1 Today's Goal

Our goal is to derive the **transform of response time**, $\widetilde{T}(s)$, for the M/G/1 with arrival rate λ and job size S. This will allow us to understand the variance of response time, and higher moments.

2 Warmup on Transforms

Warmup Question: Let A_S denote the number of arrivals of a Poisson Process with rate λ during r.v. time S. What is $\widehat{A_S}(z)$?

Warmup Question: What is $\widetilde{T}(s)$ for an M/M/1 with arrival rate λ and service rate μ?

3 M/G/1 Embedded DTMC

Question: Why can't we analyze the M/G/1 just like we did the M/M/1?

Definition: The **embedded DTMC** for the M/G/1 considers the M/G/1 only at points in time when a departure occurs.

State i indicates that the last departure left behind i jobs.

Question: Why is the embedded chain a DTMC?

Definition:

$$\begin{aligned}\pi_i^{embed} &= \text{probability of being in state } i \text{ of the embedded DTMC.}\\ &= \text{fraction of departures that leave behind } i \text{ jobs.}\end{aligned}$$

Question: How does π_i^{embed} compare to $\pi_i^{M/G/1}$?

Definition: Let $a_j = \mathbf{P}\{j \text{ arrivals during } S\}$, where S is a service time.

Question: Write the stationary equation for state j in the embedded DTMC.

Question: How can we solve these stationary equations to get the limiting probabilities?

4 Outline for deriving M/G/1 response time

1. STEP 1: Do not solve for limiting probabilities. Instead, derive $\widehat{N}(z)$.

2. STEP 2: Find a way to convert $\widehat{N}(z)$ into $\widetilde{T}(s)$. This conversion is called "Distributional Little's Law."

5 Prove $\widehat{N}(z) = \frac{(z-1)\pi_0\widehat{A_S}(z)}{z-\widehat{A_S}(z)}$

Now show how to get from

$$\widehat{N}(z) = \frac{(z-1)\pi_0 \widehat{A_S}(z)}{z - \widehat{A_S}(z)}$$

to

$$\widehat{N}(z) = \frac{\widetilde{S}(\lambda - \lambda z)(1-\rho)(1-z)}{\widetilde{S}(\lambda - \lambda z) - z}$$

6 Commercial Break: Announcements

1. Mor's office hours today: GHC 7207: 5:30 p.m. - 7 p.m.

2. Zhouzi teaching Friday's class: Covering Chpt 27 on busy periods. This is your LAST Friday class.

7 From $\widehat{\mathbf{N}}(\mathbf{z})$ to $\widetilde{\mathbf{T}}(\mathbf{s})$: Distributional Little's Law

Question: What is $\widehat{A_S}(z)$?

Question: Does the above expression hold for all r.v.s, S?

Question: Let $S = T$, where T denotes the response time in an M/G/1. Write the expression for $\widehat{A_T}(z)$.

Question: What is a much nicer name for the random variable A_T?

8 Deriving $\widetilde{\mathbf{T}}(\mathbf{s})$

Using:

$$\widehat{N}(z) = \frac{\widetilde{S}(\lambda - \lambda z)(1-\rho)(1-z)}{\widetilde{S}(\lambda - \lambda z) - z},$$

show that

$$\widetilde{T}(s) = \frac{\widetilde{S}(s)(1-\rho)s}{\lambda \widetilde{S}(s) - \lambda + s}$$

9 Deriving $\widetilde{\mathbf{T_Q}}(\mathbf{s})$

Question: What is $\widetilde{T_Q}(s)$?

Question: Shouldn't there be a term for excess?

10 Variance of $\mathbf{T_Q}$

We can differentiate the transform to get the moments of T_Q. Here are the answers:

$$\mathbf{E}[T_Q] = \frac{\rho}{1-\rho} \cdot \frac{\mathbf{E}\left[S^2\right]}{2\mathbf{E}[S]}$$

$$\mathbf{Var}(T_Q) = (\mathbf{E}[T_Q])^2 + \frac{\lambda \mathbf{E}\left[S^3\right]}{3(1-\rho)}$$

11 The M/G/1/setup queue (if time)

It is often the case that the first job to start a busy period experiences an initial setup time, I, before its service is started. During the setup time, there is no work being done, and more jobs may queue up.

