IMA Journal of Applied Mathematics (2007) **72**, 43−51
doi:10.1093/imamat/hxl027
Advance Access publication on December 5, 2006

The Kelly criterion for spread bets

S. J. CHAPMAN†
Oxford Centre for Industrial and Applied Mathematics,
Mathematical Institute, 24–29 St Giles', Oxford OX1 3LB, UK

[Received on 20 January 2006; accepted on 22 June 2006]

The optimal betting strategy for a gambler betting on a discrete number of outcomes was determined by Kelly (1956, A new interpretation of information rate. *J. Oper. Res. Soc.*, **57**, 975–985). Here, the corresponding problem is examined for spread betting, which may be considered to have a continuous distribution of possible outcomes. Since the formulae for individual events are complicated, the asymptotic limit in which the gamblers edge is small is examined, which results in universal formulae for the optimal fraction of the bank to wager, the probability of bankruptcy and the distribution function of the gamblers total capital.

Keywords: gambling; strategy; optimal bankruptcy; expectation.

1. The Kelly criterion

Suppose a gambler undertakes bets on the outcomes of a series of events which are described by independent, identically distributed random variables W_n, which take the value 1 (corresponding to a win) with probability p and 0 (corresponding to a loss) with probability $q = 1 - p$. If s is the 'starting price' of the bet (so that the profit is s times the stake on a win, with the stake returned) and tax is charged at a rate t (which the wise punter pays on the stake rather than on the return), then the payoff on a bet of size S is

$$-tS - S + W(1+s)S$$

and the expected profit is

$$S(-t - 1 + p(1+s)) = \epsilon S, \tag{1}$$

say. Kelly (1956) addressed the question of how a gambler with a fixed initial amount of money (the 'bank') should bet to maximize his return (given that he has an edge over the bookmaker, so that $\epsilon > 0$). On any finite number of bets, the return is maximized by wagering the whole bank each time. However, if this strategy is followed, the first losing bet wipes out the bank, so that there is a high probability of finishing with nothing and a small probability of amassing a fortune. As the number of bets tends to infinity, the probability of bankruptcy tends rapidly to one.

If we let f_n represent the bank after n bets and bet a fraction α of the bank each time, then

$$f_n = f_{n-1}(1 - \alpha t - \alpha + W_n(s+1)\alpha),$$

so that

$$f_n = \prod_{m=1}^{n} (1 - \alpha t - \alpha + W_m(s+1)\alpha). \tag{2}$$

†Email: chapman@maths.ox.ac.uk

Kelly showed that the optimal strategy is to maximize not the expectation of f_n but that of $\log(f_n)$. Taking logarithms, we have

$$\log f_n = \sum_{m=1}^{n} \log\left(1 - \alpha t - \alpha + W_m(s+1)\alpha\right), \qquad (3)$$

in which the right-hand side is now a sum of independent random variables, to which the law of large numbers applies. We can calculate the expectation of the sum easily as

$$E(\log f_n) = n(q\log(1-\alpha t-\alpha) + p\log(1-\alpha t-\alpha+(s+1)\alpha)).$$

Maximizing this expectation with respect to α, we require

$$\frac{\partial E(\log f_n)}{\partial\alpha} = n\left(-\frac{q(1+t)}{1-\alpha-\alpha t} + \frac{p(s-t)}{1+\alpha s-\alpha t}\right) = 0.$$

Solving for α gives

$$\alpha = \frac{ps-q-t}{(s-t)(1+t)} = \frac{\epsilon}{(s-t)(1+t)}, \qquad (4)$$

with the expected growth rate of the bank as

$$\frac{E(\log f_n)}{n} = q\log\left(\frac{q(1+s)}{s-t}\right) + p\log\left(\frac{p(1+s)}{1+t}\right).$$

