Modified Kelly Fraction

To Bet or Not to Bet

Dani Chu, Yifan Wu, and Tim B. Swartz

Department of Statistics and Actuarial Science,
Simon Fraser University

Contact Information:

Dani Chu
dani_chu@sfu.ca
danichusfu.github.io

Yifan Wu
yifan_wu@sfu.ca

Dr. Tim Swartz
tim@stat.sfu.ca
people.stat.sfu.ca/t̃im/

SFU

Abstract

This research considers an extension of the Kelly criterion used in sports wagering. By recognizing that the probability p of placing a correct wager is unknown and often overestimated, modified Kelly criteria are obtained that take the uncertainty into account. Estimators are proposed that are developed from a decision theoretic framework. We observe that the resultant betting fractions can differ markedly based on the choice of loss function. In the cases that we study,the modified Kelly fractions are smaller than original Kelly.

Introduction

- The Kelly criterion (Kelly 1956) provides a gambler with the optimal fraction of a bankroll for wagering on a given bet with probability of winning p.
- Experienced gamblers claim that the Kelly fraction is too high and often leads to financial loss (Murphy 2015).
- The simple but overlooked explanation is that the input p used in determining the Kelly fraction is an unknown quantity.
- In this research we provide a systematic approach for obtaining a "modified Kelly criteria" rather than ad-hoc adjustments such as the half-Kelly or quarter Kelly.

Profitable Systems & Kelly Criterion

If a gambler is successful at making wagers of size $\$x$ with win probability p and European odds θ (where a winning bet returns $\$x \cdot \theta$ and 0 otherwise) then the system is profitable if

$$(-x)(1-p) + (x\theta - x)p > 0 \to p > 1/\theta$$

The optimal fraction of bankroll per wager is known as the Kelly criterion and is given by

$$k(p) = \begin{cases} \frac{p\theta-1}{\theta-1} & p > 1/\theta \\ 0 & p \leq 1/\theta \end{cases}$$

Often sports gamblers are using historical data to estimate p which leads to an estimate of the Kelly criterion $k(\hat{p})$. We therefore introduce a statistical model for the number of winning historical matches X as

