Extremal $GI/GI/1$ Queues Given First Two Moments

Yan Chen
(joint work with Ward Whitt)
Industrial Engineering & Operations Research Department
Columbia University

Young European Queueing Theorists XII

Monday 3rd December, 2018

Motivation

- Queueing Network Analyzer in Whitt (1983).

Motivation

- Queueing Network Analyzer in Whitt (1983).
- Evaluate Quality of Queueing Approximations.

Motivation

- Queueing Network Analyzer in Whitt (1983).
- Evaluate Quality of Queueing Approximations.
- Propose Tight Upper & Lower Bounds for Queueing Characteristics (*Extremal Queueing Models*).

Motivation

- Queueing Network Analyzer in Whitt (1983).
- Evaluate Quality of Queueing Approximations.
- Propose Tight Upper & Lower Bounds for Queueing Characteristics (*Extremal Queueing Models*).

Research Question: How to Explore Extremal Queues?

Basic Settings

$GI/GI/1$ Model:

- unlimited waiting room and FCFS discipline.

Basic Settings

$GI/GI/1$ Model:

- unlimited waiting room and FCFS discipline.
- mean inter-arrival time and service time $(m_1 = 1, s_1 = \rho)$.

Basic Settings

$GI/GI/1$ Model:

- unlimited waiting room and FCFS discipline.
- mean inter-arrival time and service time $(m_1 = 1, s_1 = \rho)$.
- second moments $m_2 = (1 + c_a^2)m_1^2, s_2 = (1 + c_s^2)s_1^2$.

Basic Settings

$GI/GI/1$ Model:

- unlimited waiting room and FCFS discipline.
- mean inter-arrival time and service time $(m_1 = 1, s_1 = \rho)$.
- second moments $m_2 = (1 + c_a^2)m_1^2, s_2 = (1 + c_s^2)s_1^2$.
- inter-arrival time U distributed as F over $[0, M_a]$.

Basic Settings

$GI/GI/1$ Model:

- unlimited waiting room and FCFS discipline.
- mean inter-arrival time and service time $(m_1 = 1, s_1 = \rho)$.
- second moments $m_2 = (1 + c_a^2)m_1^2, s_2 = (1 + c_s^2)s_1^2$.
- inter-arrival time U distributed as F over $[0, M_a]$.
- service time V distributed as G over $[0, \rho M_s]$.

Basic Settings

$GI/GI/1$ Model:

- unlimited waiting room and FCFS discipline.
- mean inter-arrival time and service time $(m_1 = 1, s_1 = \rho)$.
- second moments $m_2 = (1 + c_a^2)m_1^2, s_2 = (1 + c_s^2)s_1^2$.
- inter-arrival time U distributed as F over $[0, M_a]$.
- service time V distributed as G over $[0, \rho M_s]$.

Related Literature:

Basic Settings

$GI/GI/1$ Model:

- unlimited waiting room and FCFS discipline.
- mean inter-arrival time and service time $(m_1 = 1, s_1 = \rho)$.
- second moments $m_2 = (1 + c_a^2)m_1^2, s_2 = (1 + c_s^2)s_1^2$.
- inter-arrival time U distributed as F over $[0, M_a]$.
- service time V distributed as G over $[0, \rho M_s]$.

Related Literature:

- Kingman (1962), Daley (1977), Ott(1987), Whitt (1984b), Daley (1992), Wolff and Wang (2003).

Basic Settings

$GI/GI/1$ Model:

- unlimited waiting room and FCFS discipline.
- mean inter-arrival time and service time $(m_1 = 1, s_1 = \rho)$.
- second moments $m_2 = (1 + c_a^2)m_1^2, s_2 = (1 + c_s^2)s_1^2$.
- inter-arrival time U distributed as F over $[0, M_a]$.
- service time V distributed as G over $[0, \rho M_s]$.

Related Literature:

- Kingman (1962), Daley (1977), Ott(1987), Whitt (1984b), Daley (1992), Wolff and Wang (2003).
- Bertsmimas and Natarajan (2007), Gupta and Osogami (2011), Osogami and Raymond (2013).

Basic Settings

- $\mathcal{P}_2(M)$: cdf's with support in the interval $[0, mM]$ given mean m and second moment $(c^2+1)m^2$.

Basic Settings

- $\mathcal{P}_2(M)$: cdf's with support in the interval $[0, mM]$ given mean m and second moment $(c^2+1)m^2$.
- $\mathcal{P}_{2,k}(M)$: cdf's with finite k support in $\mathcal{P}_2(M)$.

