

# Modeling Arrivals to the Agents Queue in a Call Center

Sivan Aldor

Paul D. Feigin  
Technion—Israel Institute of Technology

September, 2007

# 1 Introduction

The arrival process of calls to a call center is usually associated with a Poisson process. The Poisson process model provides simple solutions to many different problems related to call center operations management. However, modern call centers have automated answering and routing services that may operate in a way that leads to a non-Poissonian output process. This output process forms the arrival process to a downstream system, such as the human agent queueing and service system. When analyzing the latter system, departure from Poisson arrivals may affect results and decisions. We illustrate this non-Poisson arrival phenomenon using real data, propose a model that could explain how it may have arisen, and consider its practical ramifications on system analysis.

The main assumptions of the Poisson arrival model are that the number of incoming calls per interval follows the Poisson distribution, and that the numbers of arrivals in disjoint intervals are independent. We are interested in studying the arrivals to the agents service queue, which consists of those customers who have passed through the VRU (Voice Response Unit) and have sought an agent's help. Note that not all the incoming calls proceed to the agent queue: in many call centers most of the incoming calls end their service at the VRU (since they do not require a human agent). Figure 1 illustrates the two stage system. Different studies have tested and determined that the first VRU arrival process indeed follows a Poisson process. Surprisingly at first, tests of the arrival process to the agents queue have shown that it does not behave like a Poisson process.

Our paper will focus on modeling the arrival process of calls to the agents service queue. The idea for the model was motivated by empirical analyses of data collected as part of the Data-MOCCA (Data Model for Call Center Analysis) project [4]. This project is being carried out at the Technion's Service Enterprise Engineering (SEE) Center. One aim of the project is to provide a standardized format for creating and maintaining call center databases, which is independent of the data source, and which stores information at the individual call level. This data repository design differs from the usual summaries routinely available from most call center systems. Data-MOCCA is currently populated using data from several call centers; however, in this paper we concentrate on data which originated from the call center of a large Bank in the USA. This data were also analyzed in other papers such as [2] and [5].

## 2 Empirical Distribution

In Brown *et al* [3], one hypothesis tested was whether the arrival process to the service queue in a call center follows an inhomogeneous Poisson process. In order to test this hypothesis they assumed that the arrival rate remains constant during

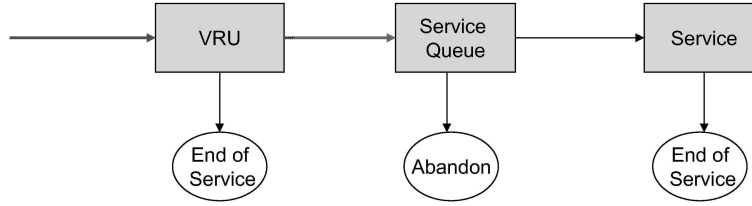


Figure 1: A simplified process flow of an incoming call to the call center.

short six-minute intervals. They tested this hypothesis over two separate hours: between 7:00-8:00, at the beginning of the working day; and between 12:00-13:00, at the peak of the working day. In the above paper they came to the conclusion that the null hypothesis is rejected; that is, the arrival process to the agent service queue does not behave like a Poisson process.

In order to fully understand how the arrival process to the agents' queue deviates from a Poisson process we compared the empirical process with the Poisson model. We considered the same two hours on June 11, 2002, which is a regular work day (i.e. there is no holiday on this date or any abnormal activity). Using the same method, we broke down the hours into six-minute intervals. Under the Poisson assumption for each six-minute block, the number of arrivals for each second should be approximately Poisson distributed. For each six-minute block we calculated the estimated arrival rate and derived the expected distribution of the number of arrivals for each second assuming a regular Poisson process. An example for a given six-minute period from each hour is given in Tables 1 and 2. The comparisons showed an interesting phenomenon: on the one hand there were “too many” seconds with no arrivals compared to a regular Poisson distribution. On the other hand, in a regular Poisson distribution with small  $\lambda$  the probability of large values is virtually zero but in our data we see several seconds with “too many” arrivals.

