Chapter 2
Kolkata Paise Restaurant Problem

2.1 Introduction

The Kolkata Paise Restaurant (KPR) problem is repeatedly played among a large number N of agents or players having no interaction amongst themselves. The agents or players choose from N' restaurants each evening independently ($N' \leq N$). In the problem the prospective customers or players each have the same set of data regarding the success or failure of the various restaurants: the data set gives the number of prospective customers arriving at each restaurant for the past evenings. Let us assume that the price for the meal to be the same for all the restaurants though the customers can have a ranking of preference for each restaurant (agreed upon by all customers). For simplicity we also assume that each restaurant can serve only one customer any evening. As already mentioned, information about the customer distributions for earlier evenings is available to everyone. Each customer will try to go to the restaurant with the highest possible rank while avoiding the crowd so as to be able to get dinner there. If any restaurant is chosen by more than one customer on any evening, one of them will be randomly chosen (each of them is anonymously treated) and will be served. The rest will not get dinner that evening. The customers collectively learn from their attempts in the past, how to avoid the crowd to get the meal from a high ranking restaurant.

Many years ago, in Kolkata, there were very popular, cheap and fixed rate "Paise Hotel" that were mostly visited by the daily workers or laborers coming to the city for construction works etc. They used to walk (to save the transport costs) to one of these restaurants for their lunches during the *tiffin time* and would miss lunch if they got to a crowed restaurant. Searching for the next restaurant would mean failing to report back to work on time! Paise is the smallest-value Indian coin. There were indeed some well-known rankings of these restaurants, as some of the restaurants would offer tastier food items compared to the others.

B.K. Chakrabarti et al., *Econophysics of the Kolkata Restaurant Problem and Related Games*, New Economic Windows, DOI 10.1007/978-3-319-61352-9_2

A more general example of such a problem can be when the public administration provides hospitals (and beds) in every locality but the locals prefer better ranked hospitals (commonly agreed by everyone) elsewhere. They would then be competing with other 'outsiders' as well as with the local patients of that locality. Unavailability of treatment in the over-crowed hospitals may be considered as lack of the service for those people and consequently as (social) wastage of service by those unattended hospitals.

One (trivial or dictator's) solution to the KPR problem may be the following: planner (or dictator) requests (or orders) everyone to form a que and each one is assigned a restaurant with rank matching the sequence of the person in the que on the first evening. Then each person will be told to go to the next ranked restaurant in the following evening (for the person in the last ranked restaurant will go to the first ranked restaurant). This shifting process (with periodic boundary condition) will continue for successive evenings. We call this dictator's solution. This is one of the most efficient solution (with utilization fraction $\bar{f}$ of the services by the restaurants equal to unity) and the system achieves this efficiency immediately (from the first evening itself). A similar solution exists for the minority game. However, this cannot be any acceptable solution of the KPR problem in reality, where each agent takes his or her own decision (in parallel or democratically) every evening, based on commonly shared information about past events. In KPR problem, the prospective customers try to evolve a learning strategy to get dinners eventually at the best possible ranked restaurant, avoiding the crowd. Generally the evolution of these strategies take considerable time (τ) to converge and even then the eventual utilization fraction $\bar{f}$ is far below unity. The KPR problem have some basic features similar to the Minority Games in that, in both cases, diversity is encouraged (compared to herding behavior), while KPR problem differs from (two-choice) Minority Game in terms of the macroscopic size of the choices available to each player or customer. Note, in the case of dictator's strategy applied to any of the games, convergence time τ vanishes and utilization fraction is unity. In all these games we intend to develop 'democratic' strategies where $\bar{f}$ is high ($\bar{f} \leq 1$) yet τ is small (preferably $\tau \sim \ln N$, when in each learning step one fails to utilize a fraction $(1 - f)$, implying $(1 - f)^{\tau} \sim 1/N$).