The preceding analysis is modified only slightly if the bet is placed with a betting exchange rather than a traditional bookmaker. On an exchange, there is no tax, but there is a commission to be payed on winning bets, which is a percentage c of the profit. Thus, in this case

$$f_n = f_{n-1}\left(1-\alpha+W_n(s(1-c)+1)\alpha\right),$$

and a similar analysis shows that

$$\alpha = \frac{ps(1-c)-q}{s(1-c)} = \frac{\epsilon}{s(1-c)},$$

where $\epsilon = ps(1-c)-q$ is the expected return on a unit stake and

$$\frac{E(\log f_n)}{n} = q\log\left(\frac{q(1+s-sc)}{s(1-c)}\right) + p\log\left(p(1+s-sc)\right).$$

2. Spread bets

Spread betting has become a popular way to gamble on the outcome of a range of sporting (and other) events. Let us consider as an example a simple spread bet on the time of the first goal in a soccer match. Before the start of the game, the bookmaker will quote a spread of (e.g.) 31–34 min. If the punter takes the view that the first goal will be scored before 31 min, he can ‘sell at 31’ and the payout on the bet (which may be positive or negative) is $S(31-X)$, where X is the actual time of the first goal and S is the unit stake. On the other hand, if the punter thinks an early goal is unlikely, he may ‘buy at 34' with a payout of $S(X-34)$. Thus, each spread bet is essentially a future (forward) contract. An important difference between the traditional bets considered in Section 1 is that the potential losses and the potential winnings are unknown *a priori*, and may even be unbounded. An analysis of the fair valuation of soccer spread bets has been given recently in Fitt *et al.* (2006).

Although the final value of the index at the conclusion of a spread betting event (the 'makeup') is discrete (taking the values 1–90 in the case of the first goal), we can approximate it well in many cases by a continuous distribution.

So, suppose now that our gambler is spread betting on a series of independent, identically distributed random events. We suppose that the random variable (which may be continuous or discrete) describing the n th event is X_n, with probability distribution function $f(x)$, and that the payoff is $g(X_n)$ per unit stake (usually $g(X)$ is linear in X as above). If the unit stake is chosen to be a fraction α of the bank each time, then

$$f_n = f_{n-1} + \alpha f_{n-1} g(X_n),$$

so that

$$f_n = \prod_{m=1}^{n} (1 + \alpha g(X_m)).$$

Taking logarithms as before gives

$$\log f_n = \sum_{m=1}^{n} \log(1 + \alpha g(X_m)).$$

Thus,

$$E(\log f_n) = nE(\log(1 + \alpha g(X_n))) = n \int \log(1 + \alpha g(x)) f(x) \mathrm{d}x.$$

Maximizing over α, we find that the optimal value satisfies

$$\int \frac{g(x) f(x) \mathrm{d}x}{1 + \alpha g(x)} = E\left(\frac{g(X)}{1 + \alpha g(X)}\right) = 0. \tag{5}$$

The Kelly criterion for spread bets has been considered previously by Haigh (2000), who obtains the criterion (5) (with $g(X)$ linear in X). Haigh proceeds to analyse a more general situation in which there may be more than one winner in an event (such as performance indices, see Section 2.1.2). In such cases, many counter-intuitive effects may arise—the optimal strategy may involve not betting on some teams even though the expected payoff is positive, and may even involve betting on teams for which the expected payoff is negative. Here, we consider that the punter has information about one team only, and refer the reader interested in multiple winners to Haigh (2000), where many interesting examples are given.