### 3 The Gatekeeper Model

Based on previous research (for example [1]), we continue to assume that calls entering the call center follow a Poisson process. However, upon exiting the VRU and entering the service queue they apparently “lose” this Poisson property. We suggest a relatively simple model to describe the observed service queue entrance process.

Our model describes a stochastic mechanism which controls the number of calls which leave the VRU and join the service queue. One can imagine a gatekeeper

|                           |        |        |         |         |         |     |
|---------------------------|--------|--------|---------|---------|---------|-----|
| Number of Calls $m_i$     | 0      | 1      | 2       | 3       | 4+      | sum |
| Actual Frequency $f_i$    | 277    | 58     | 20      | 3       | 2       | 360 |
| Theoretical Probabilities | 0.7428 | 0.2208 | 0.03281 | 0.00325 | 0.00026 | 1   |
| Expected Frequency        | 267.43 | 79.48  | 11.81   | 1.17    | 0.09    | 360 |

Table 1: Actual and expected number of seconds with given number of calls during 7:00-7:06. The estimated arrival rate is  $\frac{\sum_i f_i \cdot m_i}{360} = 0.2972$ .

|                           |         |         |       |        |       |       |      |     |
|---------------------------|---------|---------|-------|--------|-------|-------|------|-----|
| Number of Calls $m_i$     | 0       | 1       | 2     | 3      | 4     | 5     | 6+   | sum |
| Actual Frequency $f_i$    | 189     | 78      | 47    | 27     | 13    | 5     | 1    | 360 |
| Theoretical Probabilities | 0.393   | 0.367   | 0.171 | 0.053  | 0.012 | 0.002 | 0    | 1   |
| Expected Frequency        | 141.567 | 132.129 | 61.66 | 19.183 | 4.476 | 0.839 | 0.15 | 360 |

Table 2: Actual and expected number of seconds with given number of calls during 12:00-12:06. The estimated arrival rate is  $\frac{\sum_i f_i \cdot m_i}{360} = 0.9333$ .

which controls the calls that enter the service queue in the following way: every second it sees calls that have finished their VRU service and are waiting to be transferred from the VRU to the service queue. It then “decides” with probability  $1-p$  to admit all these waiting calls to the agent service queue, or with probability  $p$  to momentarily “deny” them this passage. This model may reflect the fact that at each second, the gatekeeper processor may be pre-occupied with some other task (e.g. dealing with another service type) and so not be able to control the admission of calls to the particular service queue.

In subsequent seconds the gatekeeper will repeat the procedure independently of its decision in the previous seconds. Of course the number of calls requiring passage consists of the number of calls which have accumulated from the previous “blocked” seconds plus the number of calls which have just completed VRU service.

It follows easily that if  $p$  equals zero then the gatekeeper model reduces to a regular Poisson process and if  $p$  equals 1 then no calls will ever pass through to obtain service.

One can imagine that each second the gatekeeper flips a coin (independently of previous seconds) with a probability of  $1-p$  to open the gate and  $p$  to keep it closed. When the gatekeeper opens the gate all the calls which were in the exit queue during that second (including the ones which just finished their VRU operations) pass through to the service queue. If he does not open the gate the calls will accumulate and then the process will start over. Figure 2 illustrates the model.

We use the following variables to formalize the model:

