A Winning Strategy for Roulette

Jerome H. Klotz

University of Wisconsin at Madison

Key words and phrases: Roulette, Bayes strategy, Dirichlet prior, convex loss, expected log return.

Abstract

We examine the statistical problem of computing a favorable betting strategy for the game of roulette with an unbalanced wheel. Using negative average expected log capital after n plays for loss function, we show that the Bayes strategy for a Dirichlet prior is asymptotically optimal. We simulate to illustrate the behavior of the strategy for a biased favorable wheel and an unbiased unfavorable wheel.

1 Introduction and Notation.

Consider the problem of betting on a roulette game with the possibility of an unknown favorable bet due to wheel imbalance. Wilson (1965), presents a nice discussion of the problem with anecdotes about groups exploiting such imbalance to make money and giving data on his own experiences. More recently, Barnhart (1992) gives interesting stories about large casino wins. Ethier (1982), considers the related problem of hypothesis testing for favorable numbers. Of more direct interest to the gambler, is the problem of determining an optimal betting strategy to maximize capital gain.

As a statistical model for roulette outcomes, let independent random column vectors $X_1, X_2, \ldots, X_n$ have a multinomial distribution $\mathcal{M}(1,\ \boldsymbol{p})$ where $X_i = (X_{i1}, X_{i2}, \ldots, X_{iK})^T$, $X_{ik} = 0$, or 1, $\sum_{k=1}^{K} X_{ik} = 1$, $\boldsymbol{p} = (p_1, p_2, \ldots, p_K)^T$, $\sum_{k=1}^{K} p_k = 1$. Thus with probability p_k, $X_{ik} = 1$ if the ball falls into cell number k on the ith play, and $X_{ik} = 0$ with probability $1 - p_k$ if it does not.

The ith play payoff for \$1.00 bet on cell number k is M_k dollars for a win ($X_{ik} = 1$) and $-\$1.00$ for a loss ($X_{ik} = 0$). For roulette in the US, there usually are a total of $K = 38$ cells for a wheel labeled $00, 0, 1, 2, \ldots, 36$ with payoff $M_k = \$35$ for all single number bets.

Starting with initial dollar capital, C_0, let C_n be the capital at the end of the nth bet, and let the strategy be the column vector $\boldsymbol{\gamma}_n = (\gamma_{n0}, \gamma_{n1}, \ldots, \gamma_{nK})^T$ where γ_{n0} is the proportion of C_{n-1} that is not bet and γ_{nk} for $k = 1, 2, \ldots, K$ is the proportion bet on cell number k for the nth gamble with $\sum_{k=0}^{K} \gamma_{nk} = 1$.

The capital at the end of n gambles for this strategy will be

$$C_n = C_{n-1}(\gamma_{n0} + \sum_{k=1}^{K} \gamma_{nk}(M_k + 1)X_{nk}) = C_0 \prod_{i=1}^{n}(\gamma_{i0} + \sum_{k=1}^{K} \gamma_{ik}(M_k + 1)X_{ik}) \tag{1}$$

Write $X[n] = (X_1, X_2, \ldots, X_n)$ and assume $\boldsymbol{\gamma}_i = (\gamma_{i0}, \gamma_{i1}, \ldots, \gamma_{iK})^T$ only depends on $(X_1, X_2, \ldots, X_{i-1})$ ($\mathcal{F}(X[i-1])$ measurable).

We first discuss the solution to the probability problem of finding the optimal strategy $\boldsymbol{\gamma}_n$, for known $\boldsymbol{p}$, to maximize the expected log capital as given by Kelly (1956). Kelly's use of the log penalizes extreme wagers that bet everything on proper subsets and prevents bankruptcy. We then consider the statistical problem of finding a strategy for unknown $\boldsymbol{p}$ and examine the properties of the Bayes strategy for balanced Dirichlet priors.

2 An Optimal Strategy for Known $(p_1, p_2, \ldots, p_K)$.

When the true cell frequencies $\boldsymbol{p}$ are known, Kelley (1956), gives the optimal strategy. Breiman (1960), and Finkelstein and Whitley (1981) also discuss the problem and give limit theorems. Because of the lack of detail in Kelley's derivation we present the solution in our notation.

For known $\boldsymbol{p}$ an optimal strategy $\boldsymbol{\gamma}_n = \boldsymbol{\gamma}(\boldsymbol{p})$ is independent of n. For our loss and unit initial capital, we equivalently maximize the expected log return for a single gamble

$$\phi(\boldsymbol{\gamma}) = \sum_{k=1}^{K} p_k ln(\gamma_0 + (M_k + 1)\gamma_k)$$

subject to the inequality constraints $\gamma_k \geq 0$ for $k = 0, 1, \ldots, K$ and equality constraint $\sum_{k=0}^{K} \gamma_k = 1$.

Using the Kuhn Tucker theorem (see for example Mangasarian (1969) section 7.2.2 or Rockafellar (1970) corollary 28.3.1) we can minimize the Lagrangian

$$\mathcal{L}(\boldsymbol{\lambda}, \boldsymbol{\gamma}) = -\sum_{k=1}^{K} p_k ln(\gamma_0 + \gamma_k(M_k + 1)) - \sum_{k=0}^{K} \lambda_k \gamma_k + \lambda_+(\sum_{k=1}^{K} \gamma_k - 1)$$

where $\boldsymbol{\lambda} = (\lambda_0, \lambda_1, \ldots, \lambda_K, \lambda_+)^T$ is the vector of Lagrange multipliers.

