Probabilistic parking functions

Mei Yin[1]

Department of Mathematics, University of Denver

November 14, 2022

[1] Joint work with Irfan Durmić, Alex Han, Pamela E. Harris, Rodrigo Ribeiro.

Parking functions were introduced by Konheim and Weiss (1966) under the name of "parking disciplines," in their study of the hash storage structure, and have since found many applications in combinatorics, probability, and computer science.

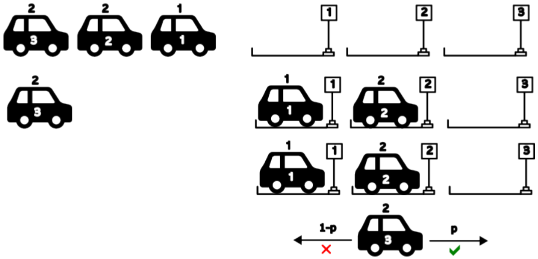
Consider a parking lot with n parking spots placed sequentially along a one-way street. A line of n cars enters the lot, one by one. The ith car drives to its preferred spot π_i and parks there if possible; if the spot is already occupied then the car parks in the first available spot after that. The list of preferences $\pi = (\pi_1, \ldots, \pi_n)$ is called a parking function if all cars successfully park. We denote by PF_n the set of parking functions of length n. It is well-known that $|\mathrm{PF}_n| = (n+1)^{n-1}$.

Introducing probability

In the classical situation, a car always drives forward when the desired spot is taken. Can we change the parking protocol so that a car can move backward sometimes?

Probabilistic parking protocol: Fix $p \in [0, 1]$ and consider a coin which flips to heads with probability p and tails with probability $1 - p$. If a car arrives at its preferred spot and finds it unoccupied it parks there. If instead the spot is occupied, then the driver tosses the biased coin. If the coin lands on heads, with probability p, the driver continues driving forward in the street. However, if the coin lands on tails, with probability $1 - p$, the car moves backward and tries to find an unoccupied parking spot.

The classical situation corresponds to $p = 1$.



Parking scheme for the preference vector $(1, 2, 2)$. The number at the top of the car represents its preference.

With probability p, all cars can park.

Preference symmetry: Having n cars enter the street from left to right with preference vector $\pi = (\pi_1, \ldots, \pi_n)$ and park under protocol with parameter p depicts the same scenario as having n cars enter the street from right to left with preference vector $\pi' = (n+1-\pi_1, \ldots, n+1-\pi_n)$ and park under protocol with parameter $1-p$.

Lack of permutation symmetry: In the case of classical parking functions, a preference vector π is either deterministically a parking function or not, and any permutation of a parking function is a parking function. Contrarily, by incorporating a probabilistic parking protocol, all preference vectors π have a positive probability of being a parking function, and permuting a preference vector might change the probability of it being a parking function.

We use $\pi \in \mathrm{PF}_n$ as a shorthand for the situation that n cars with preference vector π park.

Main Theorem 1

Magic: The probabilities of being a parking function for all preference vectors add up in a way so that dependence on p is canceled and there is invariance to the forward probability p for the randomly selected vector: $\mathrm{P}(\pi \in \mathrm{PF}_n | \pi \in [n]^n) = (n+1)^{n-1}/n^n$, as in the classical situation.

$\pi \in [3]^3$	$(1,1,1)$	$(1,1,2)$	$(1,1,3)$	$(1,2,1)$	$(1,2,2)$	$(1,2,3)$	$(1,3,1)$
$\mathrm{P}(\pi \in \mathrm{PF}_3 \| \pi)$	p^2	p^2	p	p	p	1	p
$\pi \in [3]^3$	$(2,1,1)$	$(2,1,2)$	$(2,1,3)$	$(2,2,1)$	$(2,2,2)$	$(2,2,3)$	$(2,3,1)$
$\mathrm{P}(\pi \in \mathrm{PF}_3 \| \pi)$	p	p	1	$p+(1-p)p$	$2p(1-p)$	$p(1-p)+(1-p)$	1
$\pi \in [3]^3$	$(3,1,1)$	$(3,1,2)$	$(3,1,3)$	$(3,2,1)$	$(3,2,2)$	$(3,2,3)$	$(3,3,1)$
$\mathrm{P}(\pi \in \mathrm{PF}_3 \| \pi)$	p	1	$1-p$	1	$1-p$	$1-p$	$1-p$
$\pi \in [3]^3$	$(1,3,2)$	$(1,3,3)$	$(2,3,2)$	$(2,3,3)$	$(3,3,2)$	$(3,3,3)$	
$\mathrm{P}(\pi \in \mathrm{PF}_3 \| \pi)$	1	$1-p$	$1-p$	$1-p$	$(1-p)^2$	$(1-p)^2$	

