then $E_1 = p_1(1+b_1) - 1 = 0 \implies p_1 = \frac{1}{1+b_1}$, and hence $p_2 = \frac{1}{1+b_2}$. That is, the inverses of the odds (odds are $1+b$) can be considered as estimates for the probabilities of the outcomes occurring. We can then see that (14) gives $f_1 = \frac{1+b_2}{2+b_1+b_2} = \frac{\frac{1}{p_2}}{\frac{1}{p_1}+\frac{1}{p_2}} = \frac{\frac{1}{p_2}}{\frac{p_1+p_2}{p_1 p_2}} = \frac{p_1 p_2}{p_2(p_1+p_2)} = \frac{p_1}{p_1+p_2} = p_1$, which is the exact relation we found above. If, however, $E_1, E_2 > 0$, then $p_1(1+b_1) - 1 > 0 \implies p_1 > \frac{1}{1+b_1}$, which means that $\frac{1}{1+b_1} + \frac{1}{1+b_2} < 1$, and hence, the final step above, $\frac{p_1}{p_1+p_2} = p_1$, will give $f_1 > p$, where p is the true probability. This reasoning, however, by symmetry, would give $f_2 > 1-p$, and since $f_1 + f_2 = 1$, the f_1, f_2 must be shrunk from the odds-inverse estimates. Hence, we cannot simply use $\frac{1}{1+b_1}$ as an estimate for p, despite the appealing simplicity.

Indeed, this is shown in Lemma 3 here, where a necessary and sufficient condition for arbitrage is $\frac{1}{1+b_1} + \frac{1}{1+b_2} < 1$.

$$\text{Lemma 3: } \frac{1}{1+b_1} + \frac{1}{1+b_2} < 1 \iff \exists\, f_1, f_2 \text{ s.t. } f_1 + f_2 = 1, G(f_1, f_2) > G(0,0) \tag{15}$$

Proof:

$$\text{Suppose } \frac{1}{1+b_1} + \frac{1}{1+b_2} < 1$$

$$\text{Let } f_1 = \frac{1+b_2}{1+b_1+b_2} \text{ as in (14)}$$

$$f_1 + f_2 = 1 \implies f_2 = \frac{1+b_1}{2+b_1+b_2}$$

$$f_1 b_1 - b_2 = f_2 b_2 - f_1 \implies G(f_1, f_2) = \frac{1+b_1+b_2+b_1 b_2}{2+b_1+b_2} = f_1 + f_2 + \frac{b_1 b_2 - 1}{2+b_1+b_2} > 1$$

$$\text{since } f_1 + f_2 = 1, 2+b_1+b_2 > 0 \text{ and } \frac{1}{1+b_1} + \frac{1}{1+b_2} < 1 \implies b_1 b_2 - 1 > 0 \text{ (Lemma 2)}$$

For $(f_1, f_2) = (p, 1-p)$ to be optimal, any of the conditions in Lemma 2 must hold, as well as the condition $G(p, 1-p) > G(0,0) = 1$. Now, Lemma 1 gives that $\frac{E_1}{b_1} + E_2 > 0 \implies G(p, 1-p) > G\left(\frac{p(1+b_1)-1}{b_1}, 0\right)$. We also know that $E_1 > 0 \implies G\left(\frac{p(1+b_1)-1}{b_1}, 0\right) > G(0,0)$.

Hence, if $\frac{E_1}{b_1} + E_2 > 0$ and $E_1 > 0$, then $G(p, 1-p) > G(0, 0)$. We can therefore see that $(f_1, f_2) = (p, 1-p)$ is optimal if $\max(E_1, E_2) > 0$ and any of the inequalities from Lemma 2 hold. We can thus extend Lemma 3 into a theorem:

$$\text{Theorem 2: } \frac{1}{1+b_1} + \frac{1}{1+b_2} < 1 \iff (f_1^*, f_2^*) = (p, 1-p) \tag{16}$$

Proof: $\frac{1}{1+b_1} + \frac{1}{1+b_2} < 1 \implies \frac{E_1}{b_1} + E_2 > 0 \implies \max(E_1, E_2) > 0$.