Basic Settings

- $\mathcal{P}_2(M)$: cdf's with support in the interval $[0, mM]$ given mean m and second moment $(c^2+1)m^2$.
- $\mathcal{P}_{2,k}(M)$: cdf's with finite k support in $\mathcal{P}_2(M)$.
- Steady-state Waiting Time:

$$W \stackrel{\mathrm{d}}{=} [W + V - U]^+,$$

with $W_0 = 0$.

$$w : \mathcal{P}_{a,2}(M_a) \times \mathcal{P}_{s,2}(M_s) \to \mathbb{R},$$

where $0 < \rho < 1$ and

$$w(F, G) \equiv E[W(F, G)].$$

Background

- Pollaczek-Khintchine formula:

$$E[W(M,G)] = \frac{\tau\rho(1+c_s^2)}{2(1-\rho)} = \frac{\rho^2(1+c_s^2)}{2(1-\rho)}.$$

Background

- Pollaczek-Khintchine formula:

$$E[W(M,G)] = \frac{\tau\rho(1+c_s^2)}{2(1-\rho)} = \frac{\rho^2(1+c_s^2)}{2(1-\rho)}.$$

- Kingman (1962) bound:

$$E[W(F,G)] \le \frac{\rho^2([c_a^2/\rho^2]+c_s^2)}{2(1-\rho)}.$$

Background

- Pollaczek-Khintchine formula:

$$E[W(M,G)] = \frac{\tau\rho(1+c_s^2)}{2(1-\rho)} = \frac{\rho^2(1+c_s^2)}{2(1-\rho)}.$$

- Kingman (1962) bound:

$$E[W(F,G)] \leq \frac{\rho^2([c_a^2/\rho^2]+c_s^2)}{2(1-\rho)}.$$

- Daley (1977) bound:

$$E[W(F,G)] \leq \frac{\rho^2([(2-\rho)c_a^2/\rho]+c_s^2)}{2(1-\rho)}.$$

Objectives

We want to answer a Long Standing Open Problem in Queueing Theory.

Objectives

We want to answer a Long Standing Open Problem in Queueing Theory.

- Given any G, what is the **extremal inter-arrival time dist** F^* attaining the UB of $\mathbb{E}[W(F,G)]$?

Objectives

We want to answer a Long Standing Open Problem in Queueing Theory.

- Given any G, what is the **extremal inter-arrival time dist** F^* attaining the UB of $\mathbb{E}[W(F, G)]$?
- Given any F, what is the **extremal service time dist** G^* attaining the UB of $\mathbb{E}[W(F, G)]$?

Objectives

We want to answer a Long Standing Open Problem in Queueing Theory.

- Given any G, what is the **extremal inter-arrival time dist** F^* attaining the UB of $\mathbb{E}[W(F,G)]$?
- Given any F, what is the **extremal service time dist** G^* attaining the UB of $\mathbb{E}[W(F,G)]$?
- What are the extremal F^* and G^* leading to overall UB of $\mathbb{E}[W(F,G)]$?

Extremal Distributions

Two-point distributions (**one parameter family**):

- $F \in \mathcal{P}_{a,2,2}(M_a)$: $c_a^2/(c_a^2 + (b_a - 1)^2)$ at b_a, $(b_a - 1)^2/(c_a^2 + (b_a - 1)^2)$ at $1 - c_a^2/(b_a - 1)$ where $1 + c_a^2 \leq b_a \leq M_a$.

Extremal Distributions

Two-point distributions (**one parameter family**):

- $F \in \mathcal{P}_{a,2,2}(M_a)$: $c_a^2/(c_a^2 + (b_a - 1)^2)$ at b_a, $(b_a - 1)^2/(c_a^2 + (b_a - 1)^2)$ at $1 - c_a^2/(b_a - 1)$ where $1 + c_a^2 \leq b_a \leq M_a$.
- $G \in \mathcal{P}_{s,2,2}(M_s)$: $c_s^2/(c_s^2 + (b_s - 1)^2)$ at ρb_s, $(b_s - 1)^2/(c_s^2 + (b_s - 1)^2)$ at $\rho(1 - c_s^2/(b_s - 1))$ where $1 + c_s^2 \leq b_s \leq M_s$.

Extremal Distributions

Two-point distributions (**one parameter family**):

- $F \in \mathcal{P}_{a,2,2}(M_a)$: $c_a^2/(c_a^2+(b_a-1)^2)$ at b_a, $(b_a-1)^2/(c_a^2+(b_a-1)^2)$ at $1-c_a^2/(b_a-1)$ where $1+c_a^2 \le b_a \le M_a$.
- $G \in \mathcal{P}_{s,2,2}(M_s)$: $c_s^2/(c_s^2+(b_s-1)^2)$ at ρb_s, $(b_s-1)^2/(c_s^2+(b_s-1)^2)$ at $\rho(1-c_s^2/(b_s-1))$ where $1+c_s^2 \le b_s \le M_s$.
- $F = F_0$ for $b_a = 1+c_a^2$ and $F = F_u$ when $b_a = M_a$.
- $G = G_0$ for $b_s = 1+c_s^2$ and $G = G_u$ when $b_s = M_s$

Extremal Distributions

Two-point distributions (**one parameter family**):

- $F \in \mathcal{P}_{a,2,2}(M_a)$: $c_a^2/(c_a^2+(b_a-1)^2)$ at b_a, $(b_a-1)^2/(c_a^2+(b_a-1)^2)$ at $1-c_a^2/(b_a-1)$ where $1+c_a^2 \le b_a \le M_a$.
- $G \in \mathcal{P}_{s,2,2}(M_s)$: $c_s^2/(c_s^2+(b_s-1)^2)$ at ρb_s, $(b_s-1)^2/(c_s^2+(b_s-1)^2)$ at $\rho(1-c_s^2/(b_s-1))$ where $1+c_s^2 \le b_s \le M_s$.
- $F = F_0$ for $b_a = 1+c_a^2$ and $F = F_u$ when $b_a = M_a$.
- $G = G_0$ for $b_s = 1+c_s^2$ and $G = G_u$ when $b_s = M_s$.