Here we are going to discuss the dynamics of a few (classical) stochastic learning strategies for the "Kolkata Paise Restaurant" problem, where N agents choose among N' ($N' = N$ in this chapter) equally priced but differently ranked restaurants every evening such that each agent tries to get dinner in the best restaurant (each serving only one customer and the rest arriving there going without dinner that evening). All agents are taking similar (but not the same) learning strategies and assume that each follow the same probabilistic or stochastic strategy dependent on the information of the past in the game. We will show that a few of these strategies lead to much better utilization of the services than most others.

2.2 Stochastic Learning Strategies

Suppose an agent chooses the kth restaurant having rank r_k on any day (t) with the probability $p_k(t)$ given by

$$p_k(t) = \frac{1}{z}\left[r_k^{\alpha} \exp\left(-\frac{n_k(t-1)}{T}\right)\right], \qquad z = \sum_{k=1}^{N}\left[r_k^{\alpha} \exp\left(-\frac{n_k(t-1)}{T}\right)\right], \tag{2.1}$$

where $n_k(t)$ is the number of agents arriving at the r_kth ranked restaurant on the tth day where $T > 0$ is a scaling (noise) factor and $\alpha \geq 0$ is an exponent. Therefore the probability of selecting a particular restaurant increases with its rank r_k and decreases with its popularity in the previous day (given by the number $n_k(t-1)$). Few properties of the strategies leading to the above probability are the following:

1. For $\alpha = 0$ and $T \to \infty$, $p_k(t) = \frac{1}{N}$ corresponds to the purely random choice case for which the average utilization fraction is around 0.63, i.e., on an average the utilization of the restaurants is 63% (see Sect. 2.2.1).
2. For $\alpha = 0$ and $T \to 0$, the agents still choose randomly but avoid completely those restaurants which had been visited in the last evening or day ($n(t-1)$) is non-zero). Thus choose (again randomly) from the remaining restaurants. Both analytically and in numerical simulations it seen that the average utilization fraction $\bar{f}$ is around 0.46 (see Sect. 2.2.3).

We discuss these limiting and also some intermediate cases, in the next few sections of this chapter.

2.2.1 Random Choice Strategies

Let us consider the case with $r_k = 1$ for all k (restaurants). Suppose there are λN agents and N restaurants. An agent can select any restaurant with equal probability. Therefore, the probability that a single restaurant is chosen by m agents is given by

$$\begin{aligned} \Delta(m) &= \binom{\lambda N}{m} p^m (1-p)^{\lambda N - m}; \; p = \frac{1}{N} \\ &= \frac{\lambda^m}{m!}\exp(-\lambda) \text{ as } N \to \infty. \end{aligned} \tag{2.2}$$

So, the fraction of restaurants not chosen by any agents is given by $\Delta(m=0) = \exp(-\lambda)$ which implies that average fraction of restaurants occupied on any evening is given by

$$\bar{f} = 1 - \exp(-\lambda) \simeq 0.63 \text{ for } \lambda = 1, \tag{2.3}$$

for random choice case in the KPR problem (Chakrabarti et al. [48]). It may be noted this value of the resource utilization factor $\bar{f}$ is obtained at the very first evening. The convergence time τ here is therefore zero convergence time.

2.2.2 Rank Dependent Strategies

Here r_k is not a constant (but dependent of k). For any real α and $T \to \infty$, an agent goes to the kth restaurant with probability $p_k(t) = r_k^\alpha / \sum r_k^\alpha$. The results for such a strategy can then be derived as follows (see Fig. 2.1 for numerical results in different cases):

If an agent selects any restaurant with probability p then probability finding a single restaurant chosen by m agents is given by

$$\Delta(m) = \binom{N}{m} p^m (1-p)^{N-m}. \tag{2.4}$$

Therefore, the probability that any restaurant with rank k is not chosen by any of the agents will be given by