The solution will satisfy the equations

$$\frac{\partial \mathcal{L}(\boldsymbol{\lambda}, \boldsymbol{\gamma})}{\partial \gamma_k} = 0 \tag{2}$$

for $k = 0, 1, \ldots, K$,

$$\sum_{k=0}^{K} \lambda_k \gamma_k = 0 \tag{3}$$

with $\lambda_k \geq 0$ for the inequality constraints $k = 0, 1, \ldots, K$, and

$$\sum_{k=0}^{K} \gamma_k = 1 \tag{4}$$

for the equality constraint. Solving equation (2) $k = 1, 2, \ldots, K$ for γ_k in terms of γ_0 and $\boldsymbol{\lambda}$ gives

$$\gamma_k = \frac{p_k}{\lambda_+ - \lambda_k} - \frac{\gamma_0}{(M_k + 1)}. \tag{5}$$

If $\gamma_k > 0$, then $\lambda_k = 0$ because of equation (3) and then

$$\gamma_k = \frac{p_k}{\lambda_+} - \frac{\gamma_0}{(M_k + 1)}.$$

If we define $\mathcal{G}_+ = \{k \ : \ 1 \leq k \leq K, \ \gamma_k > 0\}$ then substituting these γ_k values in equation (4) gives

$$\gamma_0 = \frac{1 - \sum_{k \in \mathcal{G}_+} p_k / \lambda_+}{1 - \sum_{k \in \mathcal{G}_+} (M_k + 1)^{-1}}. \tag{6}$$

For $\gamma_0 > 0$ equation (3) gives $\lambda_0 = 0$. Using equation (2) for $k = 0$, after substituting the above value of γ_0 we obtain $\lambda_+ = 1$.

When $\gamma_k = 0$, equation (5) with $\lambda_+ = 1$ gives

$$\lambda_k = 1 - p_k(M_k + 1)/\gamma_0 \geq 0 \tag{7}$$

so $\gamma_k = 0$ implies $\gamma_0 \geq p_k(M_k + 1)$. Then for this solution

$$\phi(\boldsymbol{\gamma}) = \sum_{k \in \mathcal{G}_+} p_k ln(p_k(M_k + 1)) + p_0 ln(\gamma_0)$$

where $p_0 = 1 - \sum_{k \in \mathcal{G}_+} p_k$.

To determine the set $\mathcal{G}_+ = \{k \ : \ 1 \leq k \leq K, \ \gamma_k > 0\}$ first sort the values and relabel them so that $p_1(M_1 + 1) \geq p_2(M_2 + 1) \geq \ldots \geq p_K(M_K + 1)$. We then have two cases.

If $1 \geq p_1(M_1 + 1)$ then $p_k \leq 1/(M_k + 1)$ for all $k = 1, 2, \ldots, K$. It follows from equation (6) with $\lambda_+ = 1$ that $\gamma_0 \geq 1 \geq p_k(M_k + 1)$ for all $k = 1, 2, \ldots, K$ and $\phi(\boldsymbol{\gamma})$ will be maximized for $\mathcal{G}_+$ empty ($\gamma_k = 0$ for $k = 1, 2, \ldots, K$, $\gamma_0 = 1$). Thus for this case we don't bet.

If $1 \leq p_1(M_1 + 1)$ we prove that if $k \in \mathcal{G}_+$ then $j \in \mathcal{G}_+$ for $1 \leq j \leq k$. Assume, to the contrary, that $j \notin \mathcal{G}_+$ for $j < k \in \mathcal{G}_+$. Then $\gamma_j = 0$ and $\gamma_0 \geq p_j(M_j + 1)$ by equation (7). But $p_j(M_j + 1) \geq p_k(M_k + 1)$ so $\gamma_0 \geq p_k(M_k + 1)$ and $0 < \gamma_k = p_k - \gamma_0/(M_k + 1) \leq 0$, a contradiction.

Thus $\mathcal{G}_+ = \{1, 2, \ldots, r\}$ where $r \geq 1$ is the largest integer with $1 - \sum_{k=1}^{r} (M_k + 1)^{-1} > 0$ and

$$p_r(M_r + 1) > \gamma_0[r] \tag{8}$$

where $\gamma_0[k] = (1 - \sum_{j=1}^{k} p_j)/(1 - \sum_{t=1}^{k}(M_t+1)^{-1})$. Then with $\gamma_0 = \gamma_0[r]$

$$\gamma_k = p_k - \gamma_0/(M_k+1) \tag{9}$$

for $k = 1, 2, \dots, r$, and $\gamma_k = 0$ for $k = r+1, \dots, K$. We also have the maximum

$$\phi(\boldsymbol{\gamma}) = \sum_{k=1}^{r} p_k ln(p_k(M_k+1)) + p_0 ln(\gamma_0) \tag{10}$$

where $p_0 = \sum_{k=r+1}^{K} p_k$. Call this maximum $\phi^*(\boldsymbol{p})$.