All preference vectors for $n = 3$ and their probability of being a parking function.

Outline of proof

1. Extend Pollak's circle argument: Consider the scenario in which n cars park on a circle with spots $[n+1]$, where each car may prefer any spot. Then given an arbitrary preference vector $\pi \in [n+1]^n$, $\pi \in \mathrm{PF}_n$ if and only if spot $n+1$ is vacant after all n cars have parked.

2. Utilize symmetry of the circle: Let $F_a^{(i)}$ denote the expected number of preference vectors which have leading preference a and leave spot i vacant after all n cars have parked. Then $F_a^{(i)} = F_b^{(j)}$ for all $a, b, i, j \in [n+1]$ satisfying $b - a \equiv j - i \pmod{n+1}$.

3. Further implications of the circle symmetry: Construct a $(n+1) \times (n+1)$ matrix, using

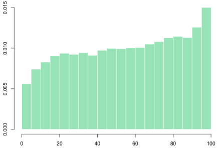
$$\begin{pmatrix} F_1^{(1)} & F_1^{(2)} & \dots & F_1^{(n+1)} \\ F_2^{(1)} & F_2^{(2)} & \dots & F_2^{(n+1)} \\ \vdots & \vdots & \ddots & \vdots \\ F_{n+1}^{(1)} & F_{n+1}^{(2)} & \dots & F_{n+1}^{(n+1)} \end{pmatrix} = \begin{pmatrix} F_1^{(1)} & F_1^{(2)} & \dots & F_1^{(n+1)} \\ F_1^{(n+1)} & F_1^{(1)} & \dots & F_1^{(n)} \\ \vdots & \vdots & \ddots & \vdots \\ F_1^{(2)} & F_1^{(3)} & \dots & F_1^{(1)} \end{pmatrix}$$

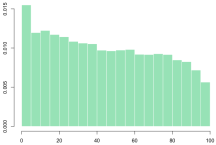
The equivalent formulation shows that every column and row of the matrix has the same sum. We then apply law of total probability.

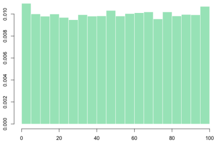
Effect of the forward probability p

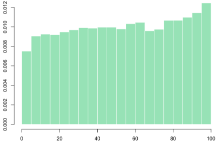
We simulate the conditional distribution of π_n (parking preference of the last car) in $100,000$ samples of preference vectors of size 100 chosen uniformly at random. The upper left plot is for $p = 0$ and the upper right plot is for $p = 1$. The middle plot is for $p = 0.5$. The lower left plot is for $p = 0.25$ and the lower right plot is for $p = 0.75$.

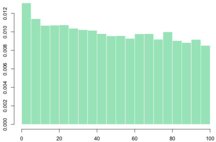
Note the preference symmetry between p and $1 - p$ as observed earlier.


0.000
0.005
0.010
0.015
0
20
40
60
80
100


0.000
0.005
0.010
0.015
0
20
40
60
80
100


0.000
0.002
0.004
0.006
0.008
0.010
0
20
40
60
80
100


0.000
0.002
0.004
0.006
0.008
0.010
0.012
0
20
40
60
80
100


0.000
0.002
0.004
0.006
0.008
0.010
0.012
0
20
40
60
80
100

Main Theorem 2

Consider the preference vector $\pi \in [n]^n$, chosen uniformly at random. Then $\mathrm{P}(\pi_n = j | \pi \in \mathrm{PF}_n)$

$$= \frac{2}{n+1} - \frac{1}{(n+1)^{n-1}} \Big[p \sum_{s=n-j+1}^{n-1} \binom{n-1}{s} (n-s)^{n-s-2} (s+1)^{s-1} + (1-p) \sum_{s=0}^{n-j-1} \binom{n-1}{s} (n-s)^{n-s-2} (s+1)^{s-1} \Big],$$

where π_n denotes the parking preference of the last car.

