The Posterior Service Time in an M/G/1 Queue with a Workload Barrier and Extreme Prior Service Times

[a*†]Percy H. Brill and [b†]Mei Ling Huang
[a]University of Windsor, Canada, [b]Brock University, Canada

September 15, 2018

Abstract

This article considers an M/G/1 queue with workload barrier at level $K > 0$. Prior service times are continuous random variables. The policy for arrivals re the barrier is: service times that would cause the workload to exceed K are truncated at K. We give the posterior pdf and expected value of the posterior service time due to the barrier, the expected number served in a busy period, and related quantities. We use a metric for the distance between the prior and posterior pdfs of service time. We specialize results to the case where prior service times have a no-mean Pareto(II) distribution.

Key Words: M/G/1 queue, bounded workload, truncated service times, regenerative process, renewal theory, integral equations, level crossing

1. Introduction

Consider the workload process in a workload-barrier M/G/1 queue with a barrier at level $K > 0$. A sample path of workload is given in Fig. 1. Denote the posterior service time by S_K (called *posterior* due to presence of barrier at level K); busy period $\mathcal{B}_K$; stationary probability of a zero workload by $P_{K,0}$; number served in $\mathcal{B}_K$ by $N_{\mathcal{B}_K}$. Then $P_{K,0}$ exists, and $\mathcal{B}_K$ is finite for all $\lambda > 0$ *due to the barrier*. Customers with truncated service, exit the system when their total time in the system reaches K.

Brill, 2015[3] deals with this model when $E(S) < \infty$, and does not consider *extreme* prior service times. The present paper extends the discussion in [3] to the case where the prior S is distributed as an *extreme* Pareto(II) random variable with shape parameter $\alpha \in (0, 1]$ (Arnold, 2015[1]; Klieber and Klotz, 2003([13])); hence in the ***no***-workload-barrier M/G/1 queue $E[S] = \infty$, $P_0 = 0$ and $E[\textit{busy period}] = \infty$. In that case results will not hold, that have formulas with explicit P_0 or $E[S]$ in the ***no***-workload-barrier M/G/1 queue where $E(S) < \infty$. This situation gives rise to the questions: "Which results will still hold in the *workload-barrier* M/G/1 queue?" and "How far apart are the stationary pdfs of the prior and posterior service time pdfs re a metric designed to measure that distance?" We use such a metric in Section 6.

In the ***no***-workload-barrier M/G/1 queue, if P_0 exists and $E[S] < \infty$, we use the same notation as above, but omit the subscript K. It is well known that if $\lambda E[S] < 1$ then P_0 =

*Corresponding author, e-mail brill@uwindsor.ca
†This research is supported by the Natural Sciences and Engineering Research Council of Canada.

$1-\lambda E\left[S\right]$ is the stationary probability of a zero workload (p. 227 & p. 235 in Gross et al., 2008[8]). Also, $E\left[N_{\mathcal{B}}\right]=1/P_0$ (p. 233 in Cooper, 1990[6]). Thus, $E\left[S\right]=\left(1-P_0\right)/\lambda$. Also, $E\left[\mathcal{B}\right]=\left(1-P_0\right)/\left(\lambda P_0\right)$ (pp. 82-84 in Brill, 2017[2]). Therefore

$$\frac{E\left[\mathcal{B}\right]}{E\left[N_{\mathcal{B}}\right]}=\frac{\frac{1-P_0}{\lambda P_0}}{\frac{1}{P_0}}=\frac{1-P_0}{\lambda}=E\left[S\right]. \tag{1}$$

Equation (1) partially motivates this article. Here we explore the validity of the hypothesis "$E\left[\mathcal{B}_K\right]/E\left[N_{\mathcal{B}_K}\right]=E\left[S_K\right]$", *even if* $E\left[S\right]$ *and* P_0 *do not exist*. The hypothesis does turn out to be true (Section 4.1 below). It also turns out that $P_{K,0}$ exists, and is a *fundamental quantity in both* $E\left[N_{\mathcal{B}_K}\right]$ and $E\left[S_K\right]$.

