Journal of Applied Statistics, Vol. 25, No. 2, 1998, 173–179

CARFAX

Imperfectly shuffled decks in bridge

EDVIN BREDRUP & LI-CHUN ZHANG, *University of Tromsø, Norway*

SUMMARY *In this paper, we study the distribution tables of imperfectly shuffled hands in a bridge game under a conditional Markov chain model, based on which a simple approximate test on the randomness of the decks is derived. The idea is to examine whether a stochastic process is a compound hypergeometric process, through the number of its ties, and it is easily adapted to similar situations.*

1 Introduction

During a bridge game, a player constantly needs to find the optimal strategy, given various situations. It is insufficient to base one's decisions purely on psychological grounds alone, and calculated judgements require knowledge on the distributions of the hands being played. For instance, a classical situation arises if we suppose that N and S have, respectively, $K\times\times\times$ and $AJ109\times$ in ♠. The optimal play is to put down A and K in the first two rounds that ♠ are played, which has a 57.92% chance of clearing the suit, because it fails only when Q is accompanied by more than one card. The same percentage is 56.22% for playing A first and 'finessing' afterwards. Given the fact that a team normally has to play more than 100 games in a tournament, one can hardly afford not to consider a margin of this size. In addition, such knowledge is of vital interest when it comes to the construction of the bidding system.

An important matter in this respect was discussed in Berger (1973), which concerned the differences between computer-dealt hands and those dealt 'manually'. Berger conjectures that the manually dealt hands, presumably imperfectly shuffled, would not follow the standard distribution tables given in the bridge literature. To study the problem, we propose here a conditional Markov chain model, the parameters of which are to be estimated via simulations. A simple test on the randomness of the decks is devised. Moreover, the consequences of the imperfectly shuffled cards are indicated at the end of the paper. In particular, the

Correspondence: E. Bredrup, Institute of Mathematics, University of Tromsø, N-9037 Tromsø, Norway.

0266-4763/98/020173-07 $7.00

underlying idea of our approach is to examine whether a stochastic process is a random process, through the number of its ties, and it can easily be adapted to other similar cases.

2 The theory

2.1 A model for imperfectly shuffled decks

To model the biased hands, we start with the unbiased hands. Throughout the text which follows, we use the notation 0 to denote ♠, 1 for ♡, 2 for ◇ and 3 for ♣. No distinction is made between the cards within the same suit. Let **X** be a vector of 52 coordinates; it is a 'deck' if it contains an equal number of 0, 1, 2 and 3, i.e. 13 of each; and there are

$$C = \frac{52!}{(13!)^4} \tag{1}$$

distinct **X** combinations in total. A deck is 'perfect' if, as one deals out the cards consecutively, the probability of the next card's being from a certain suit is strictly proportional to the number of cards that remain from the same suit. More explicitly, let $m_{i,j}$ be the number of j cards after the ith card is dealt. We have

$$P[X_{i+1} = j \mid (m_{i,0}, m_{i,1}, m_{i,2}, m_{i,3})] = \frac{(13 - m_{i,j})}{(52 - i)} \tag{2}$$

where

$$i = \sum_{j=0}^{3} m_{i,j} \qquad \text{for } 0 \leqslant i \leqslant 51$$

Noticing the similarity between the configurations of a perfect deck and the samples from a compound hypergeometric distribution, we call the process generated by consecutively dealing out the cards of a perfect deck a 'compound hypergeometric process' (CHP). Some aspects of a CHP are worth noticing. Denote by Ω the sample space of **X** of the size C. We have the following.

Lemma 1. $P[\mathbf{X} = \mathbf{x}] = 1/C, \forall \mathbf{x} \in \Omega$.

Lemma 2. A CHP is strictly stationary.

