Rational Abandonments from an $M/G/1$ Queue

Yoav Kerner
IE&M, BGU

Joint work with Eliran Schertzer

Congested Games Wrokshop
IMS, Singapore
December 2015

Outline

- The stochastic environment and the actions
- Motivation and literature
- Continuously observable $M/G/1$ queue

A queue

- Arrivals: Stochastic point process.
- Random service times

Here:

- Arrival process is Poisson.
- Service times are i.i.d.

Actions:

- Join the queue or not
- If joined, when to abandon if still waiting?

Cost and Reward Model

The simplest model

- Homogenous value of service (V)
- Homogenous linear cost implied by waiting (C)
- Common knowledge:
  Stochastic and operational features of the queue, C, V
- Rationality: If you waited t time units stay iff

$$E(W-t|W>t)\,C<V \quad (almost)$$

Cost and Reward Model

The simplest model

- Homogenous value of service (V)
- Homogenous linear cost implied by waiting (C)
- Common knowledge:
  Stochastic and operational features of the queue, C, V
- Rationality: If you waited t time units stay iff

$$E\left(W - t \middle| W > t\right) C < V$$

Main difficulty:
Deriving payoffs and utilities

Cost and Reward Model

The simplest model

- Homogenous value of service (V)
- Homogenous linear cost implied by waiting (C)
- Common knowledge:
  Stochastic and operational features of the queue, C, V
- Rationality: If you waited t time units stay iff

$$E\left(W - t|W > t\right)C < V$$

The simplest queue:

$$M/M/1 \Rightarrow W \text{ is memoryless}$$
$$\Downarrow$$
As long you wait, future is stochastically the same

Unobservable $M/M/1$ with abandonments

Without abandonments W is memoryless

With abandonments even better

Unobservable $M/M/1$ with abandonments

Without abandonments W is memoryless

With abandonments W is IFR

$X \sim F$, ($F(0) = 0$).
X (or F) is Increasing Failure Rate (IFR) if

$$\mathrm{P}(X > t + s | X > t) = \frac{1 - F(t+s)}{1 - F(t)} \downarrow t \quad \forall s$$

Alternatively, the failure (hazard) rate function is increasing:

$$h(t) = \frac{f(t)}{1 - F(t)} \uparrow t$$

Unobservable $M/M/1$ with abandonments

Without abandonments W is memoryless

With abandonments W is IFR

$$\Downarrow$$

If you joined never leave

EXTEND QUEUEING MODEL AND/OR COST MODEL

Hassin & Haviv '96

- Unobservable $M/M/1$
- Linear waiting cost
- Service value drops to 0 after waiting time T

Solution: Nash equilibrium joining probability.
If you joined, stay until until $T \wedge W$.

The reason: Utility rate function is

$$V\mu P(in\ service\ at\ time\ t) - C$$

$P(in\ service\ at\ time\ t)$ is increasing with t

Mandelbaum & Shimkin '00

- $M/M/m$
- Linear waiting cost
- Customers are discharged with (known) probability q

If you are not discharged, your remaining waiting time $\downarrow$

The more you stay, the posterior of being discharged $\uparrow$

$$\text{Possibility for indifference along an interval}$$
$$\Downarrow$$
$$\text{Mixed Nash Equilibrium}$$

Haviv & Ritov '01

- $M/M/m$
- Convex waiting cost

Remaining Waiting time decreasing (stochastically)

Cost of waiting the next time unit increasing.

Possibility for indifference along an interval

$$\Downarrow$$

Mixed Nash Equilibrium

The model: Observable $M/G/1$

- Linear waiting cost C
- Service value V
- FCFS
- DFR service time distribution

Examples:

$$F(t) = 1 - \alpha e^{-\mu_1 t} - (1-\alpha)e^{-\mu_2 t} \quad \mu_1, \mu_2 > 0, \alpha \in (0,1)$$

$$F(t) = 1 - \left(\frac{\beta}{\beta + t}\right)^{\alpha} \quad \alpha, \beta > 0$$

The model: Observable $M/G/1$

- Linear waiting cost C
- Service value V
- FCFS
- DFR service time distribution
- λ - arrival rate
- X - service time
- $F(x) = \mathrm{P}(X \leq x), \bar{F}(x) = 1 - F(x)$
- $h(x) = \frac{dF(x)}{\bar{F}(x)}$ The hazard function

Observable $M/G/1$

- Linear waiting cost C
- Service value V
- FCFS
- DFR service time distribution

After arriving and starting to wait, what can happen next?

1. A service completion
2. An abandonment of someone in front.
3. Nothing

Strategies

Upon arrival, you decide whether to join.

This decision can be based on the queue length.

If you joined, you decide when to leave.

This decision can be based on the dynamics since you arrived: The queue length, departures, abandonments.

Behavior after service completion

After observing a service completion, information from the past is irrelevant (besides the queue length).

Decisions whether to stay or leave are taken continuously.