Main Results for Extremal Queues

- (a) Given any parameter vector $(1, c_a^2, \rho, c_s^2)$ over bounded support $[0, M_a]$ and $[0, \rho M_s]$, the pair (F_0, G_u) attains the tight UB of the steady-state mean $E[W(F,G)]$, while a pair $(F_0, G_{u,n})$ attains the tight UB of the transient mean $E[W_n(F,G)]$, where $G_{u,n}$ is a two-point distribution with $G_{u,n} \Rightarrow G_u$ as $n \to \infty$.

Main Results for Extremal Queues

- (b) For the unbounded interval of support $[0, \infty)$, the tight UB of $E[W]$ is not attained directly, but is obtained asymptotically in the limit as $M_s \to \infty$ in part (a).
- (c) The mean $E[W(F_0, G_u)]$ does not approach the mean in the associated extremal $F_0/D/1$ queue as $M_s \to \infty$, yet approach to $\lim_{M_s \to \infty} E[(W(F_0, G_u)]$ $(\mathbb{E}[W(F_0, G_{u^*})])$.

Main Results for Upper Bound

Overall Upper Bound Inequalities:

Main Results for Upper Bound

Overall Upper Bound Inequalities:

$$\mathbb{E}[W(F,G)] \leq \mathbb{E}[W(F_0,G_{u^*})] (\text{Tight UB})$$

Main Results for Upper Bound

Overall Upper Bound Inequalities:

$$
\begin{aligned}
\mathbb{E}[W(F,G)] &\leq \mathbb{E}[W(F_0, G_{u^*})] (\text{Tight UB}) \\
&\leq \frac{2(1-\rho)\rho/(1-\delta)c_a^2 + \rho^2 c_s^2}{2(1-\rho)} (\text{UB Approx})
\end{aligned}
$$

Main Results for Upper Bound

Overall Upper Bound Inequalities:

$$
\begin{aligned}
\mathbb{E}[W(F,G)] &\le \mathbb{E}[W(F_0, G_{u^*})] (\text{Tight UB}) \\
&\le \frac{2(1-\rho)\rho/(1-\delta)c_a^2 + \rho^2 c_s^2}{2(1-\rho)} (\text{UB Approx}) \\
&< \frac{\rho^2([(2-\rho)c_a^2/\rho] + c_s^2)}{2(1-\rho)} (\text{Daley(1977)})
\end{aligned}
$$

Main Results for Upper Bound

Overall Upper Bound Inequalities:

$$
\begin{aligned}
\mathbb{E}[W(F,G)] &\le \mathbb{E}[W(F_0, G_{u^*})] (\text{Tight UB}) \\
&\le \frac{2(1-\rho)\rho/(1-\delta)c_a^2 + \rho^2 c_s^2}{2(1-\rho)} (\text{UB Approx}) \\
&< \frac{\rho^2([(2-\rho)c_a^2/\rho] + c_s^2)}{2(1-\rho)} (\text{Daley}(1977)) \\
&< \frac{\rho^2([c_a^2/\rho^2] + c_s^2)}{2(1-\rho)} (\text{Kingman}(1962))
\end{aligned}
$$

Main Results for Upper Bound

Overall Upper Bound Inequalities:

$$\begin{aligned}
\mathbb{E}[W(F,G)] &\leq \mathbb{E}[W(F_0,G_{u^*})](\text{Tight UB}) \\
&\leq \frac{2(1-\rho)\rho/(1-\delta)c_a^2+\rho^2c_s^2}{2(1-\rho)}(\text{UB Approx}) \\
&< \frac{\rho^2([(2-\rho)c_a^2/\rho]+c_s^2)}{2(1-\rho)}(\text{Daley(1977)}) \\
&< \frac{\rho^2([c_a^2/\rho^2]+c_s^2)}{2(1-\rho)}(\text{Kingman(1962)})
\end{aligned}$$

where $\delta \in (0,1)$ and $\delta = \exp(-(1-\delta)/\rho)$.

Main Results for Upper Bound

Table: A comparison of the bounds and approximations for the steady-state mean $\mathbb{E}[W]$ as a function of ρ for the case $c_a^2 = c_s^2 = 4.0$.

ρ	Tight LB	HTA	**Tight UB**	**UB Approx**	δ	MRE	Daley	Kingman
0.10	0.00	0.044	0.422	0.422	0.000	0.003%	0.44	2.24
0.20	0.00	0.200	0.904	0.906	0.007	0.19%	1.00	2.60
0.30	0.00	0.514	1.499	1.51	0.041	0.60%	1.71	3.11
0.40	0.00	1.07	2.304	2.33	0.107	0.94%	2.67	3.87
0.50	0.25	2.00	3.470	3.51	0.203	1.15%	4.00	5.00
0.60	1.00	3.60	5.295	5.35	0.324	1.07%	6.00	6.80
0.70	2.42	6.53	8.441	8.52	0.467	0.93%	9.33	9.93
0.80	5.50	12.80	14.92	15.02	0.629	0.67%	16.00	16.40
0.90	15.25	32.40	34.72	34.84	0.807	0.35%	36.00	36.20
0.95	35.13	72.20	74.62	74.76	0.902	0.18%	76.00	76.10
0.98	95.05	192.1	194.6	194.7	0.960	0.07%	196.0	196.0
0.99	195.0	392.0	394.5	394.7	0.980	0.04%	396.0	396.0