For any $s, r (\geqslant 0)$ and $1 \leqslant i_1 \leqslant i_2 \ldots \leqslant i_s \leqslant 51$ such that $i_s + r \leqslant 52$, and for any $\mathbf{x}_s = (x_{i_1}, \ldots, x_{i_s})$, there are exactly the same number of vectors of **X** that contain $\mathbf{X}_s = \mathbf{x}_s$ as there are that contain $\mathbf{X}s + r = (X_{i_1 + r}, \ldots, X_{i_s + r}) = \mathbf{x}_s$. It follows from Lemma 1 that $P[\mathbf{X}_s = \mathbf{x}_s] = P[\mathbf{X}_{s+r} = \mathbf{x}_s]$. Therefore, a CHP is also second-order weakly stationary. Now, $\forall 1 \leqslant i \leqslant 51$, (x_i, x_{i+1}) is a 'tie' if $x_i = x_{i+1}$. Let $T_i = 1$ if (x_i, x_{i+1}) is a tie; and 0 otherwise. Let K be the total number of ties in a deck, i.e. $K = \sum_{i=1}^{51} T_i$. We have the following.

Lemma 3. Given a CHP, $E[K] = 12$.

It follows from Lemma 2 that $P[X_{i+1} = 0 \mid X_i = 0] = 12/51$, i.e.

$$E[K] = 4 \sum_i P[X_i = X_{i+1} = 0] = 4 \sum_i (1/4)(12/51) = 12$$

Meanwhile, $\mathbf{Y} = (Y_1, \ldots, Y_{52})$ is a 'multinomial process' (MP) if all its coordinates are independent copies of a multinomial random variable with parameters $p_j = P[Y_i = j]$ subject to the constraint $\Sigma_{j=0}^{3} p_j = 1$. Denote by D the case in which $\mathbf{Y}$ happens to be a deck. Let Ω_D be the sample space of $\mathbf{Y}$ conditional to D, i.e. $\Omega_D = \Omega$. We have the following.

Lemma 4. $P[\mathbf{Y} = \mathbf{y} | D] = 1/C, \forall \mathbf{y} \in \Omega_D$.

Lemma 5. The MP is strictly stationary.

Lemma 6. Given an MP, $E[K|D] = 12$, provided that $p_j = 1/4$ for $0 \leqslant j \leqslant 3$.

In short, conditional to D, an MP with equal probabilities is a CHP. Notice that, given a CHP at stage i, the probability of $X_{i+1} = j$ depends on $(X_1, X_2, \ldots, X_i)$ through $m_{i,j}$, rather than through x_i directly. However, this is seldom the case in the reality of bridge. Even when the cards are placed according to the rules, with each player's cards in front of them, the cards usually do not follow each other randomly. This is because it is common that one suit should be played for several consecutive rounds before shifting to another. This suggests that the dependence between two consecutive cards in any imperfectly shuffled decks is likely to be greater than that in a perfect deck. Therefore, consider a Markov chain of $(X_1, X_2, \ldots, X_{52})$, such that

$$P[X_1 = j] = 1/4, \qquad j = 0, 1, 2, 3 \tag{3}$$

and, $\forall 1 \leqslant i \leqslant 51$, we have

$$P[X_{i+1} = j | X_i = r] = p, \qquad \text{if } j = r \tag{4}$$

and

$$P[X_{i+1} = j | X_i = r] = (1 - p)/3, \qquad \text{if } j \neq r \tag{5}$$

We propose to model the imperfect process (IP), generated by consecutively dealing out the cards of an imperfectly shuffled deck, as the Markov chain defined above conditional to D. Notice that we recover the MP with equal probabilities when p is set to be 1/4, and we recover the CHP if D is satisfied in addition. Otherwise, the distribution of the conditional markov chain (CMC) over Ω becomes a function of p, indicating deviation from a CHP. In fact, simulation shows that the CMC is approximately second-order weakly stationary, which means that the tendency towards ties is about the same everywhere within the chain.

2.2 *Parameter estimation*

Assume n mutually independent IPs. Under the CMC model, the log likelihood is given as

