# An explicit solution to the problem of optimizing the allocations of a bettor's wealth when wagering on horse races



Peter Boleslaw Smoczynski
Thompson Rivers University

# An Explicit Solution of the Problem of Optimizing the Allocations of a Bettor's Wealth when Wagering on Horse Races

PETER SMOCZYNSKI* and DAVE TOMKINS*

**2000 Mathematics Subject Classification:** 49K45, 65C50, 93E99, 90C46, 90C25, 90C90.

**Keywords:** optimal wealth allocation, constrained convex optimization, random processes, Kelly criterion, horse races.

**Abstract**

An explicit formula for the optimal strategy for betting allocation on horse races is given. The formula for maximal value of the logarithm of average geometric growth rate is also given. The solution is obtained with the help of KKT theory. Application of the formulas requires one to construct the optimal set of horses to bet on. For a horse to be included in the set, the expected revenue rate must be greater than the fraction of unallocated wealth. A simple way, without solving any equations, for determination of the optimal set of horses is given.

## 1. Introduction

In the seminal paper [1] Kelly considered repeated betting on horse races. He described the optimal strategy for a single horse and gave incomplete solution to the problem with many horses and he assumed the track take to be zero. A more general problem has been considered in [4] & [5]. In [4] it is proved that, when the number of atoms of the probability distribution is smaller or equal

---

* Department of Mathematics and Statistics, Thompson Rivers University, Main Campus, Box 3010, 900 McGill Road, Kamloops, BC., Canada V2C 5N3, psmoczynski@tru.ca, tomkins@tru.ca

than the number of degrees of freedom then the problem of finding the optimal solution is reducible to that of solving two systems of equations. It is concluded that it is impossible to find an explicit solution. In [5] several interesting general theorems are proved about the general case, however explicit solutions are not given. In [6] an elegant algorithm for approximate solutions of the general case is given. In [2], independently of [4], the two systems of equations were introduced and investigated for the problem of horse races with non-zero track take. The method described there produces many candidate solutions. In this paper we use the KKT theory to find the optimal solution for horse races with non-zero track take under the condition that the horseplayer cannot borrow money and cannot sell the bets short. We give a complete explicit solution of a problem set in 1956 by Kelly.

We solve the problem in three steps: at first in Section 6 we solve the optimal system, somewhat similar like in [2]. Then in Section 7 we analyze the conditions of solvability and we describe in simple language how to find the unique optimal solution without solving any equations. In order to do this we reorder the horses according to their expected revenue rates in Section 7. The optimal set $S^{opt}$ of horses contains all the horses whose expected revenue rates are bigger than the fraction of the wealth that the horseplayer does not bet, also called the reserve rate.

## 2. Horse races, track take and dividends

n horses compete in a horse race repeated infinitely many times. The respective probabilities for each horse to win are the same at each race, $\pi_k$ for k-th horse, although different horses will be winning in turns. Each time the horseplayer bets an allocation fraction $f_k$, always the same, of his wealth on k-th horse to win. After each race the track management takes some fraction (tt = the track take) of the total money bet on all the horses and the dividend fraction $D = 1 - tt \in (0, 1]$ of the remaining money is divided between those who placed bets on the winning horse, proportionally to their bets. All the other bettors lose their money.

Let b = [$b_1$, $b_2$, … $b_n$] $\in R^n$ be the vector with components equal to the total sum of money bet on each horse. The total sum to be divided between the bettors and the management of the track establishment is $\sum_m b_m = \sum_{m=1}^{m=n} b_m$. The vector

(2.1) $\beta = b \Big/ \sum_m b_m$

satisfies the condition $\sum_m \beta_m = 1$, hence its components may be interpreted as the belief probabilities that describe the collective belief of the bettors about the respective probabilities of winning for each of the horses. The revenue rates vector is

(2.2) $r_k = \frac{D}{b_k} \sum_m b_m = D \big/ \beta_k = Q_k + 1$,

and the expected revenues are

(2.3) $er_i = \frac{D\pi_k}{\beta_k} = \pi_i (Q_i + 1)$.