2.1 *Examples*

2.1.1 *Time of first goal.* We model a soccer match by assuming that the probability of a goal being scored in the time period $(t, t + \delta t)$ is $\mu \delta t$, and that this is independent of the history of the match so far. For simplicity, we assume that μ is constant; μ is then equal to the expected number of goals in the game divided by 90. Then the time of the first goal, X, is exponentially distributed for $X < 90$, with probability density function $f(x) = \mu\,\mathrm{e}^{-\mu x}$. If there is no goal, then the makeup is 90, so that the probability that $X = 90$ is given by $\mathrm{e}^{-90\mu}$. The expected value of X is

$$\frac{1-\mathrm{e}^{-90\mu}}{\mu}.$$

If this is greater than the offer price or less than the bid price, then the punter has an edge. Let us consider the former case (the latter can be treated similarly). The punter should stake a proportion α of his capital, where

$$\begin{aligned} 0 &= E\left(\frac{(X-K)}{1+\alpha(X-K)}\right) \\ &= \int_0^{90} \frac{\mu(x-K)\mathrm{e}^{-\mu x}\,\mathrm{d}x}{1+\alpha(x-K)} + \frac{(90-K)\mathrm{e}^{-90\mu}}{1+\alpha(90-K)} \\ &= \frac{1}{\alpha} - \frac{\mu\,\mathrm{e}^{(1/\alpha-K)\mu}}{\alpha^2}\left(\mathrm{Ei}\left(\frac{(1-K\alpha)\mu}{\alpha}\right) - \mathrm{Ei}\left(\frac{(1+\alpha(90-K))\mu}{\alpha}\right)\right) \\ &\quad - \frac{\mathrm{e}^{-90\mu}}{\alpha(1+\alpha(90-K))}, \end{aligned} \tag{6}$$

where K is the offer price and

$$\mathrm{Ei}(z) = \int_z^\infty \frac{\mathrm{e}^{-t}\,\mathrm{d}t}{t}$$

is the exponential integral. The solution of (6) gives α as a function of μ and K.

2.1.2 *Win index bets.* Many spread bets are based on a performance index, in which, e.g. a team receives 25 points for a win, 10 points for a draw and nothing for a loss. The expected value of X is then $25p_\mathrm{w} + 10p_\mathrm{d}$, where p_w is the probability of a win and p_d is the probability of a draw. If this is greater than the offer price K, then α should be chosen such that

$$0 = E\left(\frac{(X-K)}{1+\alpha(X-K)}\right) = \frac{p_\mathrm{w}(25-K)}{1+\alpha(25-K)} + \frac{p_\mathrm{d}(10-K)}{1+\alpha(10-K)} - \frac{(1-p_\mathrm{w}-p_\mathrm{d})K}{1-\alpha K}. \tag{7}$$

We note again here that (7) is appropriate for the two-team situation such as a soccer match. Similar markets exist in horse racing, giving 25 points for a win, 10 points for a second and five points for a third. In that case, if the punter has an estimate of the probabilities of a win, second and third for a single horse, then (7) is appropriate. On the other hand, if the punter has these estimates for all horses, then, because the outcomes for different horses are correlated, the stakes on each horse must be optimized together rather than independently (see Haigh 2000).

3. The limit of small edge

The preceding formulae (6) and (7) are unwieldy, and also (in the second case) depend on the punter estimating the odds of all possible outcomes of the event. It is therefore of interest to see what simplifications are possible in the limit that the expected payoff on a unit stake,

$$\epsilon = \int g(x) f(x) \mathrm{d}x,$$

is small (which is almost always the case). In this case, α will also be small and expanding (5) for small α gives

$$0 \sim \int g(x) f(x)(1 - \alpha g(x) + \cdots) \mathrm{d}x = \epsilon - \alpha\sigma^2 + \mathrm{O}(\epsilon^2),$$

where σ^2 is the variance of the payoff $g(X)$. Thus, to leading order in ϵ,

$$\alpha = \frac{\epsilon}{\sigma^2}, \tag{8}$$

and the expected growth rate of the capital (per bet) is

$$E(\log(1 + \alpha g(X))) = E\left(\log\left(1 + \frac{\epsilon g(X)}{\sigma^2}\right)\right) \sim E\left(\frac{\epsilon g(X)}{\sigma^2}\right) = \frac{\epsilon^2}{\sigma^2}.$$

The beauty of a formula like (8) is that ϵ and σ^2 are easy to estimate even for complicated bets. In particular, for bets made up of many independent components (such as the sum of the winning lengths at a horse racing meeting), ϵ and σ^2 are simply the sum of the expectation and variance for each event.