Main Reduction Theorem

Theorem

(reduction to a three-point distribution) Consider the class of $GI/GI/1$ queues with cdf $F \in \mathcal{P}_{a,2}$ with cdf $G \in \mathcal{P}_{s,2}$ where $0 < \rho < 1$, the following suprema are attained as indicated:

(a) For any specified $G \in \mathcal{P}_{s,2}$, there exists $F^*(G) \in \mathcal{P}_{a,2,3}(M_a)$ such that

$$
\begin{aligned}
w_a^{\uparrow}(G) &= \sup\{w(F,G) : F \in \mathcal{P}_{a,2}(M_a)\} \\
&= \sup\{w(F,G) : F \in \mathcal{P}_{a,2,3}(M_a)\} = w(F^*(G),G).
\end{aligned}
$$

Main Reduction Theorem

Theorem

(reduction to a three-point distribution) Consider the class of $GI/GI/1$ *queues with cdf* $F \in \mathcal{P}_{a,2}$ *with cdf* $G \in \mathcal{P}_{s,2}$ *where* $0 < \rho < 1$, *the following suprema are attained as indicated:*

(a) For any specified $G \in \mathcal{P}_{s,2}$, *there exists* $F^*(G) \in \mathcal{P}_{a,2,3}(M_a)$ *such that*

$$\begin{aligned} w_a^{\uparrow}(G) &\equiv \sup\left\{w(F,G) : F \in \mathcal{P}_{a,2}(M_a)\right\} \\ &= \sup\left\{w(F,G) : F \in \mathcal{P}_{a,2,3}(M_a)\right\} = w(F^*(G), G). \end{aligned}$$

Main Reduction Theorem

Theorem

(*continued*)

(*b*) *For any specified* $F \in \mathcal{P}_{a,2}$*, there exists* $G^*(F) \in \mathcal{P}_{s,2,3}(M_s)$ *such that*

$$\begin{aligned} w_s^{\uparrow}(F) &\equiv \sup\{w(F,G) : G \in \mathcal{P}_{s,2}(M_s)\} \\ &= \sup\{w(F,G) : G \in \mathcal{P}_{s,2,3}(M_s)\} = w(F, G^*(F)). \end{aligned}$$

(*c*) *There exists* (F^{**}, G^{**}) *in* $\mathcal{P}_{a,2,3}(M_a) \times \mathcal{P}_{s,2,3}(M_s)$ *such that*

$$\begin{aligned} w^{\uparrow} &\equiv \sup\{w(F,G) : F \in \mathcal{P}_{a,2}, G \in \mathcal{P}_{s,2}(M_s)\} \\ &= \sup\{w(F,G) : F \in \mathcal{P}_{a,2,3}(M_a), G \in \mathcal{P}_{s,2,3}(M_s)\} \\ &= w(F^{**}, G^{**}) = w_a^{\uparrow}(G^{**}) = w_s^{\uparrow}(F^{**}). \end{aligned}$$

Ideas of Proof for Case (a)

- Start with $W \stackrel{\mathrm{d}}{=} [W + V_G - U_F]^+$.

Ideas of Proof for Case (a)

- Start with $W \stackrel{\mathrm{d}}{=} [W + V_G - U_F]^+$.
- $E[W] = E[(W + V - U)^+] = \int_0^{M_a} \phi(u) dF$ for ϕ expressed as the double integral

$$\phi(u) \equiv \int_0^\infty \int_0^\infty (x + v - u)^+ \, dG(v) \, dH(x), \quad 0 \le u \le M_a.$$

Ideas of Proof for Case (a)

- Start with $W \stackrel{\mathrm{d}}{=} [W + V_G - U_F]^+$.
- $E[W] = E[(W + V - U)^+] = \int_0^{M_a} \phi(u) dF$ for ϕ expressed as the double integral

$$\phi(u) \equiv \int_0^{\infty} \int_0^{\infty} (x + v - u)^+ \, dG(v) \, dH(x), \quad 0 \le u \le M_a.$$

- Optimize over F: $\sup_{F \in \mathcal{P}_{a,2}(M_a)} \int_0^{M_a} \phi(u) dF$. It can be written as

$$\sup \{E[(W + V - U)^+] : F_U \in \mathcal{P}_{a,2}(M_a)\}.$$

Ideas of Proof for Case (a)

- Start with $W \stackrel{\mathrm{d}}{=} [W + V_G - U_F]^+$.
- $E[W] = E[(W + V - U)^+] = \int_0^{M_a} \phi(u) dF$ for ϕ expressed as the double integral

$$\phi(u) \equiv \int_0^\infty \int_0^\infty (x + v - u)^+ \, dG(v) \, dH(x), \quad 0 \le u \le M_a.$$

- Optimize over F: $\sup_{F \in \mathcal{P}_{a,2}(M_a)} \int_0^{M_a} \phi(u) dF$. It can be written as

$$\sup \{ E[(W + V - U)^+] : F_U \in \mathcal{P}_{a,2}(M_a) \}.$$

- Update $W_1 = (W + V - U_{F_1})^+$. $\mathbb{E}[W_1] \ge \mathbb{E}[W]$
- Repeat to obtain $W_2 = (W_1 + V - U_{F_2})^+$. $\mathbb{E}[W_2] \ge \mathbb{E}[W_1]$?