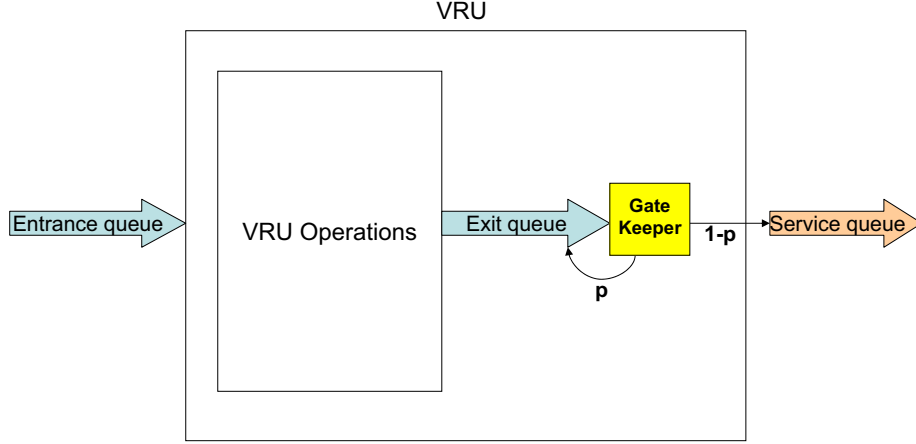


Figure 2: The gatekeeper model. There are three queues: the entrance queue to the VRU; the exit queue from the VRU; and the service queue. The gatekeeper stands between the exit queue from the VRU and the service queue, each second allowing all calls in the exit queue to pass through to the service queue with probability  $1 - p$ .

- $Y_n$  is the number of calls which *enter* the exit queue during the  $n^{th}$  second (i.e. the calls which just finish their VRU operations during the  $n^{th}$  second). We assume that for a short interval the rate of arrivals to this queue is fixed and that the distribution of this variable is Poisson, i.e.

$$Y_n \sim \text{Poisson}(\lambda) \quad (1)$$

- $B_n$  is the gatekeeper variable, i.e.

$$B_n = \begin{cases} 1 & \text{if the gate is blocked during the } n^{th} \text{ second} \\ 0 & \text{if the gate is opened during the } n^{th} \text{ second} \end{cases} \quad (2)$$

- $X_n$  is the number of *accumulated* calls in the exit queue at the end of the  $n^{th}$  second (including  $Y_n$ ) that did not pass through to the service queue, i.e.

$$X_n = B_n \cdot (X_{n-1} + Y_n) \quad (3)$$

It can be easily seen that if the gatekeeper opens the gate at the  $n^{th}$  second then  $X_n$  equals zero, meaning that the exit queue is empty.

- $T_n$  is the number of calls which passed from the VRU exit queue to the service queue in the  $n^{th}$  interval, i.e.

$$T_n = (1 - B_n) \cdot (X_{n-1} + Y_n) \quad (4)$$

It is easily seen that if the gatekeeper closed the gate at the  $n^{th}$  second then  $T_n$  equals zero, meaning that no new calls entered the service queue.

The model assumptions are:

- The gatekeeper's decision is not influenced by the number of calls which entered the exit queue at any point of time and vice versa, i.e.

$$B_n \perp Y_m \quad \text{for all } m, n \quad (5)$$

where  $\perp$  denotes statistical independence. It then follows that the number of calls accumulated in the exit queue has no influence on the future number of calls which will enter the exit queue and vice versa, i.e.

$$Y_n \perp X_m \quad \text{for } n > m \quad (6)$$

From Definition (3) it follows that  $X_n$  is a simple Markov chain with transitions probabilities:

$$\begin{aligned} \Pr(X_n = i | X_{n-1} = j) &= \Pr(B_n = 1) \cdot \Pr(Y_n = i - j) \\ &= p \cdot \frac{e^{-\lambda} \cdot \lambda^{i-j}}{(i-j)!}, \quad \text{for } i \geq j > 0 \end{aligned} \quad (7)$$

$$\Pr(X_n = 0 | X_{n-1} = j) = 1 - p + p \cdot e^{-\lambda} \cdot I_{\{j=0\}}, \quad \text{for } j \geq 0 \quad (8)$$

The relation between  $X_n$  and any previous  $X_m$  (i.e. for  $n > m$ ) can be formalized in the following manner:

$$X_n = X_m \cdot \prod_{i=m+1}^n B_i + \sum_{i=m+1}^n \left( Y_i \cdot \prod_{l=i}^n B_l \right) \quad (9)$$

$$E(X_n) = p^{n-m} \cdot E(X_m) + \frac{\lambda p \cdot [1 - p^{n-m}]}{1 - p} \quad (10)$$

$$E(X_n \cdot X_m) = p^{n-m} \cdot E(X_m^2) + E(X_m) \cdot \frac{\lambda p}{1 - p} \cdot (1 - p^{n-m}) \quad (11)$$