Usually the dividend rate is between 0.80 & 0.85 and $Q_k$ represents the odds of horse k (a profit of $Q_k$ dollars for each \$1 bet, if the horse k wins).

## 3. The objective function & constraints on bets

We consider a horseplayer who has made accurate estimates of the probabilities $\pi$ = [$\pi_1$, $\pi_2$, … $\pi_n$] of each horse winning. We also assume that horseplayer's bets are small enough as not to alter the estimates $\beta$ of the same probabilities by other players. When k-th horse wins then horseplayer's wealth grows by a factor $1 - \sum_m f_m + r_k f_k$ , the revenue rate on each dollar bet on k-th horse, see (2.2). The horseplayer's strategy, described by the vector $f$ = [$f_1$, $f_2$, … $f_n$] of allocation fractions bet on horses, constitutes a self-financing strategy if his wealth changes without any further investments after the initial wealth is reserved for betting. After W repetitions of the same race his wealth grows by a factor $\left(\prod_{k=1}^{k=n}\left(1-\sum_m f_m + D\, f_k/\beta_k\right)^{\frac{w_k}{W}}\right)^W$ where $W=\sum_k W_k$ and $w_k$ is the number of times the k–th horse won. $L_w(f)=\sum_k \frac{w_k}{W}\ln\left(1-\sum_m f_m + D\, f_k/\beta_k\right)$ is the average of independent identically distributed random variables, each taking the value $\ln\left(1-\sum_m f_m + D\, f_k/\beta_k\right)$ with probability $\pi_k$. By the strong version of the Law of the Large Numbers $L_W(f)$ converges almost surely to

(3.1) $\quad L(f) = \sum_k \pi_k \ln\left(1-\sum_m f_m + D\, f_k/\beta_k\right)$.

Since $L(f) \le \sum_k \pi_k \ln(1+D/\beta_k)$, under the constrains on $f$ described at the end of Section 3, then also $\prod_{k=1}^{k=n}\left(1-\sum_m f_m + D\, f_k/\beta_k\right)^{\frac{w_k}{W}} = e^{L_W(f)}$ converges almost surely to $e^{L(f)}$ as $e^x$ is Lipschitz continuous on $(-\infty, \sum_k \pi_k \ln(1+D/\beta_k)]$. Thus the horseplayer is interested in maximizing the objective function $L(f)$. $\mathrm{Dom}(L) = \{ f \in R^n \mid 0 < 1-\sum_m f_m + D\, f_k/\beta_k$ for every k $\} \subseteq R^n$ is a convex set as an intersection of half-spaces. If any of $\beta_k = 0$ then it is easy to find the optimal strategy: bet a penny on k-th horse and $L(f) = \infty$. The rest of the paper is concerned with the case when all $\beta_k > 0$.

L(f) is of $C^\infty$ class on its domain and, and if all $\pi_k > 0$ then it is strictly concave-downwards as a finite linear combination, with non-negative coefficients, of strictly concave logarithms. At least one of the revenue rates (2.3) must be bigger than 1, otherwise no bet would be made. Consequently, Dom(L) must be unbounded in this direction and L(f) must grow to infinity in this direction. Therefore *L(f) achieves a unique maximum on every convex bounded subset of Dom(L) and the maximum is located on the boundary of the subset*.

We consider only betting North American style, only backing a horse to win, never laying a horse to lose. The first condition is equivalent to a constraint $\sum_m f_m \le 1$, the second to $0 \le f_m$ for all m. Thus we shall look for the maximum value of the objective function (3.1) on the restricted set

(3.2) $\quad P = \{ f \mid 0 \le f_m \ \&\ 0 \le 1 - \sum_m f_m \} \subseteq \mathrm{Dom}(L(f))$.