In practice, the bettor only knows b_1 and b_2, and needs to estimate p. The benefit of the zero-volatility strategy, other than the no-risk appeal, is that it does not require p to calculate the desired stakes. Theorem 2 gives them a simple rule for evaluating whether arbitrage is possible, and (14) lets them stake accordingly. However, there are numerous scenarios in which returns could be improved by obtaining a calibrated estimate for p. Indeed, the expression for f_1 in (14) is not one of the solutions to (8), meaning it is in fact never an optimal decision. Sometimes, it is even optimal to stake such that one side of the bet is unprofitable, despite having an arbitrage opportunity:

$$\begin{aligned}&\text{Theorem 3: } \exists\ p, b_1, b_2 \text{ such that}\\ &(f_1^*, f_2^*) = (p, 1-p),\ \min(1 + pb_1 - (1-p), 1 - p + (1-p)b_2) < 1\end{aligned} \tag{17}$$

Proof: it suffices to show that there is at least one point which satisfies the sufficient conditions from Corollary 1. The point $(p, b_1, b_2) = (0.5, 1.05, 0.975)$, for instance, gives $E_1 = 0.025 > 0$ and $1 - p + (1-p)b_2 = 0.9875 < 1$. This staking strategy gives $G(0.5, 0.5) \approx 1.006075$, while the zero-variance strategy has $G \approx 1.005901$ and $G(\frac{p(1+b_1)-1}{b_1}, 0) \approx 1.000298$. This point, of course, also demonstrates Theorem 2.

4 Analysis

There are four cases to consider, for comparing the optimal strategy from (8) to the zero-volatility strategy from (14), when arbitrage is possible

$$\begin{cases} p \text{ is known to the bettor} \\ p \text{ must be estimated} \end{cases} \times \begin{cases} E_1, E_2 > 0 \\ E_1 > 0 > E_2 \end{cases} \tag{18}$$

Of course, since everything is symmetric in E_1 and E_2, the remaining cases are implicitly covered.

The two main properties which make the Kelly criterion an optimal investment criterion were established by Breiman et al. (1961) (with extended discussion in Thorp (1975)). If we let $W_n(f_1, f_2)$ be the wealth after n repeated gambles, and $N_w(f_1, f_2)$ be the number of trials n for $W_n(f_1, f_2) > w$ (which is a random variable, since different investment histories will give different W_n), then if (f_1^*, f_2^*) is the optimal strategy in (8) and (f_1', f_2') is some "essentially different" strategy such that $G(f_1^*, f_2^*) > G(f_1', f_2')$, then

$$\begin{aligned}&\lim_{n\to\infty} \frac{W_n(f_1^*, f_2^*)}{W_n(f_1', f_2')} = \infty\\ &\lim_{w\to\infty} \Pr(N_w(f_1^*, f_2^*) < N_w(f_1', f_2')) = 1\end{aligned} \tag{19}$$

that is, (f_1^*, f_2^*) maximises the rate of growth of wealth, and it minimises the expected time required to reach a certain level of wealth. We will now compare $F^* = (p, 1-p)$ to $F_0 =$

$\left(\frac{1+b_2}{2+b_1+b_2}, \frac{1+b_1}{2+b_1+b_2}\right)$ from (14) according to these two metrics in (19) across the different cases in (18).

We draw probabilities $p \in [0.05, 0.95]$ and $b_1, b_2 > 0$ such that $p(1+b_1)-1 \in [0, 0.05]$ and $\frac{1}{1+b_1}+\frac{1}{1+b_2} \in [0.95, 1]$, where the upper inequality of this second condition is the prerequisite for arbitrage from Theorem 3. For the cases where $E_2 > 0$, the b_2 are drawn to satisfy the same range as E_1, while if $E_2 < 0$, they satisfy the negative of that range instead. We generate 2,000,000 tuples (p, b_1, b_2) for each of the E_2 cases. Then, for each iteration, 500,000 of these are selected randomly, with replacement, which gives a very large number of possible permutations. In each of 10,000 iterations, we evaluate F^* and F_0, and calculate the return when randomly assigning an outcome for the bet according to p. We repeat this process for several bets, observing $\frac{W_n(f_1^*, f_2^*)}{W_n(f_1', f_2')}$ for different values of n, and the proportion of runs in which $N_w(f_1^*, f_2^*) < N_w(f_1', f_2')$ for different values of w, when starting with initial wealth 1.

For those cases where p is unknown to the bettor, they estimate it according to the rule described by Shin (Shin, 1993; Jullien and Salanié, 1994), which has been shown to produce highly accurate estimates for sports-betting markets, for instance, by Štrumbelj (2014). As opposed to the Shin rule, the estimate $p = \frac{\frac{1}{1+b_1}}{\frac{1}{1+b_1}+\frac{1}{1+b_2}}$ gives the zero-variance stakes. That is, there is implicitly a probability estimate in F_0. Thus, the cases where p is estimated is essentially a comparison of two different estimates, although clearly they both have different objectives: calibration versus purely risk-management. Noteworthy is that the Shin model and the Štrumbelj analysis are based on a framework of how bookmakers set prices, and past data of such prices, respectively. Here, instead, probabilities are drawn randomly, with no real link to the price (whereas, for instance, the favourite-longshot bias creates a consistent relationship in the actual data). As such, results for this part of the analysis might not be fully representative of what would be observed in practice.