### 3.1 Steady state behavior at one second resolution

Looking at the steady state of the above Markov chain we can derive the following results, assuming that  $0 < p < 1$ :

$$E(X) = p \cdot [E(X) + E(Y)] \Leftrightarrow E(X) = \frac{\lambda \cdot p}{1 - p} \quad (12)$$

$$E(X^2) = E(B^2 \cdot (X + Y)^2) \Leftrightarrow E(X^2) = \frac{\lambda p}{1 - p} + \frac{\lambda^2 p(1 + p)}{(1 - p)^2} \quad (13)$$

$$\text{Var}(X) = \left( \frac{\lambda p}{1 - p} \right) \left[ 1 + \frac{\lambda}{1 - p} \right] \quad (14)$$

$$\text{CV}(X) = \sqrt{\frac{\lambda + 1 - p}{\lambda p}} \quad (15)$$

For the case  $p = 0$ , where the gatekeeper always keeps the gate open,  $X_n \equiv 0$  and the chain is degenerate; and for the case  $p = 1$ , the chain is transient.

Let  $\pi_k$  be the (stationary) probability that the Markov chain  $X_n$  is in state  $k$ , i.e.  $\pi_k = \Pr(X_n = k)$ . It can be seen that  $\pi_k$  can be calculated from the following recursion:

$$\pi_k = \frac{p}{1 - p \cdot e^{-\lambda}} \sum_{i=0}^{k-1} \pi_i \cdot P(Y = k - i) \quad (16)$$

$$\pi_0 = \frac{1 - p}{1 - p \cdot e^{-\lambda}} \quad (17)$$

What is observed as arrivals to the agents service queue is  $T_n$ , the number of calls which passed through the gatekeeper during the  $n^{\text{th}}$  second. From Definition (4) we derive the following conditional probabilities:

$$\Pr(T_n = 0 | X_{n-1} = j) = p + (1 - p) \cdot e^{-\lambda} \cdot I_{\{j=0\}} \text{ for } j \geq 0 \quad (18)$$

$$\begin{aligned} \Pr(T_n = i | X_{n-1} = j) &= (1 - p) \cdot \Pr(Y_n = i - j) \\ &= (1 - p) \cdot \frac{e^{-\lambda} \cdot \lambda^{(i-j)}}{(i-j)!}, \quad \text{for } i \geq j > 0 \end{aligned} \quad (19)$$

The following characteristics of  $T_n$  can also be derived from Definition (4):

$$E(T_n) = (1 - p) \cdot [E(X_{n-1}) + \lambda] \quad (20)$$

$$\begin{aligned} E(T_n \cdot T_m) &= (1 - p) [E(X_{m-1}) \cdot \lambda(1 - p^{n-m}) \\ &\quad + \lambda^2 p \cdot (1 - p^{n-m-1}) + \lambda^2(1 - p)] \end{aligned} \quad (21)$$

Looking at the steady state we arrive at the following results for  $p < 1$ :

$$E(T) = (1 - p) \cdot [E(X) + E(Y)] = (1 - p) \cdot \left[ \frac{\lambda}{1 - p} + \lambda \right] = \lambda \quad (22)$$

$$E(T^2) = E((1 - B)^2(X + Y)^2) = \frac{\lambda^2(1 + p)}{1 - p} + \lambda \quad (23)$$

$$\text{Var}(T) = \frac{2p\lambda^2}{1 - p} + \lambda \quad (24)$$

$$\text{CV} = \frac{\sqrt{\frac{2p\lambda^2}{1 - p} + \lambda}}{\lambda} = \sqrt{\frac{2p}{1 - p} + \frac{1}{\lambda}} \quad (25)$$

For the special case where  $p = 0$ ,  $T_n$  becomes a regular Poisson variable with the result that  $\text{Var}(T_n)$  equals  $\lambda$  and the coefficient of variation equals  $\frac{1}{\sqrt{\lambda}}$ .