3 Average Return EC_n and Variance.

For known $\boldsymbol{p}$, the optimal strategy is constant so that for a single play ($n = 1$), we have

$$E(C_1/C_0) = E(\gamma_0 + \sum_{k=1}^{K} \gamma_k(M_k+1)X_{1k}) = \gamma_0 + \sum_{k=1}^{K} p_k\gamma_k(M_k+1))$$

Using independence, for general n,

$$E(C_n/C_0) = E\prod_{i=1}^{n}(\gamma_0 + \sum_{k=1}^{K} \gamma_k(M_k+1)X_{ik}) = [E(C_1/C_0)]^n. \tag{11}$$

Similarly

$$E[(C_1/C_0)^2] = \sum_{k=1}^{K} p_k(\gamma_0 + \gamma_k(M_k+1))^2$$

and

$$E[(C_n/C_0)^2] = \{E[(C_1/C_0)^2]\}^n.$$

It follows that

$$Var(C_n/C_0) = \{E[(C_1/C_0)^2]\}^n - \{[E(C_1/C_0)]^2\}^n = \{E(C_1/C_0)\}^{2n}[(1+\Delta_1^2)^n - 1] \tag{12}$$

where

$$\Delta_1 = \sigma(C_1/C_0)/E(C_1/C_0) = \frac{[\sum_{k=1}^{K} p_k\gamma_k^2(M_k+1)^2 - (\sum_{k=1}^{K} p_k\gamma_k(M_k+1))^2]^{1/2}}{[\gamma_0 + \sum_{k=1}^{K} p_k\gamma_k(M_k+1)]}$$

is the coefficient of variation for a single play.

The standard deviation can be quite large in comparison to the expectation so that often very large or very small returns occur after a period of play.

Some insight into the variation can be obtained from the central limit theorem for independent, identically distributed random variables. Taking logs in equation (1) for the optimal strategy $\boldsymbol{\gamma}_n = \boldsymbol{\gamma}$, we can write

$$C_n = C_0 e^{n\phi^*(\boldsymbol{p}) + \sqrt{n}\sigma Z_n}$$

- a large sample log-normal representation where $Z_n \to \mathcal{N}(0,1)$ in distribution as $n \to \infty$. Here

$$\sigma^2 = \sum_{k=1}^{K} p_k ln^2(\gamma_0 + \gamma_k(M_k+1)) - [\sum_{k=1}^{K} p_k ln(\gamma_0 + \gamma_k(M_k+1))]^2.$$

4 An Example.

To illustrate the strategy calculation and its expected return for a known biased model with a small number of cells ($K = 4$), consider bets on green, the first dozen, second dozen, and third dozen (see for example Scarne (1961) page 365). Specifically, we combine green outcomes $\{0, 00\}$ into a cell with index $k = 1$, outcomes $\{1, 2, \ldots, 12\}$ into a cell with index $k = 2$, outcomes $\{13, 14, \ldots, 24\}$ into a cell with index $k = 3$, and outcomes $\{25, 26, \ldots, 36\}$ into a cell with index $k = 4$.

The payoff amounts for a \$1.00 bet are $M_1 = \$17.00$ for green, with $M_2 = M_3 = M_4 =$ \$2.00 for each of the dozens.

Let the true cell frequencies be $\boldsymbol{p} = (3/38, 14/38, 12/38, 9/38)^T$ as compared to $(2/38, 12/38, 12/38, 12/38)^T$ for a completely balanced wheel. Table 1 gives the optimal strategy calculation in this case.

Table 1. $\boldsymbol{\gamma}$ computation.

k	p_k	M_k	$p_k(M_k + 1)$	$\gamma_0[k]$	γ_k
1	3/38	17	54/38≐1.42	315/323≐.98	3/95
2	14/38	2	42/38≐1.42	189/209≐.90	8/95
3	12/38	2	36/38≐0.95	81/95≐.85	3/95
4	9/38	2	27/38≐0.71	†	0

†: $\sum_{j=1}^{k}(1 - (M_j + 1)^{-1} < 0, \quad \gamma_0 = \gamma_0[3] = 81/95$

It is interesting to note that $\gamma_3 = 3/95 > 0$ despite $p_3(M_3 + 1) = 36/38 < 1$.

To calculate the expected return for starting capital of $C_0 = \$1000$ after $n = 100$ games we obtain from equation (11)

$$E(C_1/C_0) = \frac{81}{95} + (\frac{3}{38})(\frac{3}{95})18 + (\frac{14}{38})(\frac{8}{95})3 + (\frac{12}{38})(\frac{3}{95})3 \doteq 1.024986$$

$$EC_{100} = \$1000 \times [EC_1/C_0]^{100} \doteq \$7,607.50 \ .$$

With

$$[(\frac{3}{38})(\frac{3}{95})^2 18^2 + (\frac{14}{38})(\frac{8}{95})^2 3^2 + (\frac{12}{38})(\frac{3}{95})^2 3^2 - (\frac{3}{38})(\frac{3}{95})18 + (\frac{14}{38})(\frac{8}{95})3 + (\frac{12}{38})(\frac{3}{95})3)]^{1/2}$$
$$= \sigma(C_1/C_0) \doteq 0.1538721$$

we have the coefficient of variation

$$\Delta_1 = \sigma(C_1/C_0)/E(C_1/C_0) \doteq 0.1507813$$

and the standard deviation from equation (12) is

$$\sigma(C_{100}) = (EC_{100})[(1 + \Delta_1^2)^{100} - 1]^{1/2} \doteq 22,139.21 \ .$$

Note the extreme variability.