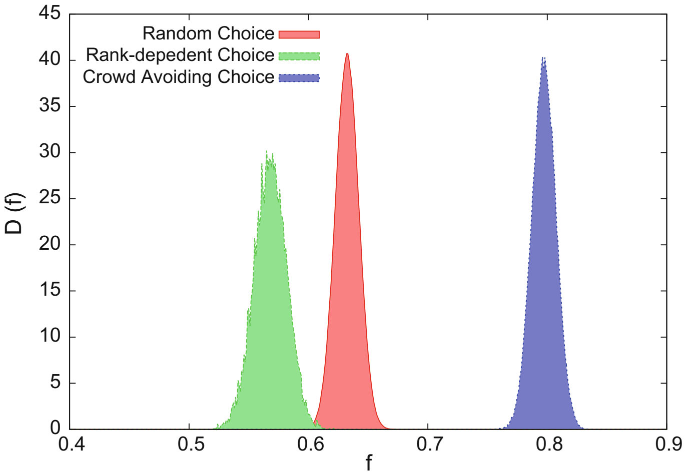

Fig. 2.1 Figure shows the probability distributions of every day utilization f ($f = 1$ denotes 100% utilization) for different strategies. All distributions are Gaussian shape with peaks at $f = 0.63$ (random choice), $f = 0.58$ (simple rank dependent choice) and $f = 0.80$ (crowed avoiding choice)

$$\Delta_k(m=0) = \binom{N}{0}(1-p_k)^N;\ p_k = \frac{r_k^\alpha}{\sum r_k^\alpha}$$
$$\simeq \exp\left(\frac{-k^\alpha N}{\widetilde{N}}\right) \text{ as } N \to \infty, \tag{2.5}$$

where r_k is set equal to k and $\widetilde{N} = \sum_{k=1}^{N} r_k^\alpha = \int_0^N k^\alpha dk = \frac{N^{\alpha+1}}{(\alpha+1)}$.

Hence

$$\Delta_k(m=0) = \exp\left(-\frac{r_k^\alpha\,(\alpha+1)}{N^\alpha}\right). \tag{2.6}$$

Therefore the average fraction of agents getting food any evening (day) in the kth ranked restaurant is given by

$$\bar{f}_k = 1 - \Delta_k\,(m=0)\,. \tag{2.7}$$

Figure 2.2 shows the numerical estimates of $\bar{f}_k$. For $\alpha = 0$, the problem reduces to the random choice case (as considerd in Sect. 2.2.1) and one gets $\bar{f}_k = 1 - e^{-1}$, giving $\bar{f} = \sum \bar{f}_k/N \simeq 0.63$. For $\alpha = 1$, we get $\bar{f}_k = 1 - e^{-2k/N}$, giving $\bar{f} = \sum \bar{f}_k/N \simeq 0.58$ (Chakrabarti et al. [48]).

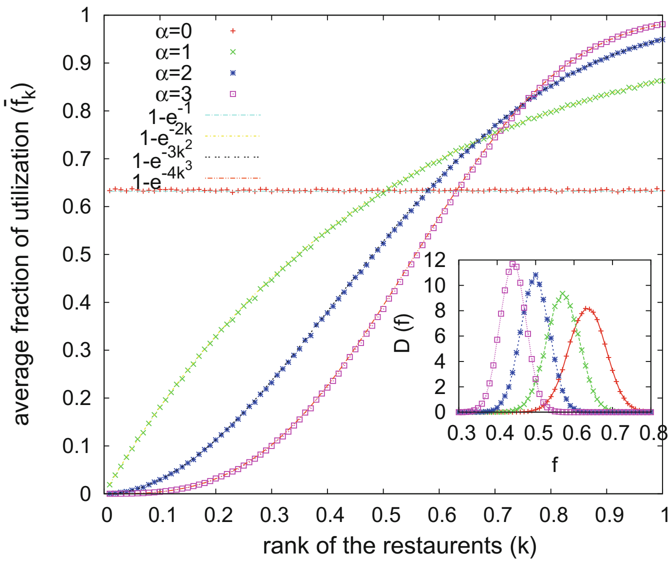

Fig. 2.2 The main figure shows average fraction of utilization ($\bar{f}_k$) versus rank of the restaurants (k) for different α values. The inset shows the distribution $D(f = \sum \bar{f}_k/N)$ of the fraction f agent getting dinner any evening for different α values