From Equation (18) we can deduce the limit distribution of  $\Pr(T_n = i)$ :

$$\Pr(T = i) = \begin{cases} p + (1 - p) \cdot e^{-\lambda} \cdot \pi_0 = \frac{p + e^{-\lambda} \cdot (1 - 2p)}{1 - p \cdot e^{-\lambda}} & \text{if } i = 0 \\ (1 - p) \cdot \sum_{j=0}^i P(Y = i - j) \cdot \pi_j & \text{if } i > 0 \end{cases} \quad (26)$$

### 3.2 Steady state behavior at coarser resolutions

An interesting question that arises from the model is “can we observe Poisson-like behavior if we examine longer intervals?”. In other words, if we count the number of arrivals during groups of successive intervals of  $n$ -seconds, rather than during each individual second, will these grouped counts have a more Poisson-like distribution. In order to answer the above question we are interested in the stationary behavior of  $\sum_{i=1}^n T_i$ . The following characteristics can be computed:

$$E\left(\sum_{i=1}^n T_i\right) = n\lambda \quad (27)$$

$$\text{Var}\left(\sum_{i=1}^n T_i\right) = n\lambda + 2\lambda^2 \cdot \frac{p(1 - p^n)}{(1 - p)^2} \quad (28)$$

$$\text{CV}\left(\sum_{i=1}^n T_i\right) = \sqrt{\frac{1}{n\lambda} + 2p \frac{(1 - p^n)}{(1 - p)^2 n^2}} \quad (29)$$

The three equations show the deviation of the expectation, variance and CV from the corresponding Poisson characteristic. The first term, in all three equations, is the Poisson characteristic. The second terms in equations (28) and (29) show how the deviation from Poisson-ness depends on  $p$ ,  $\lambda$  and  $n$ . For large enough values of  $n$  the CV characteristic will approximate that of the Poisson distribution

regardless of the value of  $p$ . Figure 3 illustrates the CV ratio between the sum of regular Poisson variables, which is just a Poisson variable with parameter  $n\lambda$  and  $\sum_{i=1}^n T_i$  for various values of  $\lambda$  and  $p$ . The reader is referred to the Appendix for the derivations of equations (28) and (29).