The main motivating factor is to explore the model characteristics when $S \underset{d}{=}$ Pareto(II) with $\alpha\in(0,1]$. The results obtained here are potentially applicable in, e.g., risk models with a dividend barrier (e.g., Brill and Yu, 2011[4]; Gerber and Shiu, 1997[7]); queues occurring in industry (Harris (1968)[9]), or on the internet (Harris et al. (2000)[10]); input amounts in dams, and demand sizes in production inventories, with finite capacities (Chapter 6 in [2]); etc.

Section 2 details the workload process when the barrier is at fixed level $K\in(0,\infty)$. Section 3 derives $E\left[N_{\mathcal{B}_K}\right]$. Section 4 states the pdf of S_K, and derives $E\left[S_K\right]$. Section 5 discusses the results when $S \underset{d}{=}$ Pareto(II), and uses a metric for the distance between the posterior and prior pdfs. Section 7 discusses a computational algorithm for obtaining the pdfs of interest.

2. Workload in the Workload-barrier M/G/1 Queue

Denote the workload process with a barrier at K by $\{W_K(t)\}_{t\geq 0}$ (Fig. 1). The state space is $[0,\infty)$; this accounts for excess over level K. Note that $S_K \underset{d}{\neq} S$, because $S_K\leq S$ (formula (2) below).

If an arrival at time τ^- "sees" $W_K(\tau^-)=0\leq y<K$, then its posterior service time S_K is related to S by

$$S_K=\min(S,K-y). \tag{2}$$

In (2), if $S>K-y$ then $W_K(\tau)=K$; if $S\leq K-y$ then $W_K(\tau)=y+S$.

The pdf $\{P_{K,0},f_K(x)\}_{0<x<K}$ is the *stationary mixed pdf of* $\{W_K(t)\}_{t\geq 0}$ *as* $t\to\infty$. Then

$$P_{K,0}=\lim_{t\to\infty}P(W_K(t)=0) \text{ and } P_{K,0}+\int_{x=0}^{K}f_K(x)dx=1. \tag{3}$$

The pdf $\{P_{K,0},f_K(x)\}_{0<x<K}$ exists since the workload is a regenerative process re busy cycles (e.g., Sigman and Wolff, 1993). Thus $P_{K,0}$ and $f_K(x), 0<x<K$ exist regardless whether $E\left[S\right]=\infty$ and $P_0=0$ or not. (The barrier has a strong effect on the statistical properties of the system.)

2.1. $\mathcal{B}_K$ is Finite $a.s.$ for Every $\lambda>0$

The article Brill, 2015[3] shows that

$$E(\mathcal{B}_K)<Ke^{\lambda K}<\infty. \tag{4}$$

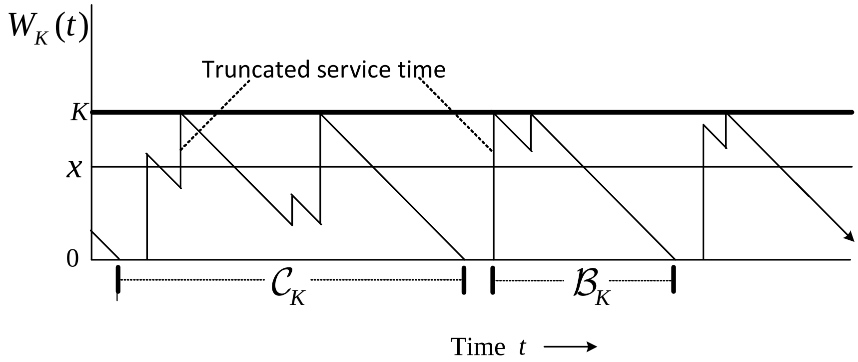

Figure 1: Sample path of $\{W_K(t)\}_{t\geq 0}$ in M/G/1 with bounded workload at level K. $\mathcal{C}_K$:= busy cycle, $\mathcal{B}_K$:= busy period, $\mathcal{I}_K$:= idle period, $\mathcal{B}_K$ $+\mathcal{I}_K = \mathcal{C}_K$