5 The Case of Unknown $(p_1, p_2, \ldots, p_K)$.

We now consider the statistical problem of determining a strategy $\boldsymbol{\gamma}[n] = (\boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2, \ldots, \boldsymbol{\gamma}_n)^{(K+1)\times n}$ for n consecutive gambles, to maximize, for unknown fixed $\boldsymbol{p}$, the expected log return

$$E_{X[n]|p} ln(C_n(\boldsymbol{\gamma}[n], X[n])).$$

As a means of deriving a class of interesting strategies, and noting the suggestion by Thomas Kurtz mentioned by Ethier (1982), we consider the Bayes strategy for the Dirichlet prior joint density

$$f_{\mathcal{P}}(\boldsymbol{p}) = \frac{\Gamma(\alpha_+)}{\Gamma(\alpha_1)\Gamma(\alpha_2)\cdots\Gamma(\alpha_K)} p_1^{\alpha_1 - 1} p_2^{\alpha_2 - 1} \cdots p_K^{\alpha_K - 1}$$

where $0 < p_k < 1$, $\sum_{j=1}^{K} p_j = 1$, $\alpha_k > 0$, $\alpha_+ = \sum_{j=1}^{K} \alpha_j$. This is a natural and mathematically convenient prior for the multinomial model and is the conjugate prior generally recommended by Raiffa and Schlaifer (1961) (see for example Wilks (1962) or DeGroot (1970)). The posterior distribution also belongs to this Dirichlet class of distributions.

A Bayes solution will minimize the Bayes risk or maximize the expected log return averaged with respect to the Dirichlet prior:

$$E_{\mathcal{P}} E_{\boldsymbol{X}[n]|\mathcal{P}} ln(C_n(\boldsymbol{\gamma}[n], X[n])). \tag{13}$$

To find the Bayes strategy, rewrite equation (13) as

$$E_{\mathcal{P}} E_{\boldsymbol{X}[n-1]|\mathcal{P}} E_{\boldsymbol{X}_n|\boldsymbol{X}[n-1],\mathcal{P}} \{ln(C_{n-1}) + ln(\gamma_{n0} + \sum_{k=1}^{K} \gamma_{nk}(M_k + 1)X_{nk})\}$$

$$= E_{\boldsymbol{X}[n-1]} E_{\mathcal{P}|\boldsymbol{X}[n-1]} \{ln(C_{n-1}) + \sum_{k=1}^{K} p_k ln(\gamma_{n0} + \gamma_{nk}(M_k + 1))\}$$

$$= E_{\boldsymbol{X}[n-1]} \{ln(C_{n-1}) + \sum_{k=1}^{K} \hat{p}_{nk} ln(\gamma_{n0} + \gamma_{nk}(M_k + 1))\}$$

where

$$\hat{p}_{nk} = E(\mathcal{P}_k | \boldsymbol{X}[n-1]) = \frac{\alpha_k + S_k[n-1]}{\alpha_+ + n - 1}$$

is the posterior expectation of p_k based on $n-1$ observations with $S_k[n-1] = \sum_{i=1}^{n-1} X_{ik}$.

It follows that the ith column of a Bayes strategy $\hat{\boldsymbol{\gamma}}[n] = (\hat{\boldsymbol{\gamma}}_1, \hat{\boldsymbol{\gamma}}_2, \ldots, \hat{\boldsymbol{\gamma}}_n)^{(K+1)\times n}$ is given by equation (9) with p_k replaced by the Bayes estimates $\hat{p}_{ik}$ for $i = 1, 2, \ldots, n$ where $S_k[0] = 0$ for the starting strategy $(i = 1)$.

6 Large Sample Optimality.

For the Bayes strategy, consider the limiting average expected log return over n gambles given by

$$\lim_{n\to\infty} E_{\boldsymbol{X}[n]|\boldsymbol{p}} \frac{1}{n} ln(C_n(\hat{\boldsymbol{\gamma}}[n], \boldsymbol{X}[n]).$$

In theorem 1 we show it attains the expected log return $\phi^*(\boldsymbol{p})$ for the optimal strategy when $\boldsymbol{p}$ is known.

Write

$$E_{\boldsymbol{X}[n]|\boldsymbol{p}} \frac{1}{n} ln(C_n(\hat{\boldsymbol{\gamma}}[n], \boldsymbol{X}[n])$$

$$= \frac{1}{n} ln(C_0) + \frac{1}{n} \sum_{i=1}^{n} E_{\boldsymbol{X}[n]|\boldsymbol{p}} ln(\hat{\gamma}_{i0} + \sum_{k=1}^{K} \hat{\gamma}_{ik}(M_k+1)X_{ik}) \tag{14}$$

and consider the nth term in the sum on the right

$$\psi_n = E_{\boldsymbol{X}[n]|\boldsymbol{p}} ln(\hat{\gamma}_{n0} + \sum_{k=1}^{K} \hat{\gamma}_{nk}(M_k+1)X_{nk}) \tag{15}$$

$$= E_{\boldsymbol{X}[n-1]|\boldsymbol{p}} E_{\boldsymbol{X}_n|\boldsymbol{X}[n-1],\boldsymbol{p}} ln(\hat{\gamma}_{n0} + \sum_{k=1}^{K} \hat{\gamma}_{nk}(M_k+1)X_{nk})$$

$$= E_{\boldsymbol{X}[n-1]|\boldsymbol{p}} \sum_{k=1}^{K} p_k ln(\hat{\gamma}_{n0} + \hat{\gamma}_{nk}(M_k+1)).$$