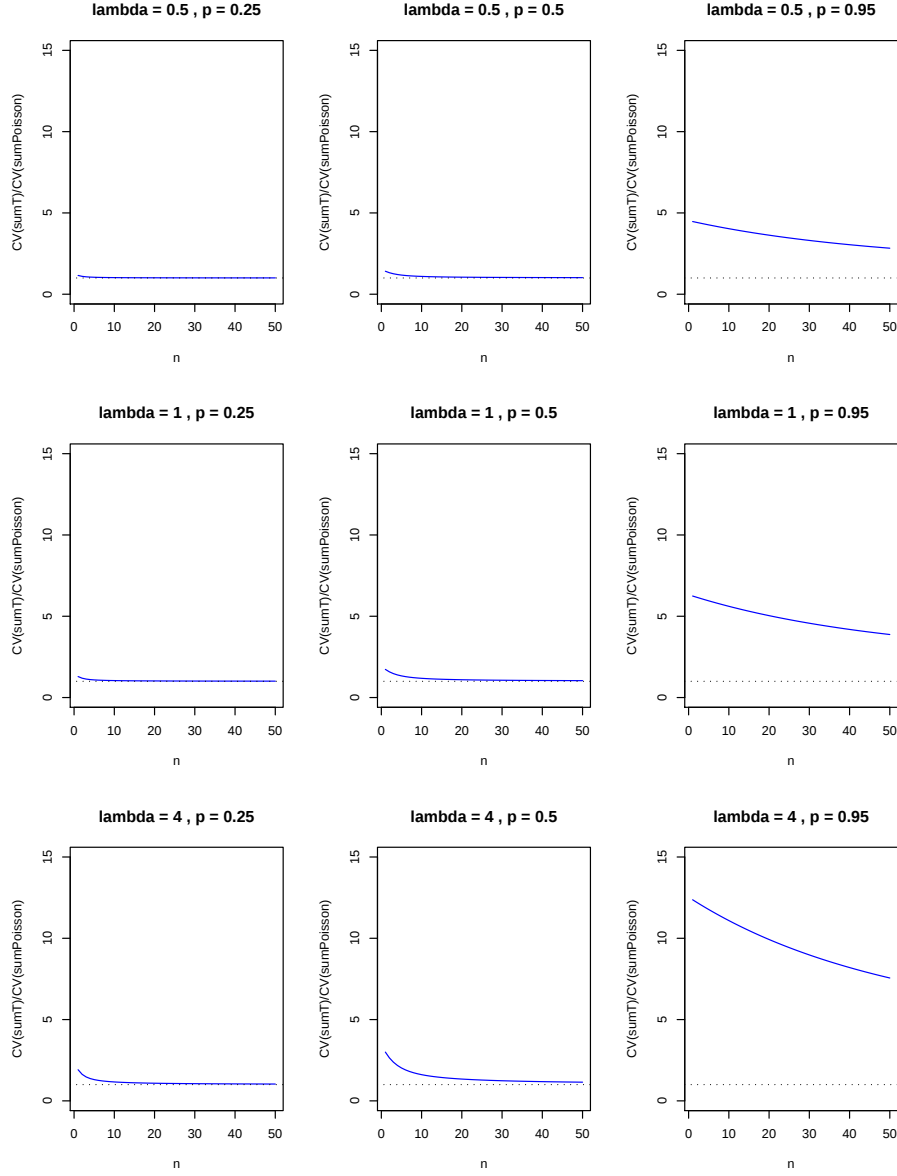


Figure 3: CV comparison. Each graph illustrates the ratio between  $CV(\sum_{i=1}^n M_i)$  (where  $M_i \sim \text{Poisson}(\lambda)$ ) and  $CV(\sum_{i=1}^n T_i)$ , as a function of  $n$ . There are nine graphs. In each row we compare models with the same  $\lambda$  but with different values of  $p$ . In each column there are models with the same value of  $p$  but with different  $\lambda$ 's. The dotted line is a horizontal line with intercept 1. As  $p$  grows so does the number of intervals ( $n$ ) required for the ratio to approach 1. As  $\lambda$  grows so does the number of intervals ( $n$ ) required for the ratio to approach 1.

## 4 Model Fitting

As described previously we analyzed two hours: 7:00-8:00 and 12:00-13:00. We divided them into twenty six-minute intervals. In each interval we calculated the number of arrivals for each second. A frequency table for each interval was then computed (the  $i^{th}$  cell in the table counts the number of seconds during which exactly  $i$  arrivals occurred). These tables served as the observed frequency tables. Using our model, calculating the expected number of arrivals per second requires two parameters: the arrival rate  $\lambda$  and the probability  $p$  of a blocked second. The arrival rate per second can be easily estimated by the average number of calls per second over the six-minute interval. The estimate of  $p$  is determined as the solution to the following minimum chi-square problem:

$$\hat{p} = \arg \min_{0 < p < 1} \sum_{i=1}^k \frac{(O_i - E_i(\hat{\lambda}, p))^2}{E_i(\hat{\lambda}, p)} \quad (30)$$

The function is the well-known chi-square statistic where:

- parameter  $k$  is the number of bins with different call arrival frequencies. Note that the number of bins varied between intervals depending on the count of seconds with a larger number (3+, 4+ , 5+) of arrivals.
- $E_i(\hat{\lambda}, p)$  is the expected number of one-second intervals with  $i$  arrivals.  $E_i(\hat{\lambda}, p) = n \cdot P(T(\hat{\lambda}, p) = i)$  where  $P(T(\hat{\lambda}, p) = i)$  is defined according to Definition (26) and  $n$  is the number of seconds in the time window under study (in our case 360).
- $O_i$  is the observed frequency for cell  $i$ .

The number of degrees of freedom of the minimum chi-square statistic is  $k-3$ , after accounting for the two estimated parameters.