The restricted domain is convex, hence *there exists exactly one maximum of L( $f$ ) on P and it is located on the boundary of P.*

## 4. KKT theory

We shall use the KKT theory [3] in order to write the system of optimality equations for the point at which L attains its maximum, since it is located at the boundary of P. The maximization problem we consider is of the following form: Find a maximum of a differentiable concave down function L(f) on a feasible convex set { $\phi \mid 0 \le g_k(f)$ } $\subseteq$ Dom(L(f)), where the functions used in constraints are also differentiable. In order to write equations for the maximum points, we form the Lagrange function $LL(f,\lambda) = L(f) + \sum_p \lambda_p g_p(f)$. KKT theory states that every constrained maximum point of L(f) is also an "unconstrained" maximum point of LL(f, $\lambda$ ). The list of all the conditions is as follows:

(4.1) $$\frac{\partial}{\partial f_i} LL(f,\lambda) = \frac{\partial}{\partial f_i} L(f) + \sum_p \lambda_p \frac{\partial}{\partial f_i} g_p(f) = 0 \text{ for all i,}$$

(4.2) $$\lambda_p g_p(f) = 0,$$

(4.4) $$0 \le \lambda_p \text{ for all p,}$$

(4.5) $$\frac{\partial}{\partial \lambda_p} LL(f,\lambda) = g_p(f) \ge 0 .$$

We shortly describe the meaning of these requirements. Requirement (4.5) is just repetition of the equations of the constraints imposed on variables $f_i$, (4.5) is listed here for the sake of symmetry and completeness. Requirement (4.2) states the following: if $0 < g_k(f)$ then the additional variables λ are all zero. In this case (4.1) reduces to the simplest requirements for an optimal point: $\frac{\partial}{\partial f_i} L(f) = 0$. The additional variables are allowed to be non-zero only if $0 = g_k(f)$, that is when the optimal point is on the boundary of the feasible region. Now (4.1) may be interpreted as follows: if the optimal point is on the boundary then the gradient $\frac{\partial}{\partial f_i} L(f)$ must be a linear combination of the gradients $\frac{\partial}{\partial f_i} g_p(f)$ of the constraint functions. The essence of this requirement is this: the gradient $\frac{\partial}{\partial f_i} L(f)$ must be perpendicular to the surface on which $0 = g_k(f)$. Intuition says that this is correct, because at an extreme value on the surface $0 = g_k(f)$ the derivative of the objective function in any direction tangent to the surface must be zero. But the gradients $\frac{\partial}{\partial f_i} g_p(f)$ are perpendicular to these tangent directions. At last we explain the meaning of (4.4). The gradients $\frac{\partial}{\partial f_i} g_p(f)$ are directed towards the interior of the feasible set. Therefore (4.1) requires that if the optimal point is on the boundary then the objective function wants to grow in the direction perpendicular to the boundary and, by (4.4), towards the exterior of the feasible set.

## 5. The optimality system

For the problem of maximization of a concave down function (3.1) on the convex set (3.2), the Lagrange function is $LL(f,\lambda) = \sum_k \pi_k \ln\left(1 - \sum_m f_m + D\, f_k / \beta_k\right) + \sum_k f_k \lambda_k + \lambda_0 \left(1 - \sum_k f_k\right)$. Thus the optimality equations are:

Explicit solution for optimal betting,

(5.1) $$\frac{\partial}{\partial f_i} LL(f,\lambda) = \frac{\pi_i D/\beta_i}{\left(1 + Df_i/\beta_i - \sum_m f_m\right)} - \sum_k \frac{\pi_k}{\left(1 + Df_k/b_k - \sum_m f_m\right)} + \lambda_i - \lambda_0 = 0,$$

(5.2) $f_i \lambda_i = 0$ for all i,

(5.3) $\lambda_0 \left(1 - \sum_k f_k\right) = 0,$

(5.4) $0 \le \lambda_i$ for all i, and $0 \le \lambda_0$,

(5.5) $0 \le f_i$ for all i.

Equations (5.1) correspond to (4.1), (5.2) & (5.3) to (4.2), (5.4) to (4.4) and (5.5) to (4.5). This system (of equations & inequalities) has gotten, in the bounded convex set (3.2), a unique solution, since L(f) is strictly concave down.