To first show $\psi_n \to \phi^*(\boldsymbol{p})$ as $n \to \infty$ we prove the following

Lemma 1 *For fixed $\boldsymbol{p}$, Bayes estimator $\hat{\boldsymbol{p}}_n$ and strategy $\hat{\boldsymbol{\gamma}}[n] = (\hat{\boldsymbol{\gamma}}_1, \hat{\boldsymbol{\gamma}}_2, \ldots, \hat{\boldsymbol{\gamma}}_n)^{(K+1)\times n}$ we have*

$$\sum_{k=1}^{K} p_k ln(\hat{p}_{nk}(M_k+1)) \le \sum_{k=1}^{K} p_k ln(\hat{\gamma}_{n0} + \hat{\gamma}_{nk}(M_k+1)) \le \sum_{k=1}^{K} p_k ln(M_k+1).$$

Proof. The right inequality holds since $(\hat{\gamma}_{n0} + \hat{\gamma}_{nk}(M_k+1)) \le (M_k+1)$.

For the left inequality,

$$\sum_{k=1}^{K} p_k ln(\hat{\gamma_{n0}} + \hat{\gamma}_{nk}(M_k+1))$$

$$= \sum_{i=1}^{r} p_{j_i} ln(\hat{p}_{nk}(M_k+1)) + \sum_{k=r+1}^{K} p_k ln(\hat{\gamma}_{n0})$$

$$\ge \sum_{k=1}^{K} p_k ln(\hat{p}_k(M_k+1))$$

since $\hat{\gamma}_{n0} \ge \hat{p}_{nk}(M_k+1)$ for $k > r$ where r is defined in equation (8) with probabilities $\hat{p}_{nk}$. This completes the proof.

Next we prove

Lemma 2 *For $0 < p_k < 1$, $k = 1, 2, \ldots, K$ and the Bayes estimates $\hat{p}_{nk}$ we have*

$$E_{\boldsymbol{X}[n]|\boldsymbol{p}} \sum_{k=1}^{K} p_k ln(\hat{p}_{nk}(M_k + 1)) \rightarrow \sum_{k=1}^{K} p_k ln(p_k(M_k + 1))$$

as $n \rightarrow \infty$.

Proof. By additivity, it suffices to prove $Eln(\hat{p}_{nk}) \rightarrow ln(p_k)$ as $n \rightarrow \infty$.

Choose ε, $0 < \varepsilon < p_k$ and define the set $A_{nk} = \{\omega \ : \ S_k[n-1] \leq (n-1)(p_k - \varepsilon)\}$. Using the definition of $\hat{p}_{nk}$ we have

$$0 < \alpha_k/(\alpha_+ + (n-1)) \leq (\alpha_k + S_k[n-1])/(\alpha_+ + (n-1)) = \hat{p}_{nk} < 1$$

and

$$ln\left(\frac{\alpha_k}{\alpha_+ + (n-1)}\right) P(A_{nk}) \leq Eln(\hat{p}_{nk}) I_{A_{nk}} \leq 0. \tag{16}$$

For $s_n = \lfloor (n-1)(p_k - \varepsilon) \rfloor$ the greatest integer not exceeding $(n-1)(p_k - \varepsilon)$, we have

$$\binom{n-1}{s_n} p_k^{s_n}(1-p_k)^{n-1-s_n} \leq P(A_{nk}) \leq \binom{n-1}{s_n} p_k^{s_n}(1-p_k)^{n-1-s_n} \left(\frac{(n-s_n)p_k}{np_k - s_n}\right)$$

using Feller (1950 page 140, (3.6)). It follows as $n \rightarrow \infty$ that

$$ln\left(\frac{\alpha_k}{\alpha_+ + (n-1)}\right) P(A_{nk}) \rightarrow 0 \ \text{ and } \ Eln(\hat{p}_{nk}) I_{A_{nk}} \rightarrow 0$$

using (16).

On the complement set A_{nk}^c we have $S_k[n-1] > (n-1)(p_k - \varepsilon)$ so that

$$1 \geq \hat{p}_{nk} = \frac{S_k[n-1] + \alpha_k}{n - 1 + \alpha_+} > \frac{(n-1)(p_k - \varepsilon) + \alpha_k}{n - 1 + \alpha_+}$$

$$\geq \frac{p_k - \varepsilon}{1 + \alpha_+/(n-1)} \geq (p_k - \varepsilon)/2$$

for n sufficiently large. Thus $ln(\hat{p}_{nk})$ is bounded on A_{nk}^c and we apply the dominated convergence theorem to obtain

$$Eln(\hat{p}_{nk}) I_{A_{nk}^c} \rightarrow ln(p_k)$$

since $\hat{p}_{nk}$ converges almost surely to p_k by the strong law of large numbers and $I_{A_{nk}^c} \rightarrow 1$ almost surely. Finally

$$Eln(\hat{p}_{nk}) = Eln(\hat{p}_{nk}) I_{A_{nk}} + Eln(\hat{p}_{nk}) I_{A_{nk}^c} \rightarrow 0 + ln(p_k)$$

to finish the proof of lemma 2.