		$\frac{W_n(F^*)}{W_n(F_0)}$				$\Pr(N_w(F^*) < N_w(F_0))$		
	$\text{sgn}(E_2)$	n 10,000	100,000	500,000	w	10^{10}	10^{100}	10^{250}
p	$+$	1.08×10^{0}	2.95×10^{0}	5.84×10^{2}		0.595	0.609	0.653
	$-$	6.23×10^{4}	3.41×10^{48}	3.28×10^{242}		0.839	0.997	1.000
$\hat{p}$	$+$	6.91×10^{-2}	3.45×10^{-12}	5.02×10^{-58}		0.345	0.127	0.035
	$-$	9.4×10^{-2}	7.10×10^{-11}	7.55×10^{-51}		0.376	0.189	0.084
p_σ	$+$	1.07×10^{0}	2.88×10^{0}	5.77×10^{2}		0.582	0.598	0.656
	$-$	5.57×10^{4}	2.94×10^{48}	5.11×10^{242}		0.834	0.996	1.000

Table 1: Comparison of the optimal strategy F^* to the zero-variance strategy F_0, in scenarios where p is and is not known with certainty, and when only one or both of the bets hold positive expected value. The ratios of W_n are medians across 10,000 iterations.

The values of w in Table 1 are obviously extremely large. The guaranteed gain per wager for F_0, with the above range of parameter values, is upwards of 5%. With the five-hundred thousand bets that are considered, $1.05^{500{,}000}$ is also rather large. Numbers at this scale will often cause computational issues. Therefore, if we would typically calculate $W_n = \prod_{i=1}^{n} r_i$ – where $r_i \in [1, 1.05]$, for instance, for F_0 – then we instead calculate $\ln(W_n) = \sum_{i=1}^{n} \ln(r_i)$, which allows us to calculate $\frac{W_n(F^*)}{W_n(F_0)} = \exp(\ln(W_n(F^*)) - \ln(W_n(F_0)))$, avoiding any such complications. This analysis of placing five-hundred thousand wagers each with a guaranteed return of some 0-5% is certainly unrealistic. However, when arbitrage bets exist, their returns are often in this range, or better. The large number of bets is invoked to satisfy the conditions in (19), in an attempt to evaluate the theoretical aspects of our framework, rather than the necessarily practically-relevant.

Table 1 displays the results from the simulation. In the first two rows, it is evident that, when p is known with certainty, then, unsurprisingly, F^* outperforms F_0, especially asymptotically, according to both of the Breiman conditions in (19). When only one of the two bets holds positive expected value ($\text{sgn}(E_2) = -$), the outperformance is particularly stark, reinforcing the insight from Theorem 1, that returns to a favourable bet can in certain situations be improved by combining it with an individually-unfavourable bet. The second two rows, however, suggest that when p is not known and has to be estimated, then the safer zero-volatility strategy is superior (or at least that the Shin estimation is inferior to the F_0 one). Of course, the Shin estimate is not bad per se, as, for instance, a non-trivial proportion of F^* runs reaches wealth of 10^{250} before F_0 does, both with E_2 negative and positive. Interestingly, when E_2 is negative, as opposed to the positive case with $\hat{p}$, having only one favourable bet decreases the extent to which F_0 gives better returns (uniformly by the second metric, ambiguously by the first), which must suggest, at least somewhat, that in these cases, the safe strategy is more likely to be outperformed (also demonstrated, as mentioned, in the case where p is known).

The third pair of rows, corresponding to p_σ, provide the results from an alternative interpretation of uncertainty in the probability estimate. In this case, the bettor knows the mean of the true probability p, while the realised value of p is randomly distributed according to a Beta distribution. This allows the bettor to be correct on average, as they may well be in practice, but subject to variability on a wager-to-wager basis nonetheless. A Beta distribution is parameterised by (α, β), with mean $\mu = \frac{\alpha}{\alpha+\beta}$ and variance $\sigma^2 = \frac{\alpha\beta}{(\alpha+\beta)^2(1+\alpha+\beta)}$. The entropy for a Bernoulli distribution with probability p is given by $H(p) = -p\log(p) - (1-p)\log(1-p)$, a concave function which peaks at $p = 0.5$. Therefore, for a given value of p, we would like its Beta distribution to have higher variance the closer it is to 0.5. We choose the function