$$X \sim \text{Binomial}(n, p)$$

Modified Kelly Criteria

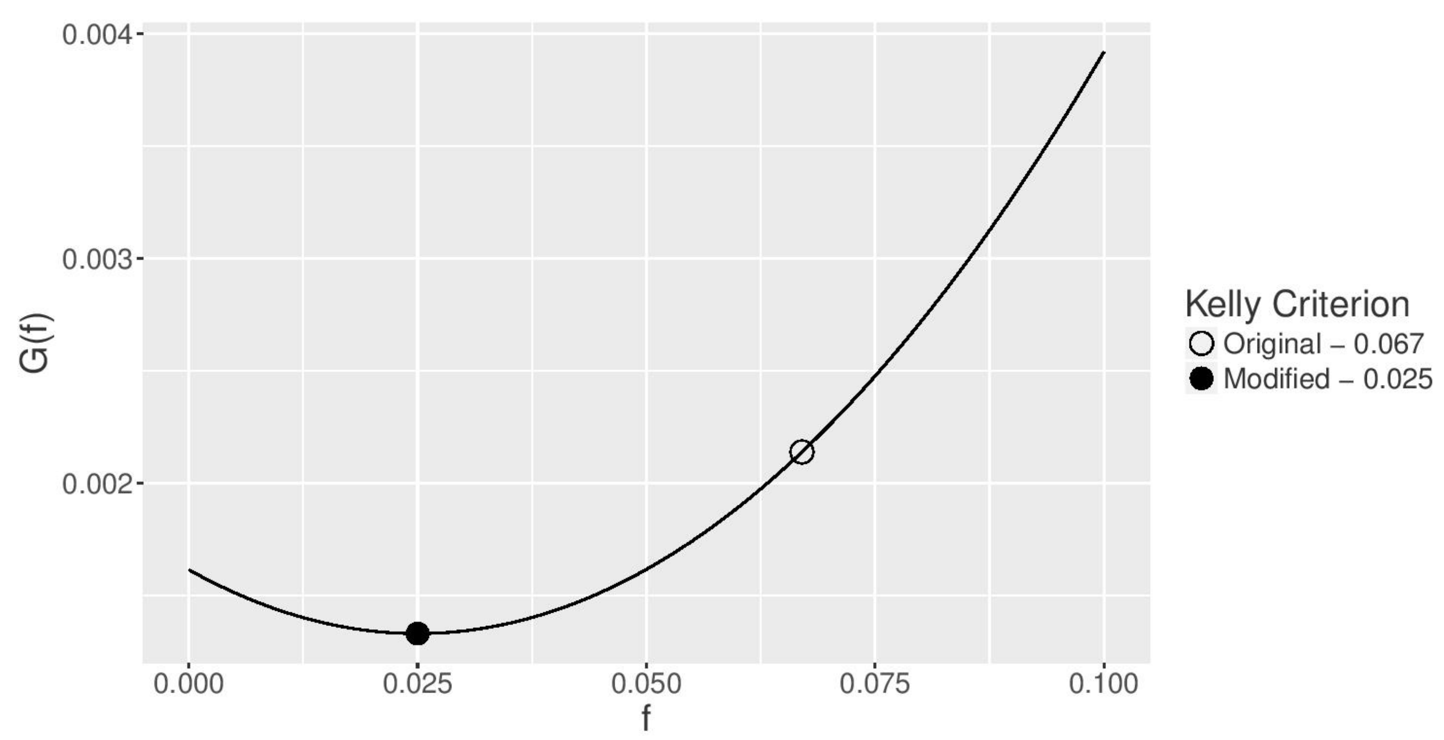

Figure 1: Plot of expected posterior loss G(f) versus f where $\theta = 1.909$, $a = b = 50$, $x = 100$ and $n = 180$

To assess the quality of an estimator we denote $l(f, p)$ as the loss incurred by estimating the true Kelly criterion $k(p)$ with the fraction $f = f(x)$. For initial bankroll B_0 and subsequent bankroll $B_1(f) = (1 - f + \theta f)^w (1-f)^{1-w} B_0$ introduce the loss function

$$l_0(f,p) = E\left[\log\left(\frac{B_1(k(p))/B_0}{B_1(f)/B_0}\right)\right]$$
$$= p\log\left(\frac{1-k(p)+\theta k(p)}{1-f+\theta f}\right) + (1-p)\log\left(\frac{1-k(p)}{1-f}\right)$$

In order to minimize the loss we use a Bayes estimator f which minimizes the Bayes risk. A Bayes estimator will minimize expected posterior loss so we must find f which minimizes

$$G(f) = \int_0^1 l_0(f,p)\pi(p \mid x)dp$$

We use $p \sim \text{Beta}(a, b)$ as a natural prior distribution. Due to the range and the concavity and when empirically defined with historical data (x successes on n trials) gives the posterior distribution

$$p \mid \sim \text{Beta}(x + a, n - x + b)$$

The Bayes estimator f_0 that minimizes Bayes risk is

$$f_0 = \begin{cases} \frac{\hat{p}\theta-1}{\theta-1} & \hat{p} > 1/\theta \\ 0 & \hat{p} \leq 1/\theta \end{cases}$$