4. Probability of bankruptcy

Since a gambler using the Kelly criterion is betting a small percentage of his bank every time, the theoretical probability of the bankruptcy is zero. However, since bet sizes are in practice constrained to be multiples of some minimum amount, we may consider that the punter is effectively bankrupt if his capital falls to say 1% of its initial value. It is of interest to estimate the probability of ever reaching a given low point, and in particular to determine its sensitivity to errors in estimating ϵ. Thus, we let $p(x)$ to be the probability that the bank hits the value ν, given that the current size of the bank is x.

Conditioning on the next bet, we find

$$p(x) = \int p(x(1 + \alpha g(y))) f(y) \mathrm{d}y.$$

Expanding for small α gives

$$\begin{aligned} p(x) &\sim \int p(x + \alpha x g(y)) f(y) \mathrm{d}y \\ &\sim \int \left(p(x) + \alpha x g(y) \frac{\mathrm{d}p}{\mathrm{d}x}(x) + \frac{\alpha^2 x^2 g(y)^2}{2} \frac{\mathrm{d}^2 p}{\mathrm{d}x^2}(x) + \cdots \right) f(y) \mathrm{d}y \\ &\sim p(x) + \alpha\epsilon x \frac{\mathrm{d}p}{\mathrm{d}x} + \frac{\alpha^2 x^2 \sigma^2}{2} \frac{\mathrm{d}^2 p}{\mathrm{d}x^2} + \cdots . \end{aligned}$$

Thus, we have, to leading order,

$$0 = x\epsilon \frac{\mathrm{d}p}{\mathrm{d}x} + \frac{\alpha x^2 \sigma^2}{2} \frac{\mathrm{d}^2 p}{\mathrm{d}x^2}, \tag{9}$$

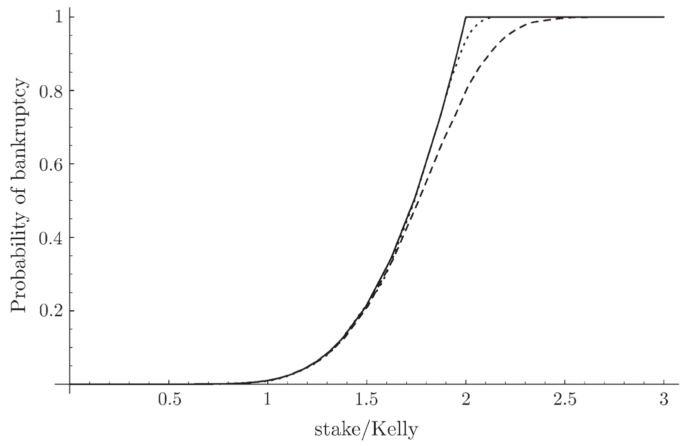

FIG. 1. Probability of the bank reducing to 1% of its initial size against stake as a multiple of the Kelly stake. Also shown are Monte Carlo simulations over 10 000 (dashed) and 100 000 (dotted) binary bets.

with boundary conditions

$$p(\nu) = 1,$$
$$p \to 0 \quad \text{as } x \to \infty.$$

The general solution of (9) is

$$p = A + Bx^k,$$

where

$$k = 1 - \frac{2\epsilon}{\alpha\sigma^2}.$$

If $k > 0$, then x^k grows as $x \to \infty$, so that we must have $A = 1$, $B = 0$ and the probability of bankruptcy is 1, regardless of the initial size of the bank. If $k < 0$, then $A = 0$ and $B = \nu^{-k}$. Thus, the probability of ever reaching a given fraction $r = \nu/x$ of the current bank is

$$p = \begin{cases} r^{\frac{2\epsilon}{\alpha\sigma^2}-1}, & \alpha < \frac{2\epsilon}{\sigma^2}, \\ 1, & \alpha > \frac{2\epsilon}{\sigma^2}. \end{cases}$$