3. $E(N_{\mathcal{B}_K})$ in the Workload-barrier M/G/1 Queue

Let: $\mathcal{C}_K$:= *busy cycle* = $\mathcal{B}_K$ + $\mathcal{I}_K$, where $\mathcal{I}_K$:= *idle period;* A_t = *number of customer arrivals in the time interval* $(0, t)$, $t > 0$; $A_{\mathcal{C}_K}$ = *number of arrivals during* $\mathcal{C}_K$. It is shown in [3] that

$$E(N_{\mathcal{B}_K}) = \frac{1}{P_{K,0}}. \tag{5}$$

Briefly, (5) follows from: $N_{\mathcal{B}_K} = A_{\mathcal{C}_K}$; $\{A_t\}_{t\geq 0}$ is a Poisson process with rate λ, $\lim_{t\to\infty} A_t/t = \lambda$;

$$E(\mathcal{C}_K) = \frac{1}{f_K(0^+)} = \frac{1}{\lambda P_{K,0}}; \tag{6}$$

and application of the renewal reward theorem, giving

$$\frac{E(A_{\mathcal{C}_K})}{E(\mathcal{C}_K)} = \lim_{t\to\infty} \frac{A_t}{t} = \lambda,$$

implying

$$\frac{E(N_{\mathcal{B}_K})}{E(\mathcal{C}_K)} = \frac{E(A_{\mathcal{C}_K})}{E(\mathcal{C}_K)} = \lambda, \text{ and } E(N_{\mathcal{B}_K}) = \lambda E(\mathcal{C}_K) = \lambda\left(\frac{1}{\lambda P_{K,0}}\right) = \frac{1}{P_{K.0}}. \tag{7}$$

proving (5).

4. Posterior pdf $b_K(\cdot)$ and $E(S_K)$

We denote the *prior* pdf and cdf (cumulative distribution function) of S by $b(x), x > 0$, and $B(x), x > 0$ respectively, with $\overline{B}(x) = 1 - B(x), x \geq 0$; and the posterior pdf of S_K by $\{\pi_K, b_K(x)\}_{0<x<K}$, where π_K := $P(S_K = K)$ (atom at level K). We assume that S is absolutely continuous. The article [3] shows that the pdf $\{\pi_K, b_K(x)\}_{0<x<K}$ is given in terms of the pdf $\{P_{K,0}, f_K(x)\}_{0<x<K}$ and the probability distribution of S by

$$\left.\begin{array}{l} \pi_K = \overline{B}(K)P_{K,0}, \\ b_K(x) = b(x)P_{K,0} + \overline{B}(x)f_K(K-x) + \int_{y=0}^{K-x} b(x)f_K(y)dy, 0 < x < K. \end{array}\right\} \tag{8}$$

The formulas in (8) *do not explicitly contain* P_0 *or* $E[S]$. *Therefore* (8) *applies even if* $P_0 = 0$ *and* $E[S] = \infty$. Hence (8) applies when S := Pareto(II) with shape parameter $\alpha \in (0, 1]$.

The article [3] shows, using (8), that the following necessary condition holds (law of total probability)

$$\pi_K + \int_{x=0}^{K} b_K(x)dx = 1; \tag{9}$$

also, in the workload-barrier M/G/1 queue

$$E(S_K) = \frac{1 - P_{K,0}}{\lambda}. \tag{10}$$

To prove (10) we use a level crossing method (Chapter 1 in [2]) to obtain an integral equation for the continuous part of the posterior stationary pdf of the *workload*, i.e.,

$$f_K(x) = \lambda P_{K,0}\overline{B}(x) + \lambda \int_{y=0}^{x} \overline{B}(x-y) f_K(y)dy,\ 0 < x < K. \tag{11}$$

Integrating both sides of (11) with respect to $x \in (0, K)$, applying (9), and dividing by λ, gives

$$\frac{1 - P_{K,0}}{\lambda} = P_{K,0} \int_{x=0}^{K} \overline{B}(x)dx + \int_{x=0}^{K} \int_{y=0}^{x} \overline{B}(x-y) f_K(y)dydx. \tag{12}$$