$$\sigma^2 = \frac{\mu(1-\mu)}{1 + k\sqrt{1 - 4(\mu - 0.5)^2}} \tag{20}$$

as the variance given the mean. The term $k\sqrt{1-4(\mu-0.5)^2}$ is concave, ranging from 0 at $\mu = 0, 1$ to k at $\mu = 0.5$. Figure 1 plots the relationship between σ^2 and μ for different values of k, chosen to give standard deviations of 0.1, 0.05 and 0.025, respectively, when $\mu = 0.5$.

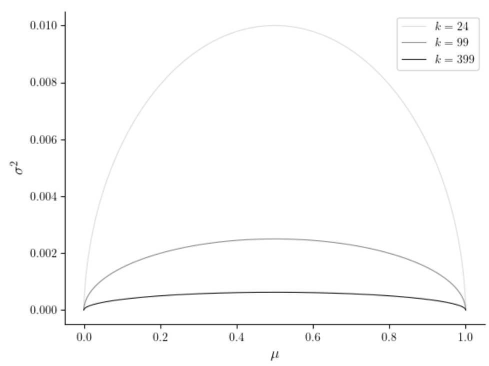

Figure 1: Variance in a Beta distribution as a function of its mean, via (20).

Using $k = 99$, the two rows in Table 1 under p_σ are given as above, only with the bettor having the estimate p for the true probability which is Beta distributed with mean p and variance $\sigma^2(p)$ from (20). Comparing these rows to the case where the true probability is known with certainty, all results are essentially the same, all being of the same order, with the certainty case being slightly favourable in ten of twelve categories. The marked outperformance for the case where one bet is unfavourable reinforces Theorem 1. This then confirms that, even if there exists uncertainty in the outcome-generating probability, as long as the estimate is accurate on average, then the optimal strategy outlined above is indeed superior to the zero-variance strategy. When the estimate is inefficient, however, F_0 is much safer.

In a sports-betting context, substantial evidence shows that bookmaker prices offer accurate probability forecasts (see, for instance, Štrumbelj (2014), Smith et al. (2009), Kaunitz et al. (2017)), especially when considering a cross-section of prices, rather than merely the most favourable offerings for each outcome. Arbitrageurs, then, would be wise to investigate the accuracy of such forecasts, and if they are deemed sufficient, to then implement the optimal strategy presented here.

5 Multiple Outcomes

Above we have discussed the case with two mutually-exclusive, collectively-exhaustive outcomes for a given event. It is of course possible that arbitrage opportunities exist for events with more than two outcomes. In these cases, the growth rate is given by

$$G(f_1, \dots, f_n) = (1 + f_1 b_1 - \dots - f_n)^{p_1} \dots (1 - f_1 - \dots + f_n b_n)^{p_n} \tag{21}$$

We speculate that in cases where it is optimal to stake with $\sum_{i=1}^{n} f_i = 1, f_i \in [0, 1] \ \forall \ i$, then $(f_1, f_2, \dots, f_n) = \left(p_1, p_2, \dots, 1 - \sum_{i=1}^{n-1} p_i\right)$. We have shown that this is the case above for $n = 2$. We can also verify that for $n = 3$, the resulting system of equations for this case is

$$\begin{aligned}
\frac{\delta \mathcal{L}}{\delta f_1} &= \frac{p_1 b_1}{1 + f_1 b_1 - f_2 - f_3} - \frac{p_2}{1 - f_1 + f_2 b_2 - f_3} - \frac{1 - p_1 - p_2}{1 - f_1 - f_2 + f_3 b_3} - \lambda = 0 \\
\frac{\delta \mathcal{L}}{\delta f_2} &= -\frac{p_1}{1 + f_1 b_1 - f_2 - f_3} + \frac{p_2 b_2}{1 - f_1 + f_2 b_2 - f_3} - \frac{1 - p_1 - p_2}{1 - f_1 - f_2 + f_3 b_3} - \lambda = 0 \\
\frac{\delta \mathcal{L}}{\delta f_3} &= -\frac{p_1}{1 + f_1 b_1 - f_2 - f_3} - \frac{p_2}{1 - f_1 + f_2 b_2 - f_3} + \frac{(1 - p_1 - p_2) b_3}{1 - f_1 - f_2 + f_3 b_3} - \lambda = 0 \\
\frac{\delta \mathcal{L}}{\delta \lambda} &= 1 - f_1 - f_2 - f_3 = 0
\end{aligned}$$