With the optimal value $\alpha = \epsilon/\sigma^2$, we have $k = -1$ and $p = r$. However, even though with the optimal α the chance of the bank reducing to 1% of its current value is only 1 in 100, if we overestimate our edge by a factor of two, then it becomes certain that we will become bankrupt. For this reason, many betters err on the side of caution and reduce the stake by a factor of two (known as betting 'half-Kelly'). The chance of the bank reducing to 1% of its initial value is shown in Fig. 1, along with Monte Carlo simulations of 10 000 and 100 000 binary bets.

5. Probability distribution function of the bank

It is also of interest to calculate the probability distribution function of the bank at future times. Defining the probability density function

$$p_n(x) = P(x < f_n < x + \delta x),$$

so that the cumulative distribution function

$$P(f_n < x) = \int_0^x p_n(y)\mathrm{d}y,$$

we have, on conditioning on the previous bet,

$$P(f_n < x) = \int P\left(f_{n-1} < \frac{x}{1+\alpha g(y)}\right) f(y)\mathrm{d}y.$$

Differentiating with respect to x gives

$$p_n(x) = \int p_{n-1}\left(\frac{x}{1+\alpha g(y)}\right)\frac{f(y)}{1+\alpha g(y)}\mathrm{d}y.$$

Thus, on expanding for small α,

$$\begin{aligned}
p_n(x) &\sim \int p_{n-1}(x - \alpha x g(y) + \alpha^2 x g(y)^2 + \cdots) f(y)(1 - \alpha g(y) + \alpha^2 g(y)^2 + \cdots)\mathrm{d}y \\
&\sim \int \Bigg(p_{n-1}(x) + (-\alpha x g(y) + \alpha^2 x g(y)^2)\frac{\mathrm{d}p_{n-1}}{\mathrm{d}x}(x) + \frac{\alpha^2 x^2 g(y)^2}{2}\frac{\mathrm{d}^2 p_{n-1}}{\mathrm{d}x^2}(x) + \cdots \\
&\qquad -\alpha g(y) p_{n-1}(x) + \alpha^2 x g(y)^2 \frac{\mathrm{d}p_{n-1}}{\mathrm{d}x}(x) + \alpha^2 g(y)^2 p_{n-1}(x) + \cdots \Bigg) f(y)\mathrm{d}y \\
&\sim p_{n-1} + (-\alpha\epsilon x + \alpha^2 x\sigma^2)\frac{\mathrm{d}p_{n-1}}{\mathrm{d}x} + \frac{\alpha^2 x^2\sigma^2}{2}\frac{\mathrm{d}^2 p_{n-1}}{\mathrm{d}x^2} \\
&\qquad -\alpha\epsilon p_{n-1} + \alpha^2 x\sigma^2\frac{\mathrm{d}p_{n-1}}{\mathrm{d}x} + \alpha^2\sigma^2 p_{n-1}\cdots .
\end{aligned}$$