Define $U_n(t)$:
Utility from staying after observing a service completion, n in front of you in the system and t time units elapsed since the service completion.

$$U_1(t) = Vh(t)dt + (1 - h(t)dt)U_1(t + dt) - Cdt + o(dt)$$

$$\Downarrow$$

$$U_1'(t) = C - h(t)(V - U_1(t))$$

Behavior after service completion

$$U_1(t) = Vh(t)dt + (1 - h(t)dt)U_1(t + dt) - Cdt + o(dt)$$

$$\Downarrow$$

$$U_1'(t) = C - h(t)(V - U_1(t))$$

$$\Downarrow$$

$$U_1(t) = V - C\frac{\int\limits_t^\infty \bar{F}(s)ds}{\bar{F}(t)} = V - CE(X - t|X > t)$$

Behavior after service completion

$$U_n(t) = U_{n-1}(0)h(t)dt + (1 - h(t)dt)U_n(t + dt) - Cdt + o(dt)$$

$$\Downarrow$$

$$U_n'(t) = C - h(t)(U_{n-1}(0) - U_n(t))$$

$$\Downarrow$$

$$U_n(t) = V - C\left(E(X - t|X > t) + (n-1)E(X)\right)$$

Behavior after service completion (cont.)

For every strategy of all, the individual utility obeys

1. $U_n(t) \downarrow t \; \forall n$
2. $U_n(t) \downarrow n \; \forall t$

#1 $\Rightarrow$ pure Nash equilibrium:

$$\text{Abandon at time } T_n \text{ that such that } U_n(T_n) = 0.$$

#2 $\Rightarrow T_n \downarrow n$.

#1+#2$\Rightarrow$ After observing a service completion, no one in front of you will abandon before you (under equilibrium).

There is n_{max} with $U_{n_{max}}(0) < 0 \Rightarrow$ a finite system.

Behavior before observing a change

Define $\hat{U}_n(t)$:
Utility from staying, n in front of you in the system and t time units elapsed since arrival, with no changes in the queue in front of you.

$$\hat{U}_n(t) = h_n(t)dtU_{n-1}(0) + (1 - h_n(t)dt)\hat{U}_n(t + dt) - Cdt$$

$$\Downarrow$$

$$\hat{U}_n'(t) = C - h_n(t)(U_{n-1}(0) - \hat{U}_n(t))$$

$h_n(t)$ - Conditional hazard function

Behavior when abandonment was observed

At service completion, if more than n_{max} are observed, I should leave.
But, if I see more than $n_{max} + 1$ at arrival, should I stay?
Some might abandon........

If the one in front of me is leaving, I should leave as well

n_{max} is the maximal number of customers the system.

Result

There exist a unique symmetric pure equilibrium, defined by

$$(A_1, \ldots, A_{n_{max}-1}, T_1, \ldots, T_{n_{max}-1})$$

If you saw n upon arrival waited A_n without seeing a service completion and/or abandonments, then abandon.

If you saw a service completion and there are n in front of you, abandon after waiting T_n

Remarks

1. Updating the expected remaining waiting time is a complicated queueing problem
2. At any instant, one compares between leaving now and staying until service.
   This looks problematic because one can leave at any time.
   Yet, the DFR assumption ensures monotone behavior.

Queueing analysis

Given strategies $(A_1, \ldots, A_{n_{max}-1}, T_1, \ldots, T_{n_{max}-1})$, a typical state in a Markov process is

$$(k, a, w_{k+1}, w_{k+2}, \ldots, w_n)$$

where

- n is the number of customers in the system, $0 \le n \le n_{max}$.
- k is the number of present customers that saw a service completion, $0 \le k < n$.
- a is time elapsed from the beginning of the current service, $a > 0$.
- w_i is the elapsed waiting time of the i^{th} customer in the in the system, $k < i \le n$.

Steady state analysis

Using supplementary variables, we have the PDE's

$$p_a'(k,a,w_{k+1},...,w_{n-1}) + \sum_{k=0}^{n-1} p'_{w_{k+1}}(k,a,w_{k+1}...,w_{n-1})$$

$$= -p(k,a,w_{k+1},...,w_{n-1})(\lambda + h(a))$$

$$+ \sum_{i=n+1}^{n_{max}} \int_{\underline{w}} p(k,a,w_{k+1},...,w_{n-1},A_n,w_i,...)d\underline{w} + p(k,a,w_{k+1},...,w_{n-1},A_n)$$

PDE's are non-homogeneous and linear.

Finite set of PDE's.

Steady state analysis

Finite set of PDE's.

From the solution we deduce

- Conditional distribution
- Conditional hazard functions

Note that after waiting x time units, the conditional distribution of the elapsed service time is **not** the one that was at arrival shifted by x.

Summary

- The model: $M/G/1$ with DFR service times, linear waiting cost, constant service reward.
- Steady state analysis of the system given any pure strategies (using supplementary variables)
- Finding pure Nash Equilibrium strategies.

In practice the steps are

1. Fix “feasible" A_i, $i < n_{max}$.
2. Analyze the steady state (most likely numerically)
3. Find new A_i, $i < n_{max}$ that are best response.
4. Go to step 2 unless a predefined convergence criterion is satisfied.

Related models

- Non-DFR service times
- Observable only at arrival $M/G/1$ with DFR service
- Bayesian $M/M/1$

Thank you