2.2.3 Strict Crowd-Avoiding Case

We consider here the case where each agent chooses on any evening (t) randomly among the restaurants in which nobody had visited in the last evening ($t - 1$). This is the case where $\alpha = 0$ and $T \to 0$ in Eq. (2.1). Numerical simulation results for the distribution $D(f)$ of the fraction f of utilized restaurants is Gaussian with a most probable value at $\bar{f} \simeq 0.46$. This can be explained in the following way: As the fraction $\bar{f}$ of restaurants visited by the agents in the last evening is completely avoided by the agents this evening, so the number of available restaurants is $N(1-\bar{f})$ for this evening and is chosen randomly by all the agents. Hence, when fitted to Eq. (2.2) with $\lambda = 1/(1 - \bar{f})$. Therefore, following Eq. (2.2), the equation for $\bar{f}$ can be written as

$$(1 - \bar{f})\left[1 - \exp\left(-\frac{1}{1-\bar{f}}\right)\right] = \bar{f}.$$

By solving above equation, we get $\bar{f} \simeq 0.46$. This result is well fitted with the numerical results for this limit ($\alpha = 0$, $T \to 0$) (Chakrabarti et al. [48]).

2.2.4 Stochastic Crowd Avoiding Case

Let the strategy be the following: if an agent goes to restaurant k in the earlier day ($t - 1$) then the agent will go to the same restaurant in the next day with probability $p_k(t) = \frac{1}{n_k(t-1)}$ and to any other restaurant $k'(\neq k)$ with probability $p_{k'}(t) = \frac{(1-p_k(t))}{(N-1)}$. Numerical results for this stochastic strategy show the average utilization fraction $\bar{f}$ to be around 0.80 and the distribution $D(f)$ to be Gaussian peaked around $f \simeq 0.8$ as shown in Fig. 2.3 (Ghosh et al. [118]).

An approximate estimate of the average utilization ratio $\bar{f}$ for this strategy in steady state may proceed as follows: Let $a_i(t)$ denote the fraction of restaurants having exactly i agents ($i = 0, \ldots, N$) visiting on any evening (t) and assume that $a_i(t) = 0$ for $i \geq 3$ at any (large enough) t, as the dynamics stabilizes in steady state. So, $a_0(t) + a_1(t) + a_2(t) = 1$, $a_1(t) + 2a_2(t) = 1$ for any (large enough) t. Hence $a_0(t) = a_2(t)$. Now $a_2(t)$ fraction of agents will make attempts to leave (each with probability 1/2) their respective restaurants in the next evening ($t + 1$), while no activity will occur on the restaurants where, only one came (a_1) in the previous evening (t). These $a_2(t)$ fraction of agents will get equally divided (each in the remaining $N - 1$ restaurants). Of these $a_2(t)$, the fraction going to the vacant restaurants (a_0 in the earlier evening) is now $a_0(t)a_2(t)$. Hence the new fraction of vacant restaurants at this stage of consideration will be $a_0(t) - a_0(t)a_2(t)$. In the restaurants having exactly two agents (a_2 fraction in the last evening), some vacancy will be created due to this process in steady state, and this fraction will be equal to $\frac{a_2(t)}{4} - a_2(t)\frac{a_2(t)}{4}$. In the steady state, where $a_i(t + 1) = a_i(t) = a_i$ for all i and t, we get $a_0 - a_0a_2 + \frac{a_2}{4} - a_2\frac{a_2}{4} = a_0$. Hence using $a_0 = a_2$ we get $a_0 = a_2 = 0.2$,