## 6. Solution of the optimality system

In this Section we solve the optimality system (5.1 – 5.5) in the set P given by (3.2). The maxima with $\sum_m f_m = 1$ are lower than the ones found here (the proof is omitted). Hence we assume in what follows that $0 < 1 - \sum_k f_k$ and then $\lambda_0 = 0$ thanks to (5.3). Thanks to equations (5.5) the indices {1, 2, …, n} may be divided into two groups: the set S of those indices i for which $f_i > 0$ and its complement with $f_j = 0$ for j not in S. So, if $i \in S$ then $\lambda_i = 0$ thanks to (5.2) and the corresponding equations of the group (5.1) may be written as follows:

(6.1) $$\frac{D\pi_i}{\left(1 + Df_i/\beta_i - \sum_{m\in S} f_m\right)} = \beta_i \left( \sum_{k\in S} \frac{\pi_k}{\left(1 + Df_k/b_k - \sum_{m\in S} f_m\right)} + \frac{\sum_{k\notin S} \pi_k}{\left(1 - \sum_{m\in S} f_m\right)} \right), i \in S.$$

We sum the part of equations corresponding to $i \in S$ and we obtain an equation

$$D\sum_{k\in S} \frac{\pi_k}{\left(1 + Df_k/\beta_k - \sum_{m\in S} f_m\right)} = \left( \sum_{k\in S} \frac{\pi_k}{\left(1 + Df_k/b_k - \sum_{m\in S} f_m\right)} + \frac{\sum_{k\notin S} \pi_k}{\left(1 - \sum_{m\in S} f_m\right)} \right) \sum_{k\in S} \beta_k$$

and we transform it into

Explicit solution for optimal betting,

$$(6.2)\qquad \sum_{k\in S}\frac{\pi_k}{\left(1+Df_k/\beta_k-\sum_{m\in S}f_m\right)}=\frac{\sum_{k\notin S}\pi_k}{\left(1-\sum_{m\in S}f_m\right)}\frac{\sum_{k\in S}\beta_k}{\left(D-\sum_{k\in S}\beta_k\right)}.$$

We substitute this into (6.1) and we obtain

$\dfrac{D\pi_i}{\left(1+Df_i/\beta_i-\sum_{m\in S}f_m\right)}=\dfrac{D\beta_i}{\left(D-\sum_{k\in S}\beta_k\right)_i}\dfrac{\sum_{k\notin S}\pi_k}{\left(1-\sum_{m\in S}f_m\right)}, i\in S$ , which may be transformed to

$$(6.3)\qquad \beta_i\left(1-\sum_{m\in S}f_m\right)+Df_i=\pi_i\left(D-\sum_{k\in S}\beta_k\right)\frac{\left(1-\sum_{m\in S}f_m\right)}{\sum_{k\notin S}\pi_k}, i\in S.$$

We sum these equations and we transform the obtained equation into

$$(6.4)\qquad \left(1-\sum_{m\in S}f_m\right)=\frac{D\sum_{k\notin S}\pi_k}{\left(D-\sum_{k\in S}\beta_k\right)}$$

and we substitute this into (6.3) and we obtain the solution after some simplifications:

$$(6.5)\qquad f_i^{opt}=\pi_i-\beta_i\frac{\sum_{k\notin S}\pi_k}{\left(D-\sum_{k\in S}\beta_k\right)}\ \ if\ \ i\in S,\quad f_j^{opt}=0\ if\ \ j\notin S.$$

When (6.2) & (6.4) are substituted into the remaining equations of the group (5.1), corresponding to $j\notin S$ , then these take the following form:

$$(6.6)\qquad \frac{\sum_{k\notin S}\pi_k}{\left(D-\sum_{k\in S}\beta_k\right)}\beta_j=\pi_j+\beta_j\lambda_j\frac{\sum_{k\notin S}\pi_k}{\left(D-\sum_{k\in S}\beta_k\right)},\ j\notin S.$$