Ideas of Proof for Case (a)

- Start with $W \stackrel{\mathrm{d}}{=} [W + V_G - U_F]^+$.
- $E[W] = E[(W+V-U)^+] = \int_0^{M_a} \phi(u) dF$ for ϕ expressed as the double integral

$$\phi(u) \equiv \int_0^{\infty} \int_0^{\infty} (x+v-u)^+ \, dG(v) \, dH(x), \quad 0 \le u \le M_a.$$

- Optimize over F: $\sup_{F \in \mathcal{P}_{a,2}(M_a)} \int_0^{M_a} \phi(u) dF$. It can be written as

$$\sup \{ E[(W+V-U)^+] : F_U \in \mathcal{P}_{a,2}(M_a) \}.$$

- Update $W_1 = (W + V - U_{F_1})^+$. $\mathbb{E}[W_1] \ge \mathbb{E}[W]$
- Repeat to obtain $W_2 = (W_1 + V - U_{F_2})^+$. $\mathbb{E}[W_2] \ge \mathbb{E}[W_1]$?

Ideas of Proof for Case (a)

- Suppose fixed point exists:

$$(W(F^*, G) + V - U_{F^*})^+ \stackrel{\mathrm{d}}{=} W(F^*, G).$$

With the property

$$\begin{aligned} &E[(W(F^*, G) + V - U_{F^*})^+] \\ &= \sup\{E[(W(F, G) + V - U_F)^+] : F \in \mathcal{P}_{a,2}(M_a)\}. \end{aligned}$$

Ideas of Proof for Case (a)

- Suppose fixed point exists:

$$(W(F^*, G) + V - U_{F^*})^+ \overset{d}{=} W(F^*, G).$$

With the property

$$E[(W(F^*, G) + V - U_{F^*})^+]$$
$$= \sup\{E[(W(F, G) + V - U_F)^+] : F \in \mathcal{P}_{a,2}(M_a)\}.$$

- Show $F^* \in \mathcal{P}_{a,2,3}(M_a)$.

Relate to Transient Mean Waiting Time

Theorem

(reduction to the transient mean) Consider $GI/GI/1$ queues given first two moments and bounded support,

(a) For any specified $G \in \mathcal{P}_{s,2}$, if there exists $F_n \in \mathcal{P}_{a,2,3}(M_a)$ such that

$$w_n(F_n, G) = w_{a,n}^{\uparrow}(G) \equiv \sup\{w_n(F, G) : F \in \mathcal{P}_{a,2}(M_a)\},$$

then the sequence $\{F_n : n \geq 1\}$ is tight, so that there exists a convergent subsequence. Moreover, if F is the limit of any convergent subsequence, then F is in $\mathcal{P}_{a,2,3}(M_a)$ and F is optimal for $E[W(F, G)]$.

Relate to Transient Mean Waiting Time

Theorem

(continued) (b) *For any specified* $F \in \mathcal{P}_{a,2}$*, if there exists* $G_n \in \mathcal{P}_{s,2}(M_a)$ *such that*

$$w_n(F, G_n) = w^{\uparrow}_{s,n}(F) \equiv \sup\{w_n(F, G) : G \in \mathcal{P}_{s,2}(M_a)\},$$

then the sequence $\{G_n : n \geq 1\}$ *is tight, so that there exists a convergent subsequence. Moreover, if* G *is the limit of any convergent subsequence, then* G *is in* $\mathcal{P}_{s,2,3}(M_s)$ *and* G *is optimal for* $E[W(F, G)]$.

(c) *If there exists* (F_n, G_n) *in* $\mathcal{P}_{a,2,3}(M_a) \times \mathcal{P}_{s,2,3}(M_s)$ *such that*

$$w_n(F_n, G_n) = w^{\uparrow}_{n} \equiv \sup\{w_n(F, G) : F \in \mathcal{P}_{a,2}(M_a), G \in \mathcal{P}_{s,2}(M_s)\},$$

then the sequence $\{(F_n, G_n) : n \geq 1\}$ *is tight.*

The Multinomial Representation

In the bounded support three-point distribution space,

$$P_k(\mathbf{v}, \mathbf{p}) \equiv \frac{k! p_1^{k_1} p_2^{k_2} p_3^{k_3}}{k_1! k_2! k_3!},$$

$$Q_w(\mathbf{u}, \mathbf{q}) \equiv \frac{w! q_1^{w_1} q_2^{w_2} q_3^{w_3}}{w_1! w_2! w_3!}.$$

The Multinomial Representation

In the bounded support three-point distribution space,

$$P_k(\mathbf{v},\mathbf{p}) \equiv \frac{k!p_1^{k_1}p_2^{k_2}p_3^{k_3}}{k_1!k_2!k_3!},$$

$$Q_w(\mathbf{u},\mathbf{q}) \equiv \frac{w!q_1^{w_1}q_2^{w_2}q_3^{w_3}}{w_1!w_2!w_3!}.$$

$$E[W_n] = \sum_{k=1}^{n}\frac{1}{k}\sum_{(\mathbf{k},\mathbf{w})\in\mathcal{I}} \max\{0,\sum_{i=1}^{3}(k_i v_i - w_j u_j)\}P_k(\mathbf{v},\mathbf{p})Q_w(\mathbf{u},\mathbf{q}).$$