Now using the lemmas we conclude

Theorem 1 *For fixed p_k, $0 < p_k < 1$ and Bayes strategy $\hat{\boldsymbol{\gamma}}[n]$ we have $\psi_n \to \phi^*(\boldsymbol{p})$ as $n \to \infty$, where ψ_n and ϕ^* are defined by equations (15) and (10), and*

$$\lim_{n\to\infty} E_{\boldsymbol{X}[n]|\boldsymbol{p}} \frac{1}{n} ln(C_n(\hat{\boldsymbol{\gamma}}[n], \boldsymbol{X}[n]) = \phi^*(\boldsymbol{p}).$$

Proof. Since $\hat{p}_{nk}$ converges almost surely to p_k, it follows that a Bayes strategy $\hat{\boldsymbol{\gamma}}_n = \boldsymbol{\gamma}(\hat{\boldsymbol{p}}_n)$, using equations (9) and (10), converges almost surely to an optimal strategy for $\boldsymbol{p}$ known

$$\hat{\boldsymbol{\gamma}}_n \to \boldsymbol{\gamma}(\boldsymbol{p})$$

as $n \to \infty$ since the equations are continuous in $\boldsymbol{p}$. Thus the function

$$h_n = \sum_{k=1}^{K} p_k ln(\hat{\gamma}_{n0} + \hat{\gamma}_{nk}(M_k + 1)) \to \phi^*(\boldsymbol{p})$$

almost surely since it is also a continuous function. If we denote the random function $g_n = \sum_{k=1}^{K} p_k ln(\hat{p}_{nk}(M_k + 1))$ and the constant $G = \sum_{k=1}^{K} p_k ln(M_k + 1)$, then by lemma 1, $g_n \leq h_n \leq G$. Since $g_n \to g = \sum_{k=1}^{K} p_k ln(p_k(M_k + 1))$, and by lemma 2, $Eg_n \to g$, we can apply the theorem of Pratt (1960) to conclude $\psi_n = Eh_n \to \phi^*(\boldsymbol{p})$.

For the second part of the theorem, the average expected log return also converges to $\phi^*(\boldsymbol{p})$ using $n^{-1}ln(C_0) \to 0$ and Toeplitz lemma (see for example Ash (1972 page 270)) applied to equation (14).

7 Performance of the Bayes Strategy.

For single number bets, symmetry considerations suggest using prior parameters the same for all cell numbers ($\alpha_k = \alpha$). If α is selected to be large, the Bayes strategy observes for quite a few games without betting. If the wheel is favorably biased, eventually the Bayes estimates $\hat{p}_{nk}$ will discover this after many games and betting will begin. If a small α is used, chance fluctuations in the counts lead to early betting on unfavorable cells resulting in capital near zero.

Because of the complexity of C_n when $\boldsymbol{\gamma}[n]$ depends on $\boldsymbol{X}[n-1]$, we simulate to get some idea of performance. Table 2 gives results for n plays using α as the prior parameter and C_0 as the initial capital. Sample averages and sample standard deviations

$$\bar{C}_n = \sum_{j=1}^{R} C_n(j)/R, \quad S(C_n) = [\sum_{j=1}^{R}(C_n(j) - \bar{C}_n)^2/(R-1)]^{1/2}$$

for $R = 10,000$ samples are given where $C_n(j)$ is the capital at the end of n plays for the jth sample replication. A biased wheel with $p_1 = 1/30$, $p_2 = \ldots = p_{38} = 29/1110$ was used. This degree of bias in the wheel is consistent with the estimates for wheel imbalance discussed by Wilson (1965).

The multiplicative random number generator $a_j = (a_{j-1} \times 69069)\ mod\ 2^{32}$ was used with the jth random number on $(0,1)$ given by $x_j = a_j/2^{32}$. The period is $2^{32}/4 = 1,073,741,824$ (see for example Marsaglia (1972)). A different starting seed integer $a_0 \neq 0$ was used to compute each entry.

Table 2. Bayes Strategy Simulation.
$\boldsymbol{p} = (1/30, 29/1110, \ldots, 29/1110)$, $M_k = \$35$, $C_0 = \$1000$, $R = 10,000$.

$\bar{C}_n$ ($S(C_n)$)	$n = 100$	200	500	1000	2000	5000
α =1	20.12	0.43	0.00	0.00	0.00	0.00
	(906.92)	(36.39)	(0.00)	(0.00)	(0.000)	(0.00)
2	199.13	119.09	0.01	0.00	0.00	0.00
	(3348.55)	(7800.21)	(0.51)	(0.00)	(0.00)	(0.00)
5	635.63	336.51	10.00	0.07	0.00	0.01
	(1681.20)	(7460.27)	(272.36)	(1.15)	(0.04)	(0.00)
10	863.79	638.70	156.42	18.07	1.69	0.12
	(613.25)	(1229.23)	(1103.10)	(296.89)	(54.02)	(5.88)
20	966.90	874.76	566.34	252.13	148.70	24.12
	(213.69)	(424.02)	(915.12)	(1468.52)	(5312.37)	(832.58)
50	998.69	991.92	948.89	867.31	877.24	6413.53
	(29.13)	(85.02)	(323.98)	(1652.45)	(4441.26)	(218520.92)
100	999.97	999.86	997.85	993.99	1084.80	10378.17
	(1.48)	(10.55)	(67.40)	(227.77)	(1238.92)	(488461.42)
200	1000.00	1000.00	1000.01	1001.83	1038.90	2005.98
	(0.00)	(0.20)	(3.34)	(30.05)	(230.72)	(3869.52)
500	1000.00	1000.00	1000.00	1000.00	1000.21	1052.87
	(0.00)	(0.00)	(0.00)	(0.18)	(5.44)	(174.80)

Figure 1 gives the histogram of 10,000 values of C_n for $\alpha = 100$, $n = 5000$, $C_0 = \$1000$ using logarithmic scale class intervals and positive counts as given in table 3.