where $\hat{p}$ is the posterior mean defined as $\hat{p} = (x+a)/(n+a+b)$

Simulation Results

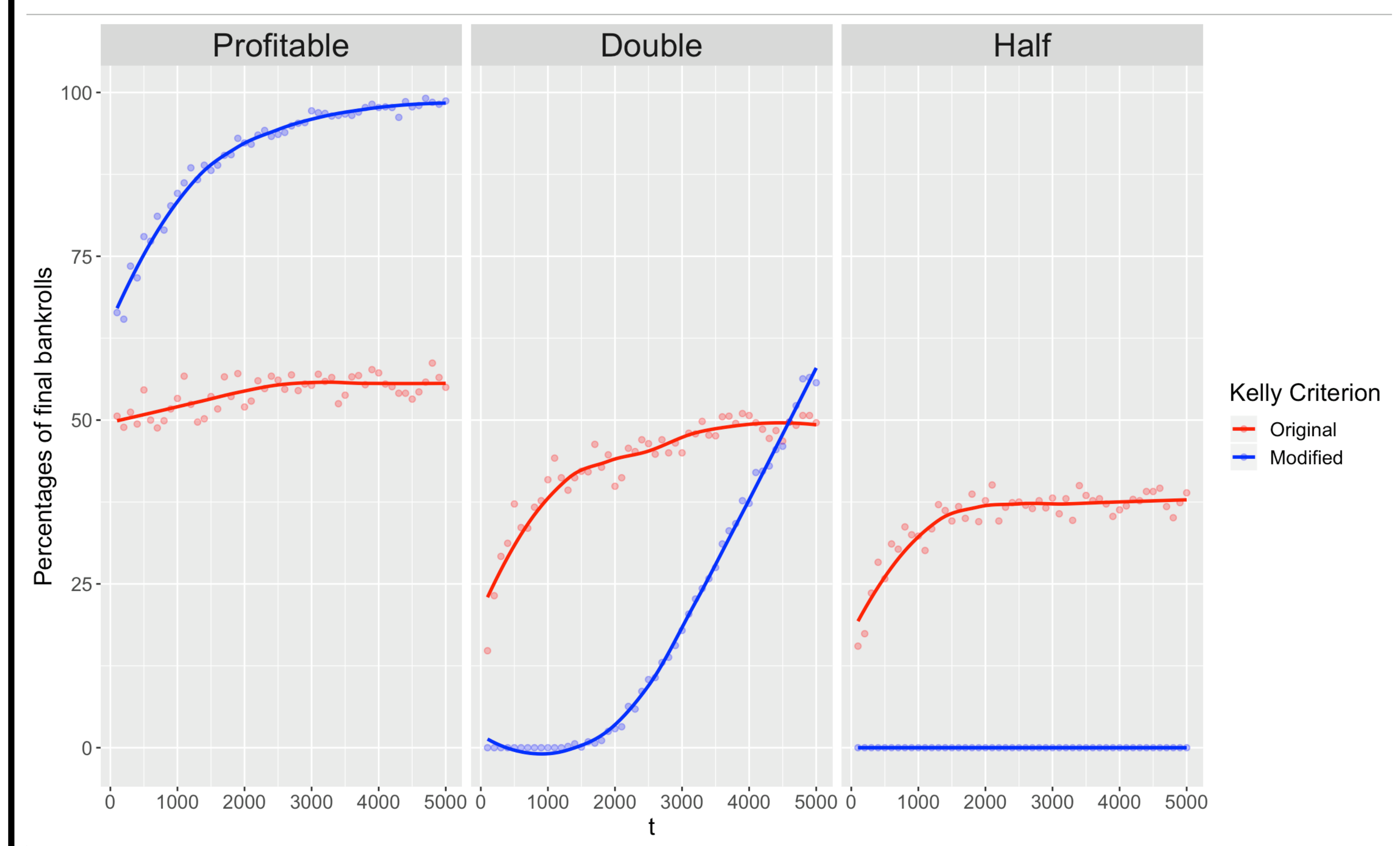

Figure 2: Percentages corresponding to the original Kelly and the modified Kelly schemes f_0 under betting seasons of length 100, ..., 5000 points are smoothed using LOESS curves

Alternative loss functions are also proposed. The selection of a desired loss function can considerably impact the resultant betting fraction.

$$l_1(f,p) = |f - k(p)|$$
$$l_2(f,p) = (f - k(p))^2$$
$$l_3(f,p) = (c_1 I_{f > k(p)} + c_2)|f - k(p)|^k,\ 1 < k < 2, c_1, c_2 > 0$$

Acknowledgements

The authors would like to thank the Natural Sciences and Engineering Research Council of Canada (NSERC) and the Big Data Key Initiative at Simon Fraser University for their financial support. The authors are most appreciative of the contribution of two anonymous reviewers who provided insightful comments and Mohammad Jafari Jozani who provided helpful discussions relating to loss functions.