The Multinomial Representation

A **tractable** formulation for optimizing $\mathbb{E}[W_n]$:

$$\begin{aligned}
\text{maximize} \quad & \sum_{k=1}^{n} \frac{1}{k} \sum_{\sum k_i = k, \sum w_j = k} \max\Big(\sum_i k_i v_i - \sum_j w_j u_i, 0\Big) P(k_1, k_2, k_3) Q(w_1, w_2, w_3) \\
\text{subject to} \quad & \sum_{j=1}^{3} v_j p_j = s_1, \quad \sum_{j=1}^{3} v_j^2 p_j = (1 + c_s^2) s_2^2, \\
& \sum_{j=1}^{3} u_j q_j = m_1, \quad \sum_{j=1}^{3} u_j^2 q_j = (1 + c_a^2) m_1^2, \\
& \sum_{j=1}^{3} p_j = \sum_{k=1}^{3} q_k = 1, \\
& M_s \ge v_j \ge 0,\ M_a \ge u_j \ge 0,\ p_j \ge 0,\ q_j \ge 0, \quad 1 \le j \le 3.
\end{aligned} \tag{1}$$

The Multinomial Representation

- all local optima in $\mathcal{P}_{a,2,2} \times \mathcal{P}_{s,2,2}$.
- $E[W(F_0, G_{u,n})]$ is larger than for other local optima.
- $G_{u,n}$ denote there is a optimal $b_s^*(n, M_s)$ (not necessary to be equal to M_s).

The Multinomial Representation

- all local optima in $\mathcal{P}_{a,2,2} \times \mathcal{P}_{s,2,2}$.
- $E[W(F_0, G_{u,n})]$ is larger than for other local optima.
- $G_{u,n}$ denote there is a optimal $b_s^*(n, M_s)$ (not necessary to be equal to M_s).

Table: Numerical values of $E[W_n(F_0, G_{u,n})]$ from the optimization and numerical search for $c_a^2 = c_s^2 = 4.0$ for $M_a = M_s = 10$

n	$\rho = 0.1$	$\rho = 0.2$	$\rho = 0.3$	$\rho = 0.4$	$\rho = 0.5$	$\rho = 0.6$	$\rho = 0.7$	$\rho = 0.8$	$\rho = 0.9$
1	0.080	0.160	0.240	0.320	0.400	0.489	0.579	0.668	0.758
5	0.269	0.538	0.813	1.095	1.414	1.777	2.140	2.505	2.882
10	0.357	0.716	1.102	1.525	2.056	2.634	3.228	3.869	4.555
15	0.386	0.778	1.220	1.744	2.410	3.137	3.949	4.832	5.776
20	0.395	0.804	1.281	1.871	2.626	3.508	4.499	5.602	6.808
25	0.399	0.814	1.313	1.948	2.781	3.782	4.933	6.242	7.693
30	0.400	0.820	1.332	1.999	2.896	3.992	5.291	6.794	8.508
35	0.400	0.822	1.343	2.032	2.979	4.163	5.590	7.270	9.185
40	0.400	0.824	1.349	2.056	3.040	4.299	5.846	7.696	9.858
45	0.400	0.824	1.354	2.072	3.088	4.411	6.067	8.075	10.423
50	0.400	0.825	1.356	2.084	3.126	4.505	6.260	8.421	11.002

Numerics and Simulation Search over $\mathcal{P}_{a,2,2} \times \mathcal{P}_{s,2,2}$

Table: Numerical estimates of $E[W_{20}]$ as a function of b_a and b_s when $\rho = 0.5$, $c_a^2 = c_s^2 = 4.0$ and $M_a = 7 < M_s = 10$.

$b_s \backslash b_a$	5.00	5.25	5.50	5.75	6.00	6.25	6.50	6.75	7.00
5.0	2.497	2.530	2.518	2.497	2.469	2.439	2.406	2.371	2.335
5.5	2.557	2.414	2.420	2.422	2.402	2.378	2.351	2.320	2.288
6.0	2.561	2.447	2.328	2.318	2.328	2.312	2.290	2.266	2.239
7.0	2.549	2.447	2.331	2.204	2.165	2.149	2.154	2.150	2.132
8.0	2.556	2.430	2.319	2.208	2.074	2.029	2.021	2.010	2.007
9.0	2.598	2.456	2.310	2.183	2.068	1.937	1.895	1.903	1.898
10.0	**2.626**	2.506	2.353	2.188	2.043	1.921	1.786	1.779	1.789

Numerics and Simulation Search over $\mathcal{P}_{a,2,2} \times \mathcal{P}_{s,2,2}$

Table: Simulation estimates of $E[W]$ as a function of b_a and b_s when $\rho = 0.5$, $c_a^2 = c_s^2 = 4.0$ and $M_a = 7 < M_s = 10$.