Preference symmetry: $\mathrm{P}(\pi_n = j | \pi \in \mathrm{PF}_n)$ under protocol with parameter p equals $\mathrm{P}(\pi_n = n+1-j | \pi \in \mathrm{PF}_n)$ under protocol with parameter $1-p$.

Convexity: The conditional distribution of π_n for a generic p can be written as a convex combination of the cases $p = 0$ and $p = 1$.

Outline of proof

1. Cars $1, \dots, n-1$ have all parked along the one-way street before the nth car enters, leaving only one open spot k for the nth car to park. Since a car cannot jump over an empty spot, the parking protocol implies that $(\pi_1, \dots, \pi_{n-1})$ is a parking function shuffle of $\alpha \in \mathrm{PF}_{k-1}$ and $\beta \in \mathrm{PF}_{n-k}$, and α and β do not interact with each other.

2. This open spot k could be either the same as j, the preference of the last car, in which case the car parks directly. Or, k could be bigger than or less than j, in which case whether the last car parks or not depends on the outcome of the biased coin flip, as it will dictate the car to go forward or backward.

3. Derivation of the explicit formula utilizes Abel's binomial theorem.

Parking function shuffle was first discussed in Diaconis and Hicks (2017) and later extended further in Kenyon and Yin (2021). Related work in Adeniran et al. (2020) and Paguyo (2022).

Let $1 \leq k \leq n$. Say that $(\pi_1, \ldots, \pi_{n-1})$ is a parking function shuffle of the parking function $\alpha \in \mathrm{PF}_{k-1}$ and $\beta \in \mathrm{PF}_{n-k}$ if $\pi_1, \ldots, \pi_{n-1}$ is any permutation of the union of the two words α and $\beta + (k, \ldots, k)$.

Example: Take $n = 8$ and $k = 4$. Take $\alpha = (2, 1, 2) \in \mathrm{PF}_3$ and $\beta = (1, 2, 4, 3) \in \mathrm{PF}_4$. Then $(2, \underline{5}, 2, \underline{8}, \underline{7}, 1, \underline{6})$ is a shuffle of the two words $(2, 1, 2)$ and $(5, 6, 8, 7)$.

Abel's extension of the binomial theorem (derived from Pitman (2002) and Riordan (1968)):

Let

$$A_n(x, y; p, q) = \sum_{s=0}^{n} \binom{n}{s} (x+s)^{s+p} (y+n-s)^{n-s+q}.$$

Then

$$A_n(x, y; p, q) = A_n(y, x; q, p),$$
$$A_n(x, y; p, q) = A_{n-1}(x, y+1; p, q+1) + A_{n-1}(x+1, y; p+1, q),$$
$$A_n(x, y; p, q) = \sum_{s=0}^{n} \binom{n}{s} s!(x+s) A_{n-s}(x+s, y; p-1, q).$$

The following special instances hold via the basic recurrences listed above:

$$A_n(x, y; -1, -1) = (x^{-1} + y^{-1})(x + y + n)^{n-1},$$

$$A_n(x, y; -1, 0) = x^{-1}(x + y + n)^n,$$

$$A_n(x, y; -1, 1) = x^{-1} \sum_{s=0}^{n} \binom{n}{s} (x + y + n)^s (y + n - s)(n - s)!,$$