- Example: Setup time for large servers in data centers: ___________
- Example: Setup time for photocopier: ___________

GOAL: How does setup time affect mean response time?
Easiest approach is to compute the transform of response time.

Approach: Start with the embedded DTMC.

$$\pi_i^{embed} = \text{probability that the last departure left behind } i \text{ jobs.}$$

By PASTA,

$$\pi_i^{embed} = \text{ time-avg fraction of time there are } i \text{ jobs} = \pi_i.$$

Notation:

$$\begin{aligned} S &= \text{job size} \\ I &= \text{initial setup time} \\ a_j &= \mathbf{P}\{j \text{ arrivals in } S \text{ seconds }\} \\ a'_j &= \mathbf{P}\{j \text{ arrivals in } S + I \text{ seconds }\} \end{aligned}$$

Question: Write the stationary equation for $\pi_j^{embed} = \pi_j$:

12 Deriving $\widehat{\mathbf{N}}^{\mathbf{setup}}(\mathbf{z})$

Starting with your stationary equation, show that

$$\widehat{N}^{setup}(z) \;=\; \pi_0 \frac{z\widehat{A}_S(z)\widehat{A}_I(z) - \widehat{A}_S(z)}{z - \widehat{A}_S(z)}.$$

13 Derive π_0 for the M/G/1/setup

Clearly π_0 is not the same as it was for the M/G/1 queue. Use $\widehat{N}^{setup}(z)$ to derive it.

Recall:

$$\widehat{N}^{setup}(z) \;=\; \pi_0 \frac{z\widehat{A}_S(z)\widehat{A}_I(z) - \widehat{A}_S(z)}{z - \widehat{A}_S(z)}.$$

You should get:

$$\pi_0 = \frac{1 - \lambda \mathbf{E}[S]}{1 + \lambda \mathbf{E}[I]}.$$

14 Deriving $\widetilde{\mathbf{T}}^{\mathbf{setup}}(\mathbf{s})$

Now use Distributional Little's Law to go from

$$\widehat{N}^{setup}(z) \;=\; \pi_0 \frac{z\widehat{A}_S(z)\widehat{A}_I(z) - \widehat{A}_S(z)}{z - \widehat{A}_S(z)}$$

to

$$\widetilde{T}^{setup}(s) \;=\; \pi_0 \frac{(\lambda - s)\widetilde{S}(s)\widetilde{I}(s) - \lambda\widetilde{S}(s)}{\lambda - s - \lambda\widetilde{S}(s)}$$

15 After substituting in π_0, you get ...

$$\widetilde{T}_Q^{setup}(s) = \frac{(1-\rho)s}{s-\lambda+\lambda\widetilde{S}(s)} \cdot \frac{\lambda-(\lambda-s)\widetilde{I}(s)}{(1+\lambda\mathbf{E}[I])s}$$

$$= \widetilde{T_Q}^{\text{M/G/1}}(s) \cdot \frac{\lambda-(\lambda-s)\widetilde{I}(s)}{(1+\lambda\mathbf{E}[I])s}.$$

What's important here is that the first term involves delay for an M/G/1 without setup and the second involves only setup-related terms.

16 Translating to mean delay we have ...

$$\mathbf{E}[T_Q]^{setup} = \mathbf{E}[T_Q]^{M/G/1} + \frac{2\mathbf{E}[I] + \lambda\mathbf{E}\left[I^2\right]}{2(1+\lambda\mathbf{E}[I])}$$

It is really neat (and rare) when results decompose in such a pretty way. Such results are referred to as **decomposition results**.

17 Discussion: When does DLL apply?

Question: We've applied DLL to an M/G/1 and an M/G/1/setup. What was needed to make DLL work, and where else might it apply?