$b_s \backslash b_a$	5.00	5.25	5.50	5.75	6.00	6.25	6.50	6.75	7.0
5.0	3.110	3.134	3.117	3.083	3.040	2.997	2.950	2.910	2.863
5.5	3.179	3.026	3.019	3.009	2.975	2.938	2.901	2.860	2.823
6.0	3.191	3.065	2.932	2.907	2.905	2.876	2.844	2.809	2.767
7.0	3.181	3.067	2.942	2.797	2.748	2.720	2.713	2.691	2.670
8.0	3.195	3.056	2.934	2.810	2.664	2.611	2.591	2.564	2.553
9.0	3.239	3.092	2.931	2.792	2.663	2.525	2.472	2.467	2.449
10.0	**3.282**	3.142	2.986	2.812	2.640	2.507	2.367	2.350	2.349

Classical $GI/GI/1$ Problem

We want to address a Long Standing Open Problem in Queueing Theory.

Classical $GI/GI/1$ Problem

We want to address a Long Standing Open Problem in Queueing Theory.

- Given any G, what is the **extremal inter-arrival time dist** F^* attaining the UB of $\mathbb{E}[W(F,G)]$?

Classical $GI/GI/1$ Problem

We want to address a Long Standing Open Problem in Queueing Theory.

- Given any G, what is the **extremal inter-arrival time dist** F^* attaining the UB of $\mathbb{E}[W(F,G)]$?
  - Answer: **three point dist** F^*.

Classical $GI/GI/1$ Problem

We want to address a Long Standing Open Problem in Queueing Theory.

- Given any G, what is the **extremal inter-arrival time dist** F^* attaining the UB of $\mathbb{E}[W(F,G)]$?
  - Answer: **three point dist** F^*.
- Given any F, what is the **extremal service time dist** G^* attaining the UB of $\mathbb{E}[W(F,G)]$?

Classical $GI/GI/1$ Problem

We want to address a Long Standing Open Problem in Queueing Theory.

- Given any G, what is the **extremal inter-arrival time dist** F^* attaining the UB of $\mathbb{E}[W(F,G)]$?
  - Answer: **three point dist** F^*.
- Given any F, what is the **extremal service time dist** G^* attaining the UB of $\mathbb{E}[W(F,G)]$?
  - Answer: **three point dist** G^*.

Classical $GI/GI/1$ Problem

We want to address a Long Standing Open Problem in Queueing Theory.

- Given any G, what is the **extremal inter-arrival time dist** F^* attaining the UB of $\mathbb{E}[W(F,G)]$?
  - Answer: **three point dist** F^*.
- Given any F, what is the **extremal service time dist** G^* attaining the UB of $\mathbb{E}[W(F,G)]$?
  - Answer: **three point dist** G^*.
- What are the **extremal** F^* and G^* leading to overall UB of $\mathbb{E}[W(F,G)]$?
  - Answer: **the pair** F_0, G_u.

Classical $GI/GI/1$ Problem

We want to address a Long Standing Open Problem in Queueing Theory.

- Given any G, what is the **extremal inter-arrival time dist** F^* attaining the UB of $\mathbb{E}[W(F,G)]$?
  - Answer: **three point dist** F^*.
- Given any F, what is the **extremal service time dist** G^* attaining the UB of $\mathbb{E}[W(F,G)]$?
  - Answer: **three point dist** G^*.
- What are the **extremal** F^* and G^* leading to overall UB of $\mathbb{E}[W(F,G)]$?
  - Answer: **the pair** F_0, G_u.

Impact of Inter-arrival Time

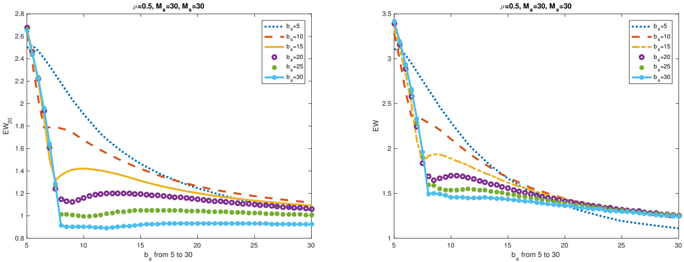


Figure: Simulation estimates of the transient mean $E[W_{20}]$ (left) and the steady-state mean $E[W]$ (right) as a function of b_a for six cases of b_s the in the case $\rho = 0.5$, $c_a^2 = c_s^2 = 4.0$ and $M_a = M_s = 30$.

Impact of Service Time

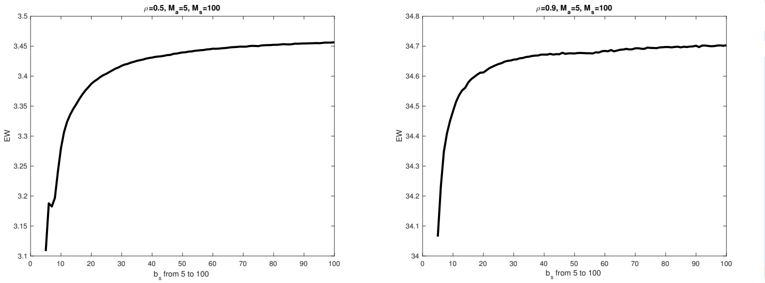
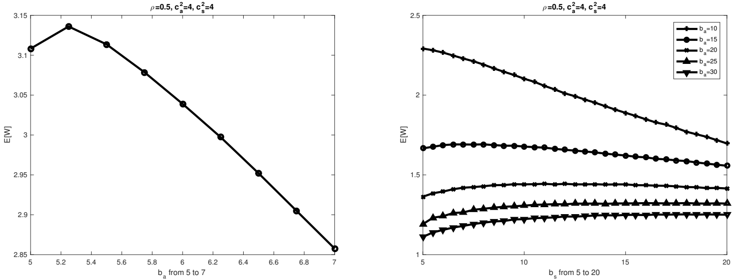


Figure: $E[W(F_0, G)]$ for $G \in \mathcal{P}_{s,2,2}$ as a function of b_s given $b_a = (1 + c_a^2)$.