$$A_n(x, y; 0, 0) = \sum_{s=0}^{n} \binom{n}{s} (x + y + n)^s (n - s)!.$$

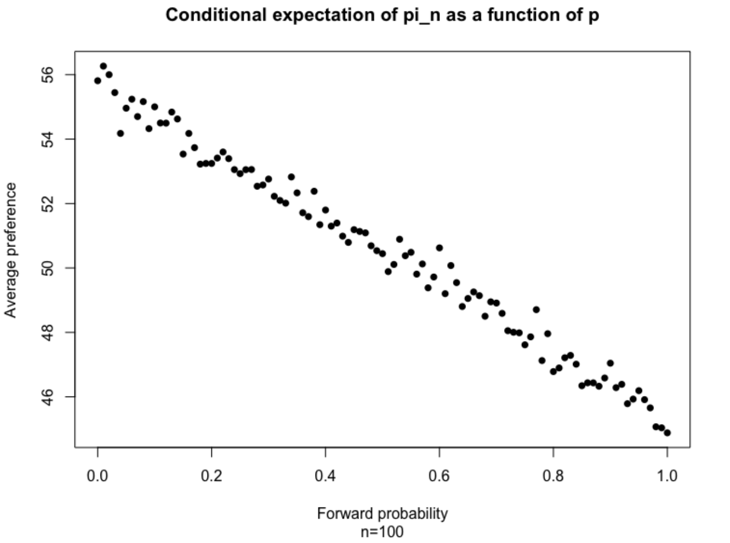
Main Theorem 3

Take n large. For preference vector $\pi \in [n]^n$ chosen uniformly at random, we have

$$\mathbb{E}(\pi_n | \pi \in \mathrm{PF}_n) = \frac{n+1}{2} - (2p-1)\Big[\frac{\sqrt{2\pi}}{4} n^{1/2} - \frac{7}{6}\Big] + o(1).$$

Preference symmetry: The sum of $\mathbb{E}(\pi_n | \pi \in \mathrm{PF}_n)$ under protocol with parameter p and $\mathbb{E}(\pi_n | \pi \in \mathrm{PF}_n)$ under protocol with parameter $1-p$ is $n+1$. $\mathbb{E}(\pi_n | \pi \in \mathrm{PF}_n) = (n+1)/2$ exactly.

Proof uses asymptotic expansion methods: Stirling's, Edgeworth expansion for Poisson random variables.

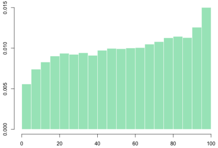


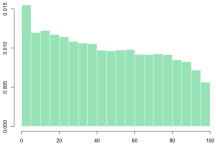
Size $n = 100$ and $100,000$ samples are drawn uniformly at random.

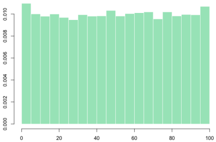
Why the forward probability $p = 1/2$ is special

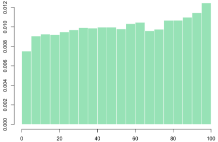
Observation: As p increases from 0 to $1/2$, the distribution of π_n places less and less mass on cars with large parking preferences, as cars are gradually losing their backwards moving bias when their desired spot is taken. Similarly, as p decreases from 1 to $1/2$, the distribution of π_n places less and less mass on cars with small parking preferences, as cars are gradually losing their forward moving bias when their desired spot is taken. For $p = 1/2$, the movement bias of a car when the desired spot is taken is gone because going left and right have the same probability.

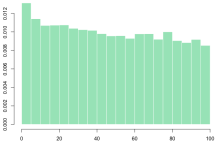
Conclusion: For $p = 1/2$ the symmetry provided by p plays an important role of speeding up the rate of convergence of π_n to a uniform distribution over $\{1, \ldots, n\}$.


0.000
0.005
0.010
0.015
0
20
40
60
80
100


0.000
0.005
0.010
0.015
0
20
40
60
80
100


0.000
0.002
0.004
0.006
0.008
0.010
0
20
40
60
80
100


0.000
0.002
0.004
0.006
0.008
0.010
0.012
0
20
40
60
80
100


0.000
0.002
0.004
0.006
0.008
0.010
0.012
0
20
40
60
80
100

Main Theorem 4

Let $Q_{n,p}(\cdot)$ be $\mathrm{P}(\pi_n = \cdot \mid \pi \in \mathrm{PF}_n)$ under the probabilistic parking protocol with parameter p and Uni_n be the uniform distribution over $\{1, \ldots, n\}$. The following bounds hold:

- For $p \neq 1/2$, $\|Q_{n,p} - \mathrm{Uni}_n\|_{\mathrm{TV}} = \Theta\left(\frac{1}{\sqrt{n}}\right)$.
- For $p = 1/2$, $\|Q_{n,p} - \mathrm{Uni}_n\|_{\mathrm{TV}} = \Theta\left(\frac{1}{n}\right)$.