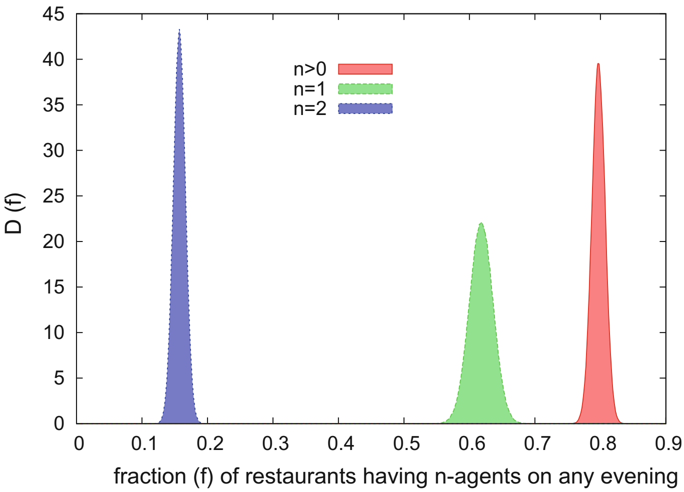

Fig. 2.3 Plot shows the numerical simulation results for a typical prospective customer distribution on any evening

giving $a_1 = 0.6$ and $\bar{f} = a_1 + a_2 = 0.8$ in the steady state. The above calculation is approximate as none of the restaurant is assumed to get more than two costumers on any day ($a_i = 0$ for $i \geq 3$). The advantage in assuming only a_1 and a_2 to be non-vanishing on any evening is that the activity of redistribution on the next evening starts from a_2 fraction of the restaurants only. This of course affects a_0 and a_1 for the next day and for steady state these changes will balance. Numerically we checked that $a_i \leq 0.03$ for $i \geq 3$ and hence the above approximation does not lead any serious error (Ghosh et al. [118]).

2.3 Convergence to a Fair Social Norm with Deterministic Strategies

If the agents or players interact among themselves, then a social norm that can evolve periodically in a organized state with periodicity N where the agents in turn get served in all the N restaurants and each agent gets served every evening. Can we find deterministic strategies (in the absence of a dictator) such that the society collectively and spontaneously achieves this? One variant of Pavlov's "win-shift lose-stay" strategy that can be adopted to achieve the fair social norm. Another variant that can be adopted to achieve the fair social norm in an asymptotic sense. But these strategies are deterministic in nature and also not quite democratic (see Appendix C).

2.3.1 A 'Fair' Strategy

A 'fair' strategy may be as follows:

(i) On the first evening $t = 0$, agents can choose any restaurants either randomly or deterministically.
(ii) If at time t agent i was in a restaurant ranked k and got served then, at time $t + 1$ (next evening), the agent moves to the restaurant ranked $k - 1$ if $k > 1$ and moves to the restaurant ranked N if $k = 1$.
(iii) If agent i was in a restaurant ranked k at time t and had not been got foods then, at time $t + 1$, the agent goes to the same restaurant.

This strategy gives a convergence to the fair social norm in less than or equal to N time steps. And after convergence is achieved, the fair social norm is retained ever after. One difficulty with this strategy is that a myopic agent will find it hard to justify the action of going to the restaurant ranked last after getting served in the top ranked restaurant. However, if the agents are not that myopic and observe the past history of strategies played by all the agents and can figure out that this one evening loss is a tacit commitment devise for this kind of symmetric strategies to work then this voluntary loss is not that implausible. We need to run experiments before arguing for or against this kind of symmetric deterministic strategies. The fair strategy can be modified to take care of this justification problem provided to achieve the fair social norm in an asymptotic sense.