Counterexamples

Conjecture

Given any $G \in \mathcal{P}_{s,2}$, the extremal inter-arrival time is F_0.
Given any $F \in \mathcal{P}_{a,2}$, the extremal service time is G_0, G_u.



Counterexamples

Conjecture

Given any $G \in \mathcal{P}_{s,2}$, *the extremal inter-arrival time is* F_0.
Given any $F \in \mathcal{P}_{a,2}$, *the extremal service time is* G_0, G_u.

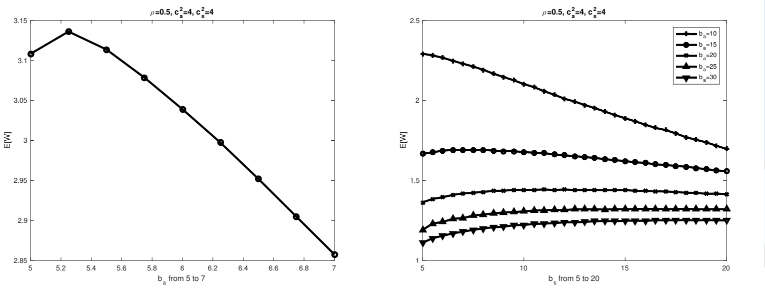


Figure: The $E[W]$: a function of b_a in $[(1 + c_a^2), M_a = 7]$ for $b_s = 5$, i.e., for G_0 (left) and as a function of b_s in $[(1 + c_s^2), M_s = 20]$ for b_a (right).

Conjectures

Theorem

(Counterexamples) Fix any service time dist G, $F^*(G) = F_0$*; Fix any inter-arrival dist* F, $G^*(F)$ *is* G_0 *or* G_u*. The both arguments are invalid.*

Conjecture

(Chen and Whitt I) Fix any G*, the extremal* $F^*(G)$ *is a two-point distribution.*

Conjecture

(Chen and Whitt II) Fix any two-point F*, the extremal* $G^*(F)$ *is a two-point distribution.*

Conjectures

Theorem

(Counterexamples) Fix any service time dist G, $F^*(G) = F_0$; Fix any inter-arrival dist F, $G^*(F)$ is G_0 or G_u. The both arguments are invalid.

Conjecture

(Chen and Whitt I) Fix any G, the extremal $F^*(G)$ is a two-point distribution.

Conjecture

(Chen and Whitt II) Fix any two-point F, the extremal $G^*(F)$ is a two-point distribution.

Conjectures

Theorem

(*Counterexamples*) *Fix any service time dist* G, $F^*(G) = F_0$; *Fix any inter-arrival dist* F, $G^*(F)$ *is* G_0 *or* G_u. *The both arguments are invalid.*

Conjecture

(*Chen and Whitt I*) *Fix any* G, *the extremal* $F^*(G)$ *is a two-point distribution.*

Conjecture

(*Chen and Whitt II*) *Fix any two-point* F, *the extremal* $G^*(F)$ *is a two-point distribution.*

Upper Bound Inequality

Overall Upper Bound:

Upper Bound Inequality

Overall Upper Bound:

$$\begin{aligned}\mathbb{E}[W(F,G)] &\leq \mathbb{E}[W(F_0, G_{u^*})] \\ &\leq \frac{2(1-\rho)\rho/(1-\delta)c_a^2 + \rho^2 c_s^2}{2(1-\rho)}.\end{aligned}$$

$\delta \in (0,1)$ and $\delta = \exp(-(1-\delta)/\rho)$.

Theorem

(an UB for $E[W(F_0, G_{u^*})]$) For the $GI/GI/1$ queue with parameter four-tuple $(1, c_a^2, \rho, c_s^2)$, if $E[W(F_0, G_{u^*})]$ is UB, then

$$E[W(F_0, G_{u^*})] \leq \frac{2(1-\rho)\rho/(1-\delta)c_a^2 + \rho^2 c_s^2}{2(1-\rho)}.$$

Upper Bound Inequality

Overall Upper Bound:

$$\begin{aligned}\mathbb{E}[W(F,G)] &\le \mathbb{E}[W(F_0,G_{u^*})] \\ &\le \frac{2(1-\rho)\rho/(1-\delta)c_a^2+\rho^2c_s^2}{2(1-\rho)}.\end{aligned}$$

$\delta \in (0,1)$ and $\delta = \exp(-(1-\delta)/\rho)$.

Theorem

(*an UB for* $E[W(F_0,G_{u^*})]$) *For the* $GI/GI/1$ *queue with parameter four-tuple* $(1,c_a^2,\rho,c_s^2)$, *if* $E[W(F_0,G_{u^*})]$ *is UB, then*

$$E[W(F_0,G_{u^*})] \le \frac{2(1-\rho)\rho/(1-\delta)c_a^2+\rho^2c_s^2}{2(1-\rho)}.$$

Thank You!