In (12) integrate the first integral by parts, interchange the order of integration in the double integral, and integrate the inner integral by parts, which leads to

$$\begin{aligned} \tfrac{1-P_{K,0}}{\lambda} &= KP_{K,0}\overline{B}(K) \\ &\quad + \textstyle\int_{x=0}^{K} x \left(P_{K,0} b(x)dx + \overline{B}(x) f_K(K-x)dx + \int_{y=0}^{K-x} b(x) f_K(y)dy \right) dx \\ \tfrac{1-P_{K,0}}{\lambda} &= K\pi_K + \textstyle\int_{x=0}^{K} x b_K(x)dx = E[S_K] \end{aligned} \tag{13}$$

upon substituting from (8), and applying the definition of expected value.

4.1. $E[\mathcal{B}_K]/E(N_{\mathcal{B}_K}) = E[S_K]$

From Section 3, $E[\mathcal{I}_K] = \frac{1}{\lambda}$ because arrivals occur in a Poisson process at rate λ and all arrivals join the system. (All arrivals join, but some get only partial service.) Therefore

$$E[\mathcal{B}_K] = E[\mathcal{C}_K] - E[\mathcal{I}_K] = \frac{1}{\lambda P_{K,0}} - \frac{1}{\lambda} = \frac{1 - P_{K,0}}{\lambda P_{K,0}}.$$

Using (10) and (13) gives

$$\frac{E[\mathcal{B}_K]}{E(N_{\mathcal{B}_K})} = \frac{\frac{1-P_{K,0}}{\lambda P_{K,0}}}{\frac{1}{P_{K.0}}} = \frac{1 - P_{K,0}}{\lambda} = E[S_K] \tag{14}$$

which parallels the result in (1) for the unbounded-workload M/G/1 queue.

5. Model where S is Distributed as $Pareto(II)$ with Shape Parameter α

The cdf, ccdf and pdf of S are, respectively,

$$\begin{aligned} B(x) &= 1-(1+x)^{-\alpha}, 0<x<\infty, \\ \overline{B}(x) &= 1-B(x) = (1+x)^{-\alpha}, 0 \le x < \infty, \\ b(x) &= \tfrac{d}{dx}B(x) = \alpha\,(1+x)^{-\alpha-1}, 0<x<\infty. \end{aligned} \tag{15}$$

From (11) and (15) we get

$$f_K(x) = \lambda P_{K,0}\,(1+x)^{-\alpha} + \lambda \int_{y=0}^{x} (1+x-y)^{-\alpha} f_K(y)dy, 0<x<K. \tag{16}$$

The normalizing condition is

$$P_{K,0} + \int_{y=0}^{K} f_K(y)dy = 1. \tag{17}$$

Once we have the solutions of (16) and (17) for $f_K(x)$ and $P_{K,0}$, we can get explicit solutions for $E\,[S_K]$, π_K, $b_K(x), 0<x<K$, and other characteristics.

6. Metric for Distance Between Prior and Posterior pdfs of Service Tine

We use a modified form of the metric developed in Brill and Huang, 2017[5], which takes into account the atom with probability π_K at K in the pdf $\{\pi_K, b_K(x)\}_{0<x<K}$, namely, formula (19) below and inequality $\rho(b_K, b)$ in (20) below. It gives a measure of distance between the prior and posterior pdfs of the service time. The prior pdf is $b(x) = \alpha\,(1+x)^{-\alpha-1}, 0<x<\infty$ (formula (15)). The posterior pdf of service, $b_K(\cdot)$, is given in (8), and repeated here, while substituting from (15), i.e., for $0<x<K$

$$\begin{aligned} &\pi_K = (1+K)^{-\alpha} P_{K,0}, \\ &b_K(x) = \alpha\,(1+x)^{-\alpha-1} P_{K,0} \\ &\qquad + (1+x)^{-\alpha} f_K(K-x) + \textstyle\int_{y=0}^{K-x} (1+x)^{-\alpha} f_K(y)dy, \\ &\quad = \alpha\,(1+x)^{-\alpha}\left[(1+x)^{-1} P_{K,0} + f_K(K-x) + \textstyle\int_{y=0}^{K-x} f_K(y)dy\right]. \end{aligned} \tag{18}$$