which is solved by $(f_1, f_2, f_3, \lambda) = (p_1, p_2, 1 - p_1 - p_2)$. We have evaluated for all cases up to $n = 30$, randomly drawing n-tuples $(b_1, \dots, b_n)$ such that $\sum_{i=1}^{n} \frac{1}{1+b_i} \in [0.95, 1]$, and finding that $(f_1, f_2, \dots, f_n) = \left(p_1, p_2, \dots, 1 - \sum_{i=1}^{n-1} p_i\right)$ gives the same value for λ from each of the n partial derivatives of $\mathcal{L}$ with respect to the f_i, rounded to ten decimal places. It remains to be shown whether this optimal strategy satisfies the properties in Breiman et al. (1961) compared to F_0 when n increases, as was the case for $n = 2$, but we cannot imagine why it would not.

6 Conclusion

In gambling contexts, especially online sports betting, it is not uncommon for arbitrage opportunities to exist. In this paper, we have outlined a Kelly-criterion (Kelly, 1956) optimal strategy for exploiting such advantageous investment opportunities, provided some of its most important properties, and evaluated it in various relevant situations. Whereas a zero-volatility strategy, where the net profits from each side of the arbitrage are equal, can be executed with knowledge only of the offered prices, the optimal strategy also requires an accurate estimate for the probability of each of the outcomes occurring. We have seen that, when the probability is known, the Kelly-optimal strategy yields significantly greater returns than the zero-volatility strategy, in accordance with the properties presented by Breiman et al. (1961). This is especially true when one of the two bets is unfavourable in isolation, and is a key result of our paper, where the returns from a favourable bet can be increased by combining it with an unfavourable wager. However, when the bettor does not have an accurate estimate for the outcome-generating probability, the zero-volatility strategy is much safer, in general achieving high returns as a benchmark. These insights are important for bettors, then, who are practicing arbitrageurs, as, if they can create accurate probability estimates, then their profits can increases significantly by following our novel strategy, compared to the risk-free benchmark.

7 References

Bernoulli, D. "Specimen theoriae novae de mensura sortis (Translated from Latin into English by L. Sommer in 1954)". *Commentarii Academiae Scientiarium Imperialis Petropolitanae*, vol. 1738, 1738, pp. 175–92.

Bernoulli, Daniel. "Exposition of a new theory on the measurement of risk". *The Kelly capital growth investment criterion: Theory and practice*, World Scientific, 2011, pp. 11–24.

betmonitor.com. "Comparing Odds the Easy Way", www.betmonitor.com/. Accessed 16 Aug. 2024.

Breiman, Leo, et al. "Optimal gambling systems for favorable games". *The Kelly Capital Growth Investment Criterion*, 1961, pp. 47–60.

Jullien, Bruno, and Bernard Salanié. "Measuring the incidence of insider trading: A comment on Shin". *The Economic Journal*, vol. 104, no. 427, 1994, pp. 1418–19.

Karush, W. "Minima of Functions of Several Variables with Inequalities as Side Constraints". *M.Sc. thesis*, 1939.

Kaunitz, Lisandro, et al. "Beating the bookies with their own numbers-and how the online sports betting market is rigged". *arXiv preprint arXiv:1710.02824*, 2017.

Kelly, John L. "A new interpretation of information rate". *the bell system technical journal*, vol. 35, no. 4, 1956, pp. 917–26.

Kuhn, H.W., and A. W. Tucker. "Nonlinear Programming". *Proceedings of 2nd Berkely Symposium*, 1951, pp. 481–92.

oddspedia.com. "Sure Bets", oddspedia.com/surebets. Accessed 16 Aug. 2024.

Shin, Hyun Song. "Measuring the incidence of insider trading in a market for state-contingent claims". *The Economic Journal*, vol. 103, no. 420, 1993, pp. 1141–53.

Smith, Michael A, et al. "Do bookmakers possess superior skills to bettors in predicting outcomes?" *Journal of Economic Behavior & Organization*, vol. 71, no. 2, 2009, pp. 539–49.

Spitznagel, Mark. *Safe Haven: Investing for Financial Storms*. John Wiley & Sons, 2021.

Štrumbelj, Erik. "On determining probability forecasts from betting odds". *International journal of forecasting*, vol. 30, no. 4, 2014, pp. 934–43.

Thorp, Edward O. "Portfolio choice and the Kelly criterion". *Stochastic optimization models in finance*, Elsevier, 1975, pp. 599–619.