$$l(p) = \bar{k} \log p + (51 - \bar{k}) \log(1 - p) - \log f(p) \tag{6}$$

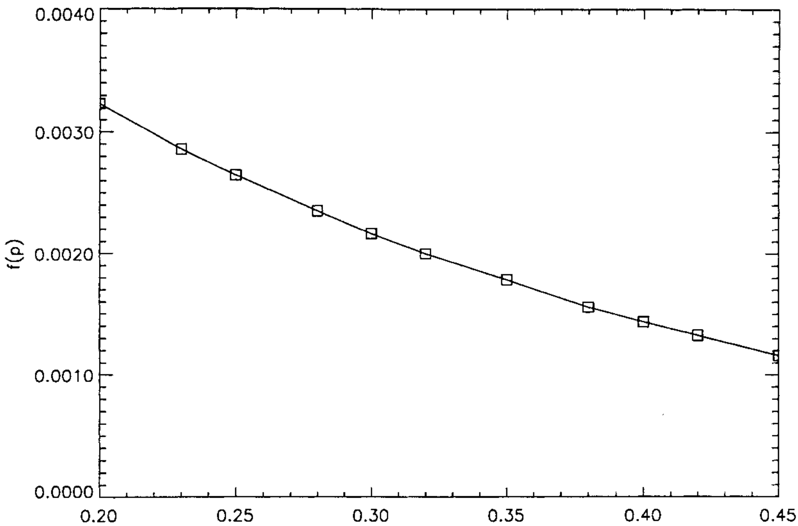

FIG. 1. Simulation results for $f(p) = P[D|p]$. The marked values are estimates based on 50 million MCs generated at each given p.

where

$$\bar{k} = \sum_{i=1}^{n} \frac{k_i}{n}$$

with k_i being the number of ties within the ith deck, and

$$f(p) = P[D|p] \tag{7}$$

It is difficult to obtain the analytical form of $f(p)$. However, it can be estimated from the simulations. Figure 1 shows a plot of the values of $f(p)$ estimated based on simulations from 50 million Markov chains at each marked p. The fitted curve indicates that $f(p)$ is decreasing in p. Notice that $f(p) = 0$ for $p = 1$, and $f(p) > 0$ even if $p = 0$.

To maximize $l(p)$, one could include the fitted $f(p)$ as known. Alternatively, one could estimate $f(p)$ over a set of fine-grid values of p, and choose the estimate with the largest $l(p)$. To find the suitable interval for the grid of p, we turn to the 'quick-fit' estimate, defined as

$$\hat{p}_q^f = \frac{13\bar{K}}{13\bar{K} + 12(51 - \bar{K})} \tag{8}$$

This can be motivated as follows. Given n independent Markov chains, $E[\bar{K}] = 51p$. Conditional to D, we adjust via a weighing mechanism

$$\hat{E}[K] = (51ap)/[(1 - p) + ap], \qquad \text{for } a \in (0, 1)$$

Conforming this to the case of independent CHPs gives us $a = 12/13$.

The distribution of the data belongs to a one-parameter exponential family, so that $\hat{p}$ is asymptotically unbiased, efficient and normally distributed with variance equal to the Cramer–Rao lower bound, i.e.

$$\mathrm{Var}(\hat{p}) = [nI(p)]^{-1} + O\left(\frac{1}{n}\right) \tag{9}$$

where $I(p)$ is the Fisher information from a single deck, i.e.

$$I(p) = E[K]/p^2 + (51 - E[K])/(1-p)^2 + [f''(p)f(p) - f'(p)^2]/f(p)^2 \tag{10}$$

In practice, $E[K]$ can be estimated by $\bar{K}$, whereas an evaluation of the last term on the right-hand side might be obtained by applying polynomial regression to the simulated values of $f(p)$, for example.

2.3 A test

Set up the test on the randomness of the shuffled decks as

$$H_0: p = 1/4, \qquad H_1: p > 1/4$$

Simulations under the null hypothesis gave us $E[\hat{p}_{qf}] \approx 1/4$ and $\mathrm{Var}(\hat{p}_{qf}) \approx 1/(256n)$. In fact, $\hat{p}_{qf}$ is so close to $\hat{p}$ for p around 1/4 that the normal approximation applies to it directly. In other words, we have (approximately)

$$Z = 16n^{1/2}(\hat{p}_{qf} - 1/4) \sim N(0, 1) \tag{11}$$

Thus, a test at level-α rejects H_0 if $Z > z_\alpha$, where z_α is the $100(1-\alpha)\%$ percentile of Z. In contrast, the Taylor expansion after inverting for $\bar{K}$ gives

$$\bar{K} > k_\alpha = 12 + (52z_\alpha)/(17n^{1/2}) + O\left(\frac{1}{n}\right) \tag{12}$$

Simulations have confirmed that the approximate critical region above is correct for $n \geqslant 4$. Moreover, $k_\alpha \to 12$ for $n \to \infty$, as it should according to Lemma 3. It should be noticed that the ideas (1) of simulating the CHP through the conditional multinomial process with equal probabilities, and (2) of examining its 'properness' through the number of ties, are easily adapted to similar situations.