With the help of (6.5) & (6.6) the inequalities (5.4) & (5.5) may be rewritten as follows:

$$(6.7)\qquad \pi_j\left(Q_j+1\right)=\frac{D\pi_j}{\beta_j}\le R(S)<\frac{D\pi_i}{\beta_i}=\pi_i\left(Q_i+1\right)\qquad \text{for all i} \in \text{S and all j} \notin \text{S.}$$

where the right-hand side inequalities are equivalent to $0<f_i^{opt}$ and these imply the nontrivial part of (5.5), while the left-hand side inequalities are equivalent to nontrivial part of (5.4). Here

$$(6.8)\qquad R(S)=\left(1-\sum_{m\in S}f_m\right)=\frac{D\sum_{k\notin S}\pi_k}{\left(D-\sum_{k\in S}\beta_k\right)}\ \ if\ \ S\neq\varnothing\ \ and\ \ R(\varnothing)=1$$

is the reserve rate, that is the fraction of horseplayer's wealth that is not bet on any horse. With the help of (6.5) & (6.8) the formula (3.1) may be rewritten as follows:

Explicit solution for optimal betting,

(6.9) $$\ln(G^{opt}) = L(f^{opt}) = \sum_{k\in S} \pi_k \ln\left(\frac{D\pi_k}{\beta_k}\right) + \sum_{k\notin S} \pi_k \ln(R(S)).$$

## 7. The optimal set of horses and the optimal allocations

In order to find the solution of the optimality system one needs to find an optimal set $S^{opt} \subseteq \{1, 2, \ldots, n\}$ such that all inequalities (6.7) and $0 \le R(S^{opt}) = 1 - \sum_{k\in S} f_k \le 1$ are satisfied. $0 \le R(S^{opt}) \le 1$ is equivalent to $\sum_{k\in S} \beta_k \le \sum_{k\in S} \pi_k \le D$, see (6.8). If the only set satisfying these inequalities is $S^{opt} = \varnothing$, then one should not bet at all. Generally, *a set* $S \subseteq \{1, 2, \ldots, n\}$ *is optimal set if for every* $i \in S$ *the expected revenue rate* $er_i > R(S)$ *and for every* $j \notin S$ *the expected revenue rate* $er_i \le R(S)$. From it follows that *there is always exactly one optimal set of horses, possibly empty* (the proofs are omitted). This is the optimal set of horses to bet on.

The optimal set $S^{opt}$ of horses may be obtained with the help of the following algorithm.

1 Calculate expected revenue rates $er_i$ according to formula (2.3).

2 Reorder the horses so that the sequence $er_i$ is non-increasing.

3 Set $S = \varnothing$, $i = 1$, $R = 1$.

4 Repeat:

if $er_i > R$ then insert i-th horse into the set S and

recalculate R according to the formula (6.8)

else set $S^{opt} = S$ and stop repetition.

The set $S^{opt}$ produced by this algorithm is the optimal set of horses. When the optimal set of horses is determined then the optimal allocation fractions should be determined from formulas (6.5) and the maximum of the average of logarithm of the growth rate from the formula (6.9).

This algorithm may be easily implemented in any programming language for use at the racetrack or for Internet wagering. The next section contains two examples of application of this algorithm.

# 8. Examples

In each example the horses are reordered according the expected revenue rates $er_i$, see (2.3), in decreasing order in the first table. Then in the second table the steps 3 & 4 of the algorithm from the preceding section are summarized.

## *Example 1*

The dividend rate is D = .8, there are 7 horses in this race, all the data are in the table:

| Horse | i | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|---|
| Win probability | $\pi_i$ | .003247 | .003247 | .003247 | .01623 | .2273 | .1623 | .5844 |
| Belief probability | $\beta_i$ | .025 | .0375 | .0625 | .125 | .25 | .3125 | .1875 |
| Expected revenue rate | $er_i$ | 0.1039 | .06926 | .04156 | .1039 | .7273 | .4156 | 2.494 |

Therefore, the order of horses according to the expected revenue rates is [7,5,6,1,4,2,3], hence, if the algorithm is run then it produces following list of the candidate sets for the optimal betting set { }, {7}, {7, 5}, {7, 5, 6}, {7, 5, 6, 1}, {7, 5, 6, 1, 4}, {7, 5, 6, 1, 4, 2}, {7, 5, 6, 1, 4, 2, 3}. The steps 3 & 4 of the algorithm from the preceding section are summarized in the following table.