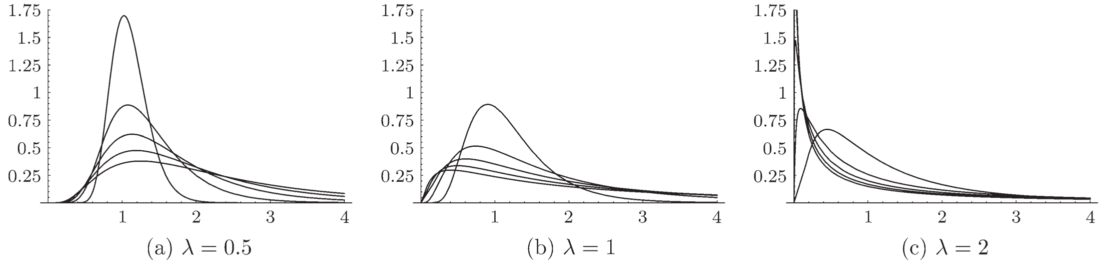

FIG. 2. Evolution of the probability density functions of the bank, at (a) half-Kelly, (b) Kelly and (c) double Kelly stakes. The initial distribution is a δ-function centred at $x = 1$, and the curves correspond to $n\epsilon^2/\sigma^2 = 0.2$, 0.6, 1.0, 1.4 and 1.8, where n is the number of bets (corresponding to times t of $0.2\lambda^2$, $0.6\lambda^2$, $1.0\lambda^2$, $1.4\lambda^2$ and $1.8\lambda^2$). The δ-function relaxes and spreads out, and in case (c) is subsequently drawn into a sharp peak near the origin.

Defining time t in terms of the number of bets as $t = n\alpha^2\sigma^2$, we have, to leading order in α,

$$\frac{\partial p}{\partial t} = \left(1 - \frac{\epsilon}{\alpha\sigma^2}\right) p + \left(2 - \frac{\epsilon}{\alpha\sigma^2}\right) x \frac{\partial p}{\partial x} + \frac{1}{2} x^2 \frac{\partial^2 p}{\partial x^2}.$$

We write α as a multiple of the Kelly value by setting $\alpha = \lambda\epsilon/\sigma^2$, and set $y = \log x$, to give

$$\frac{\partial p}{\partial t} = \left(1 - \frac{1}{\lambda}\right) p + \left(\frac{3}{2} - \frac{1}{\lambda}\right) \frac{\partial p}{\partial y} + \frac{1}{2} \frac{\partial^2 p}{\partial y^2}.$$

If we assume that initially the bank is of magnitude 1, then $p(y, 0) = \delta(y)$, where δ is the Dirac δ-function. Now, writing $z = y + (3/2 - 1/\lambda)t$ and $p = \mathrm{e}^{(1-1/\lambda)t}\hat{p}(z, t)$ gives

$$\frac{\partial \hat{p}}{\partial t} = \frac{1}{2} \frac{\partial^2 \hat{p}}{\partial y^2},$$

with solution

$$\hat{p} = \frac{\mathrm{e}^{-z^2/2t}}{\sqrt{2\pi t}},$$

so that

$$p = \mathrm{e}^{(1-1/\lambda)t} \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{(\log x + (3/2 - 1/\lambda)t)^2}{2t}\right).$$

Figure 2 shows the evolution of the probability density function of the bank for $\lambda = 1/2$, $\lambda = 1$ and $\lambda = 2$. We see that even when using the Kelly optimal stake, the most likely value of the bank is less than 1.

6. Conclusion

We have examined the extension of the Kelly criterion to spread bets on the outcome of continuous random variables. Although the formulae for the optimal fraction of the bank to stake each time are easy to calculate given the probability distribution function of the event, they are unwieldy and difficult to use in practice. We therefore derived an approximate formula valid when the expected profit on a unit stake (the ‘edge’ over the bookmaker) is small, which is usually the case in practice. We found that the optimal fraction of the bank to stake is just the expected profit on a unit stake divided by its variance.

We then derived simple expressions for the probability of bankruptcy and for the distribution function of the bank after n bets. These showed in particular that the chance of the bank reducing to 1% of its initial value was only 1 in 100, but that overestimating the edge over the bookmaker by a factor of two would result in certain bankruptcy.

Acknowledgements

The author is grateful to the referees for drawing the paper of Haigh to his attention, and for other useful comments.

References

Fitt, A. D., Howls, C. J. & Kabelka, M. (2006) The valuation of soccer spread bets. *J. Oper. Res. Soc.*, **57**, 975–985.

Haigh, J. (2000) The Kelly criterion and bet comparisons in spread betting. *Statistician*, **49**, 531–539.

Kelly Jr., J. L. (1956) A new interpretation of information rate. *Bell Syst. Tech. J.*, **35**, 917–926.