2.3.2 Asymptotically Fair Strategy

This strategy works as follows:

(i) At time (evening or day) $t = 0$, each agent or prospective customer can choose any restaurants either randomly or deterministically.
(ii) If at time t agent i was in a restaurant ranked k and got served then, at time $t + 1$, the agent moves to the restaurant ranked $k - 1$ if $k > 1$ and goes to the same restaurant if $k = 1$.
(iii) If agent i was in a restaurant ranked k at time t and did not get food (more prospective customers and not chosen) then, at time $t + 1$, the agent goes to the restaurant ranked N.

Numerical studies indicate (Appendix C) that after a convergence time (τ) of order N, the customers typically move periodically along the rank order of the restaurants.

2.4 Summary and Discussion

In KPR problem each agent makes decision in each day t independently and is based on the information about the rank k of the restaurants and their previous day prospective customer crowd size given by the numbers $n_k(t - 1) \ldots n_k(0)$. Here we discussed the several stochastic strategies where each agent chooses the kth ranked restaurant with probability $p_k(t)$ described by Eq. (2.1). The utilization fraction f_k of the kth ranked restaurants on every evening is found and their average (over k) distributions $D(f)$ are shown in Fig. 2.2 for some special cases. Numerically we find their distributions to be Gaussian with the most probable utilization fraction $\bar{f} \simeq 0.63$, 0.58 and 0.46 for the cases with $\alpha = 0$, $T \to \infty$; $\alpha = 1$, $T \to \infty$; and $\alpha = 0$, $T \to 0$ respectively. For the stochastic crowd-avoiding strategy, we get the best utilization fraction $\bar{f} \simeq 0.8$. The analytical estimates for $\bar{f}$ for the stochastic crowd-avoiding strategy agree very well with the numerical observations. In all these cases, we assume $N' = N$, that is the number of choices for each of the N agents is the same as the number of agents or players.

We discuss ways to achieve the fair social norm either exactly in the presence of incentive problem or asymptotically in the absence of incentive problem. For $N \to \infty$ limit, implementing or achieving such a norm in a decentralized way is impossible. The KPR problem has similarity with the two restaurants Minority Game Problem as in both the games, herding behavior is punished and diversity's encouraged. Also, both games involve learning of the agents from the past successes. We have observed that, KPR has some simple exact solution limits, a few of which are discussed here. In none of these cases considered here, learning strategies are individualistic; rather each agent chooses following the probability given by Eq. (2.1). In a few learning strategy, the average utilization fraction $\bar{f}$ and their distributions are obtained numerically and compared with the analytic estimates, which are reasonably close. But the real challenge is to design algorithms of learning mixed strategies by the agents so that the fair social norm emerges eventually even when every one independently decides on the basis of their own information.

All the stochastic strategies, being parallel in computational mode, converge to solution at smaller time steps ($\sim \sqrt{N}$ or weakly dependent on N) while for deterministic strategies the convergence time is typically of order of N, which is useless in the truly macroscopic ($N \to \infty$) limits. However, deterministic strategies are useful for small N and rational agents can design appropriate punishment schemes for the deviators.

In brief, the KPR problem has a dictated solution that leads to one of the best possible solution to the problem, with each agent getting his dinner at the best ranked restaurant with a period of N days, and with best possible value of $\bar{f}$ ($= 1$) starting from the first evening itself. However the parallel decision strategies (employing evolving algorithms by the agents, and past informations, e.g., of $n(t)$), which are necessarily parallel among the agents and stochastic (as in democracy), are less efficient ($\bar{f} \ll 1$; the best one the stochastic crowd-avoiding strategy, giving $\bar{f} \simeq 0.8$ only). We note that most of the "smarter" strategies lead to much lower efficiency or less utilization.

http://www.springer.com/978-3-319-61351-2

Econophysics of the Kolkata Restaurant Problem and Related Games
Classical and Quantum Strategies for Multi-agent, Multi-choice Repetitive Games
Chakrabarti, B.K.; Chatterjee, A.; Ghosh, A.; Mukherjee, S.; Tamir, B.
2017, XI, 208 p. 43 illus., 33 illus. in color., Hardcover
ISBN: 978-3-319-61351-2