Here TV denotes total variation (TV) distance. For two distributions P and Q over $\{1, \ldots, n\}$,

$$\|P - Q\|_{\mathrm{TV}} := \frac{1}{2} \sum_{j=1}^{n} |P(j) - Q(j)|.$$

Outline of proof

1. $Q_{n,p}$ is derived in Main Theorem 2, with $Q_{n,1}(j) = Q_{n,0}(n+1-j)$ (preference symmetry) and $Q_{n,p} = pQ_{n,1} + (1-p)Q_{n,0}$ (convexity).

2. We establish a chain of inequalities concerning rate of convergence:

$$|2p-1| \|Q_{n,1} - \mathrm{Uni}_n\|_{\mathrm{TV}} \leq \|Q_{n,p} - \mathrm{Uni}_n\|_{\mathrm{TV}} \leq \|Q_{n,1} - \mathrm{Uni}_n\|_{\mathrm{TV}}.$$

3. We identify a clever choice of test function f in the alternative definition of TV distance:

$$\|P - Q\|_{\mathrm{TV}} = \frac{1}{2} \sup \left\{ \sum_{j=1}^{n} f(j)\,(P(j) - Q(j)) : \max_{j} |f(j)| \leq 1 \right\}.$$

4. We apply Abel's binomial theorem and Main Theorem 3.

Related combinatorial results

(solution to an open problem posed by Novelli and Thibon)
OEIS A220884: Let $P_n(q)$ be the generating polynomial $\prod_{k=2}^{n}[(n+1-k)q+k]$. Then

$$P_n(q) = \sum_{k=0}^{n-1} U_n(k) q^k,$$

where $U_n(k)$ is the number of preference vectors in $[n+1]^n$ with leading preference fixed at spot 1 containing k unlucky cars. This corresponds to the situation when n cars park on a circle with spots $[n+1]$.

1			
2	1		
6	8	2	
24	58	37	6

OEIS A071208: Let $Q_n(q)$ be the generating polynomial $\prod_{k=1}^{n}[(n-k)q+k]$. Then

$$Q_n(q) = \sum_{k=0}^{n-1} V_n(k) q^k,$$

where $V_n(k)$ is the number of preference vectors in $[n]^n$ containing k unlucky cars. This corresponds to the situation when n cars park on a one-way street with spots $[n]$.

Equivalent formulation in terms of leaders in a labeled tree due to Knuth's bijection between parking functions and trees.

1			
2	2		
6	15	6	
24	104	104	24

Further directions for research

1. Other parking statistics? Their correlations?
2. Other modified parking protocols? Higher dimensions?
3. Study of asymptotics in related combinatorial models?

Thank You! Questions?

Further directions for research

1. Other parking statistics? Their correlations?
2. Other modified parking protocols? Higher dimensions?
3. Study of asymptotics in related combinatorial models?

Thank You! Questions?

Main References:

- Persi Diaconis and Angela Hicks. Probabilizing parking functions. Adv. Appl. Math. 89: 125-155 (2017)
- Irfan Durmić, Alex Han, Pamela E. Harris, Rodrigo Ribeiro, and Mei Yin. Probabilistic parking functions. arXiv: 2211.00536 (2022)
- Richard Kenyon and Mei Yin. Parking functions: From combinatorics to probability. arXiv: 2103.17180 (2021)
- Alan G. Konheim and Benjamin Weiss. An occupancy discipline and applications. SIAM J. Appl. Math. 14: 1266-1274 (1966)
- Jim Pitman. Forest volume decompositions and Abel-Cayley-Hurwitz multinomial expansions. J. Comb. Theory Ser. A 98: 175-191 (2002)
- John Riordan. Combinatorial Identities. John Wiley & Sons Inc. New York (1968)