Table 3. Counts for 10,000 values of C_n.

Interval	$[2^6, 2^7)$	$[2^7, 2^8)$	$[2^8, 2^9)$	$[2^9, 2^{10})$	$[2^{10}, 2^{11})$	$[2^{11}, 2^{12})$	$[2^{12}, 2^{13})$	$[2^{13}, 2^{14})$	$[2^{14}, 2^{15})$
Count	25	862	2687	2463	1626	1047	601	328	170
Interval	$[2^{15}, 2^{16})$	$[2^{16}, 2^{17})$	$[2^{17}, 2^{18})$	$[2^{18}, 2^{19})$	$[2^{19}, 2^{20})$	$[2^{20}, 2^{21})$	$[2^{21}, 2^{22})$	$[2^{25}, \infty)$	
Count	78	53	31	12	10	5	1	1	

A log-normal distribution is suggested as in the case of known $\boldsymbol{p}$.

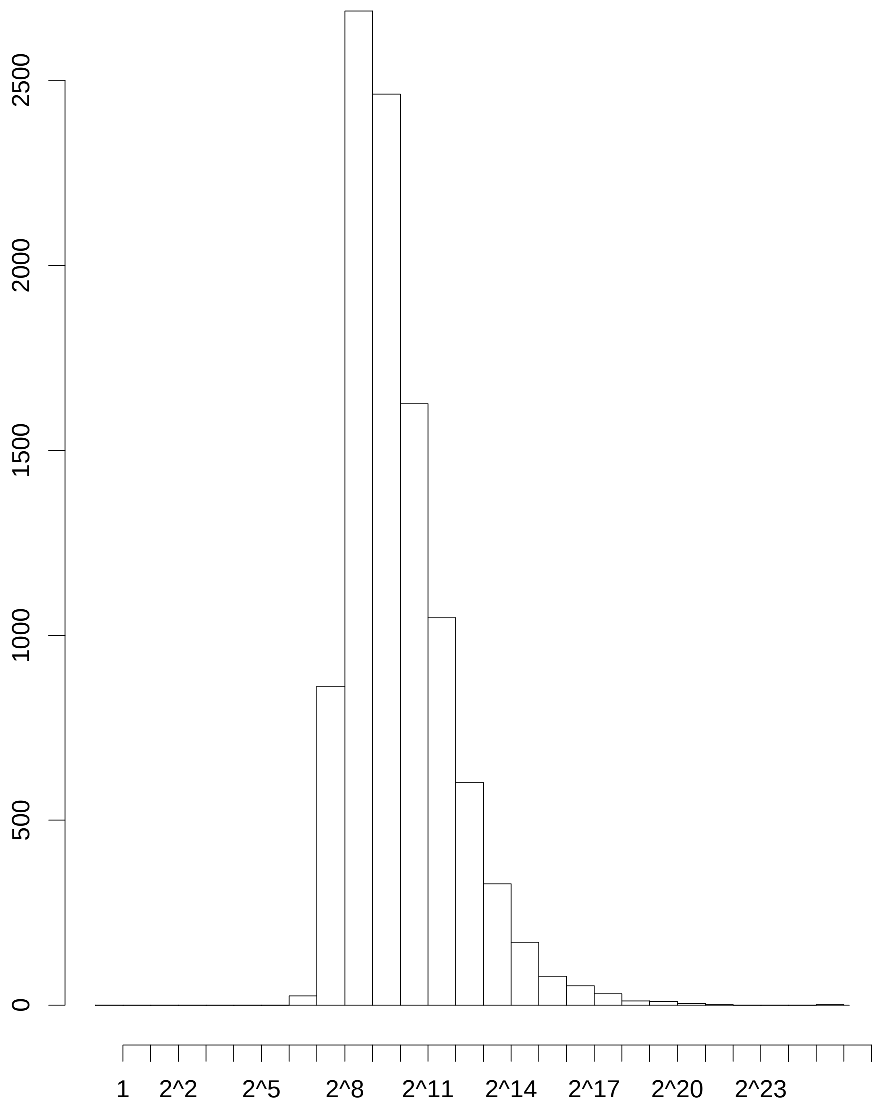

Figure 1. Counts of 10,000 values of C_n with $\alpha = 100$, $n = 5000$, $C_0 = \$1000$.

To compare these results with those for $\boldsymbol{p} = (1/30, 29/1110, \ldots, 29/1110)^T$ assumed known we calculate

$$\boldsymbol{\gamma} = (174/175, 1/175, 0, \ldots, 0)^T$$

and

$$E(C_1/C_0) = (174/175 + (1/30)(1/175)36) \doteq 1.001142857$$

$$\Delta_1 = \frac{[(1/30)(1/175)^2 36^2 - ((1/30)(1/175)36)^2]^{1/2}}{E(C_1/C_0)} \doteq 0.036884690 \ .$$

Table 4 gives corresponding expectations and standard deviations.