| Horse | i | 7 | 5 | 6 | 1 | 4 | 2 | 3 |
|---|---|---|---|---|---|---|---|---|
| S before including horse i | $S_i$ | { } | { 7 } | { 7, 5 } | {7, 5, 6 } | {7,5,6,1} | - | - |
| Reserve rate | $R(S_i)$ | 1 | .5428 | .4156 | .4156 | .7273 | - | - |
| Expected revenue rate | $er_i$ | 2.494 | .7273 | .4156 | 0.1039 | .1039 | .06926 | .04156 |
| To bet or to not to bet? | | Yes | Yes | No | No | No | No | No |
| Allocation fraction | $f_i^{opt}$ | .4870 | .0974 | 0 | 0 | 0 | 0 | 0 |

The optimal set $S^{opt} = S_5 = \{ 7, 5 \}$of the horses to bet on consists of all the horses for which the value in the row of the expected revenue $er_i$ rates are bigger than those in the row of the reserve rates R(S). For horse 6, $er_6 = .4156 = R(\{7,5\})$, hence this horse and all the other horses with smaller $er_i$ are not included in the optimal set. The logarithm of the maximum possible growth rate is $L(f^{opt}) = .2964$.

Explicit solution for optimal betting,

### *Example 2*

This is the example from [2]. The dividend rate is D = .85, there are 5 horses in this race, all the data are in the table:

| Horse | i | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|
| Win probability | $\pi_i$ | .25 | .1 | .1 | .4 | .15 |
| Belief probability | $\beta_i$ | .17 | .05667 | .034 | .34 | .3993 |
| Expected revenue rate | $er_i$ | 1.25 | 1.5 | 2.5 | 1. | .3193 |

Therefore, the order of horses according to the expected revenue rates is [3,2,1,4,5], hence the sets of the candidate sets for the optimal betting set are { }, {3}, {3,2}, {3,2,4}, {3,2,4,1}, {3,2,4,1,5}. The steps 3 & 4 of the algorithm from the preceding section are summarized in the following table.

| Horse | i | 3 | 2 | 4 | 1 | 5 |
|---|---|---|---|---|---|---|
| S before including horse i | $S_i$ | { } | {3} | {3,2} | {3,2,4} | {3,2,4,1} |
| Reserve rate | $R(S_i)$ | 1 | .9375 | .8955 | .7933 | .5114 |
| Expected revenue rate | $er_i$ | 2.5 | 1.5 | 1.25 | 1. | .3193 |
| To bet or to not to bet? | | Yes | Yes | Yes | Yes | No |
| Allocation fraction | $f_i^{opt}$ | .1477 | .06591 | .07955 | .1955 | 0 |

The optimal set $S^{opt} = S_1$ = {3,2,4,1}of the horses to bet on consists of all the horses for which the value in the row of the expected revenue rates $er_i$ are bigger than those in the row of the reserve rates. The logarithm of the maximum possible growth rate is $L(f^{opt}) = .08736$.

## 9. References


[1] Kelly, J.L.(1956). A New Interpretation of Information Rate, *IRE Transactions on Information Theory*, Vol. IT-2, No. 3, Sept. *Bell System Technical Journal* Vol. 35, pp 917-926.

[2] Enns, Ernest and Tomkins, D. D. (1993). Optimal betting allocations. *The Mathematical Scientist*, Vol. 18, No. 1, pages 37-42

Explicit solution for optimal betting,

[3] S. Boyd and L. Vandenberghe (2004). *Convex Optimization*, Cambridge University Press

[4] James H. Vander Weide; David W. Peterson; Steven F. Maier (1977). A Strategy Which Maximizes the Geometric Mean Return on Portfolio Investments, *Management Science*, Vol. 23, No. 10. June, pp. 1117-1123.

[5] Thomas M. Cover, Joy A. Thomas (1991). *Elements of Information Theory*, John Wiley & Sons

[6] Cover T. (1984). An algorithm for maximizing expected log investment return, *IEEE Transactions on Information Theory*, Vol. 30, issue 2, pp. 369 - 373