3 Twenty-four actual imperfectly shuffled decks

After one tournament at the Student Bridge Club in Tromsø, one of the authors collected all the 24 decks used that night, shuffled them one by one, and observed that $\bar{K} = 452/24 = 18.83$. The critical region in this case is $\bar{K} > 13.6$ at the 5% level. Maximizing $l(p)$ by simulation gives us $\hat{p} = 0.384$, as compared with the quick-fit $\hat{p}_{qf} = 0.388$, with estimated standard error 4.4×10^{-3}. Further simulations of 50 million Markov chains at $p = 0.384$ have resulted in 76 402 decks giving us the results in Table 1 of the distribution of the suits in one bridge hand.

In particular, when N and S (N–S) hold nine trumps, the distribution of the remaining cards is given as

2–2: 41.8%, 3–1: 50.1%, 4–0: 8.1%

Compare this with the standard table from Kelsey and Glauert (1980) at $p = 0.25$ (Table 2).

TABLE 1. Approximate distribution of a bridge hand given imperfectly shuffled decks[a]

Configuration	Rate (100%)
4–4–3–2	24.6
5–4–3–1	12.3
6–4–2–1	3.65
5–5–2–1	2.60
6–4–3–0	0.92
5–3–3–2	16.6
5–4–2–2	10.7
4–4–4–1	2.97
7–3–2–1	1.25
5–5–3–0	0.62
4–3–3–3	12.9
6–3–2–2	5.01
6–3–3–1	2.88
5–4–4–0	0.96
Others	2.04

[a]Simulated at $p = 0.384$.

TABLE 2. Distribution of a bridge hand given perfectly shuffled decks, where $p = 0.25$

Configuration	Rate (100%)
4–4–3–2	21.6
5–4–2–2	10.6
6–4–2–1	4.70
4–4–4–1	2.99
5–4–4–0	1.24
6–5–2–0	0.65
5–3–3–2	15.5
4–3–3–3	10.5
6–3–3–1	3.45
7–3–2–1	1.88
5–5–3–0	0.90
7–2–2–2	0.51
5–4–3–1	12.9
6–3–2–2	5.64
5–5–2–1	3.17
6–4–3–0	1.33
6–5–1–1	0.71
Others	1.74

In particular, when N and S hold nine trumps, the distribution of the remaining cards is given as

2–2: 40.7%, 3–1: 49.7%, 4–0: 9.6%

Judging from Tables 1 and 2, it seems that imperfectly shuffled decks reduce the probabilities of longer suits. If we return to the problem in the first section, where N and S hold nine cards in one suit, with all the top cards except *Q*, it appears that the optimal play would give a 58.58% probability of success on this occasion. In contrast, the rate is 56.26% for playing *A* first and 'finessing' afterwards. indeed, if the imperfectly shuffled decks do reduce the chances for longer suits as noted, then the greater bias it is, the greater success this gives for the optimal play. (The same may not be said about the finessing.) Note also that, as the bias grows larger, it may be the case that only $\hat{p}$ is admissible, because we did not study in this paper how $\hat{p}_{qf}$ would behave under such conditions. Finally, because there are obviously no 'standard' imperfectly shuffled decks, we refer to Bredrup and Zhang (1993) for the distribution tables of one bridge hand under the various values of p.

REFERENCES

BERGER, P. D. (1973) On the distribution of hand patterns in bridge: man-dealt versus computer-dealt, *Canadian Journal of Statistics*, 1, pp. 261–266.

BREDRUP, E. & ZHANG, L. (1993) On models of imperfectly shuffled decks in bridge, *Preprint*, University of Tromsø.

KELSEY, H. & GLAUERT, M. (1980) *Bridge Odds for Practical Players* (London, Victor Gollance).