Table 4. Optimal Strategy Values.

n	EC_n	$\sigma(C_n)$
100	1121.00	427.79
200	1256.64	702.45
500	1770.22	1746.55
1000	3133.67	5331.35
2000	9819.88	36960.92
5000	302180.01	9039452.88

In addition to results for a biased wheel, it is of interest to see how the Bayes strategy performs in the equiprobable case $\boldsymbol{p} = (1/38, 1/38, \ldots, 1/38)$ when there is no favorable bet ($M_k = \$35$). Table 5 gives $\bar{C}_n$ and $(S(C_n))$ for $R = 10,000$ simulations for some (α, n) values.

Table 5. Bayes Simulation
$p_k = 1/38,\ R = 10,000.$

$\bar{C}_n$ $(S(C_n))$	n =1000	2000	5000
α =50	696.34	340.47	39.19
	(460.16)	(570.81)	(113.28)
100	955.19	810.83	393.26
	(148.20)	(294.36)	(349.26)
200	998.73	987.02	880.10
	(15.15)	(61.25)	(204.69)
500	1000.00	1000.00	999.47
	(0.00)	(0.12)	(7.39)

8 Practical Considerations.

The Bayes strategy assumes that an arbitrary fraction of the capital (e.g. $\$1000 \times 1/175 = \$5.714$) can be bet. In reality there is usually a minimum bet and $C_n \times \gamma_k$ must be an integer multiple of this minimum bet. This restriction should diminish the exponential rate of capital increase for a favorably biased wheel. The effect can be reduced by starting with a large initial capital C_0. In addition, a requirement to bet at least four times the minimum bet on each play as well as a maximum bet limitation further complicates implementing the strategy.

Another difficulty is the computation required to determine bets. It may be not be possible to bring a computer into some casinos although today's palm top portables are unobtrusive. According to Barnhart (1992), the use of an electronic device to aid gambling is a **felony** in Nevada.

The Bayes strategies that do well in table 2 have a long period initially with very little betting. A practical approximation might initially observe the wheel for a long period without betting, if this is permitted, and then use a fixed strategy for $\boldsymbol{p}$ estimated from the initial counts -a "wheel clocking" approach.

A major difficulty is finding a wheel with sufficient favorable bias and avoiding gaffed ones. Even if a favorable wheel is found such as described by Wilson (1965 page 33), there is no guarantee that the casino will not change the wheel if there are large winnings.

We did not include a fixed overhead cost to observe the wheel when no bets are placed (e.g. food, lodging, parking, etc.). Including such an overhead cost would change the Bayes solution and make favorable returns even more difficult.

9 Conclusions.

If true frequencies are known accurately and a favorable bias exists, the optimal strategy expected return increases exponentially with the number of games played (see table 4). When frequencies are unknown, an exponential increase also occurs on favorably biased wheels but a considerable number of trials are required (see table 2). Despite all the difficulties the earning possibilities using prior parameter values in the range $200 \leq \alpha \leq 500$ are quite interesting for 2000 or more plays. However, because of extreme variation, large losses as well as large winnings are possible.

Acknowledgements.

Thanks go to Stewart Ethier for helpful corrections and simplifications to an earlier draft. Thanks also go to Mark Finkelstein, Richard Johnson, Tom Leonard, Bin Yu, and David Blackwell for comments and suggestions for improvement.

References

[1] Ash, Robert B. (1972) *Real Analysis and Probability.* San Diego: Academic Press, Inc.

[2] Barnhart, R.T. (1992) *Beating the Wheel. Winning Strategies at Roulette.* New York: Lyle Stuart.

[3] Breiman, L. (1961) Optimal gambling systems for favorable games. *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability* **I** 65-78. University of California Press

[4] Ethier, S.N. (1982) Testing for favorable numbers on a roulette wheel. *Journal of the American Statistical Association* **77**, 660-665.

[5] Feller, William (1950) *An Introduction to Probability Theory and Its Applications, Vol I.* New York: John Wiley & Sons.

[6] Finkelstein, Mark and Whitley, Robert (1981) Optimal strategies for repeated games.*Advances in Applied Probability* **13**, 415-428.

[7] Kelley, J.L.(1956) A new interpretation of information rate. *Bell System Technical Journal* **36**, 917-926.

[8] Mangasarian, O. L. (1969) *Nonlinear Programming.* New York: McGraw Hill.

[9] Marsaglia, George (1972) The structure of linear congruential sequences, in *Applications of Number Theory to Numerical Analysis.* Edited by S.K. Zaremba. New York: Academic Press, 249-286.

[10] Pratt, John W. (1960) On interchanging limits and integrals. *Annals of Mathematical Statistics* **31**, 74-77.

[11] Raiffa, H. and Schlaifer, R. (1961) *Applied Statistical Decision Theory.* Division of Research, Graduate School of Business Administration, Harvard University, Boston.

[12] Rockafellar, R. T. (1970) *Convex Analysis.* Princeton: Princeton University Press.

[13] Scarne, John (1961) *Scarne's Complete Guide to Gambling.* New York: Simon & Schuster.

[14] Srivastava, M.S and Khatri, C.G. (1979) *An Introduction to Multivariate Statistics.* New York: North Holland.

[15] Wilks, S.S. (1962) *Mathematical Statistics.* New York: John Wiley & Sons.

[16] Wilson, Allan N. (1965) *The Casino Gambler's Guide.* New York: Harper & Row.