The distance measure between $b_K(\cdot)$ and $b(\cdot)$ is defined as

$$\rho(b_K, b) = \frac{1}{2}\left(\int_{x=0}^{K} |b_K(x) - b(x)|\,dx + \left|\pi_K - \int_{K}^{\infty} b(x)dx\right|\right). \tag{19}$$

It can be shown from (19) that

$$0 < \rho(b_K, b) < 1. \tag{20}$$

7. Computational Algorithm

We use a computational algorithm to solve equations (16) and (17) for the stationary pdf of workload, $\{P_{K,0}, f_K(x)\}_{0<x<K}$. We substitute that solution into (8) to obtain π_K and $b_K(\cdot)$. This requires additional computation, because $P_{K,0}$ and $f_K(x)$ are themselves the

outputs of a computation. Then obtain the distance measure between $b_K(\cdot)$ and $b(\cdot)$ in (19), by using the resulting π_K and $b_K(\cdot)$ in (18).

Acknowledgement

This work is supported by the Natural Sciences and Engineering Research Council of Canada (NSERC).

References

[1] Arnold, B.C. (2015). Pareto Distributions, Second Edition, CRC Press, Taylor & Francis Group, Boca Raton.

[2] Brill, P.H., 2017. Level Crossing Methods in Stochastic Models, Second Edition. Springer. New York.

[3] Brill, P.H., 2015. Note on the service time in an M/G/1 queue with bounded workload. Statistics & Probability Letters, Volume **96**, January, 162-169.

[4] Brill, P. H., Yu, K., 2011. Analysis of Risk Models Using a Level Crossing Technique. Insurance: Mathematics and Economics. **49(3)**, 298–309.

[5] Brill, P.H., Huang, M.L. 2017. Approximating the finite-time t probability distributions in an extreme renewal process. In JSM Proceedings, Statistical Computing Section. Alexandria, VA: American Statistical Association, pp. 1393- 1401.

[6] Cooper, R.B., 1990. Introduction to Queueing Theory, 3^{rd} ed.. CEE PRESS Books, Washington, D.C.

[7] Gerber, H.U., Shiu., E.S.W., 1997. The Joint Distribution of the Time of Ruin, the Surplus Immediately Before Ruin, and the Deficit at Ruin. Insurance: Mathematics and Economics. **21**, 129-137.

[8] Gross, D., Shortle, J. F., Thompson, J. M., Harris, C. M., 2008. Fundamentals of Queueing Theory, 4^{th} ed.. Wiley, New York.

[9] Harris, C.M., (1968). The Pareto Distribution as a Queue Service Discipline, Operations Research, **16**(2), pp. 307–313.

[10] Harris, C.M., Brill, P.H., Fischer, M.J. (2000). Internet-Type Queues with Power-Tailed Interarrival Times and Computational Methods for Their analysis, INFORMS J. on Computing, 12(4), 261-271.

[11] Huang, M.L., Coia, V., Brill, P.H. (2013). A Cluster Truncated Pareto Distribution and Its Applications, ISRN Probability and Statistics, Volume 2013, Article ID 265373, 10 pages, http://dx.doi.org/10.1155/2013/265373.

[12] Karlin, S., Taylor, H.M., 1975. *A First Course in Stochastic Processes, 2^{nd} ed.*. Academic Press, New York.

[13] Kleiber, C., Kotz, S. (2003), Statistical Size Distribution in Economics and Actuarial Sciences, John Wiley, New York.

[14] Ross, S.M., 2010. Introduction to Probability Models, 10th ed. Elsevier, Oxford, U.K.

[15] Sigman, K. (1999). A Primer on Heavy-tailed Distributions, *Queueing Systems*, **33**, 261-275**.**

[16] Sigman, K., Wolff, R. W., 1993. A Review of Regenerative Processes. SIAM Review. **35**(2), 269-288.

[17] Wolff, R.W., 1982. Poisson Arrivals See Time Averages. Operations Research,. **30(**2**)**, 223-231.