For each 6-minute interval we estimated  $\hat{\lambda}$  and  $\hat{p}$ . We used a standard minimization solver to estimate  $\hat{p}$ . Consequently, for each interval we obtained a chi-square statistic which now could be used to test the goodness of fit of our theoretical distribution. The results of the above test are shown in Table 3. Further examination of the p-values reveals that none of the hypotheses can be rejected, suggesting that the gatekeeper model fits well. It is important to note that fixing  $p = 0$ , corresponding to a Poisson arrival process, does not fit the data at all. For almost all six-minute intervals, the p-value is significant, rejecting the Poisson process model for arrivals to the agents service queue.

In Figure 4 we plot  $p$  (the probability for a blocked second) versus the arrival rate  $\lambda$ . For most of the intervals,  $\hat{p}$  has a value between 0.25 and 0.4.

| Starting time | Ending time | $\hat{p}$ | $\hat{\lambda}$ | Chi-Square statistic | df | p-value  |
|---------------|-------------|-----------|-----------------|----------------------|----|----------|
| 07:00         | 07:06       | 0.421     | 0.297           | 1.6740               | 2  | 0.433005 |
| 07:06         | 07:12       | 0.564     | 0.203           | 4.5196               | 3  | 0.210547 |
| 07:12         | 07:18       | 0.315     | 0.242           | 0.3374               | 1  | 0.561288 |
| 07:18         | 07:24       | 0.323     | 0.258           | 1.7829               | 2  | 0.410054 |
| 07:24         | 07:30       | 0.290     | 0.275           | 0.6344               | 2  | 0.728181 |
| 07:30         | 07:36       | 0.338     | 0.256           | 2.0061               | 1  | 0.156662 |
| 07:36         | 07:42       | 0.256     | 0.283           | 1.0632               | 1  | 0.302469 |
| 07:42         | 07:48       | 0.261     | 0.317           | 4.3491               | 2  | 0.113656 |
| 07:48         | 07:54       | 0.284     | 0.367           | 3.8414               | 2  | 0.146504 |
| 07:54         | 08:00       | 0.412     | 1.372           | 5.0688               | 3  | 0.166111 |
| 12:00         | 12:06       | 0.318     | 0.933           | 3.5012               | 4  | 0.477693 |
| 12:06         | 12:12       | 0.361     | 0.975           | 3.6791               | 4  | 0.451161 |
| 12:12         | 12:18       | 0.321     | 0.869           | 1.6599               | 3  | 0.645867 |
| 12:18         | 12:24       | 0.343     | 0.947           | 3.8446               | 5  | 0.571988 |
| 12:24         | 12:30       | 0.324     | 0.963           | 9.0676               | 4  | 0.059430 |
| 12:30         | 12:36       | 0.258     | 0.922           | 6.9987               | 6  | 0.320964 |
| 12:36         | 12:42       | 0.324     | 0.908           | 7.8848               | 6  | 0.246665 |
| 12:42         | 12:48       | 0.317     | 0.933           | 1.3427               | 5  | 0.930473 |
| 12:48         | 12:54       | 0.281     | 0.838           | 5.7747               | 4  | 0.216615 |
| 12:54         | 13:00       | 0.323     | 1.033           | 4.9943               | 5  | 0.416575 |

Table 3: Goodness of fit between the model results and the observed values. This table summaries the estimated  $p$ , estimated arrival rate  $\lambda$ , chi-square statistic and the appropriate p-value for each 6-minute interval between 7:00-8:00 and between 12:00-13:00.

## 5 Discussion

From the above analyses of data from a US bank call center, it is clear that when modeling the arrival process to the agents queue, one cannot assume that it is a Poisson process — at least not at the resolution of one second. The gatekeeper model suggested here does however provide an adequate description of the arrival process, with a probability of about 0.30 that for any given second the accumulated potential arrivals will be detained (blocked) for another second.

Of course, we have no reason to believe that we have discovered the actual mechanism that disturbs the Poisson-ness of the arrival process. On the other hand, it does give the software and hardware engineers some food for thought.

For macro decisions concerning call center operations, this aberration from the usual assumption of Poisson arrivals is probably of marginal importance. As can be seen from Figure 3, when considering the arrivals at a resolution of say 5

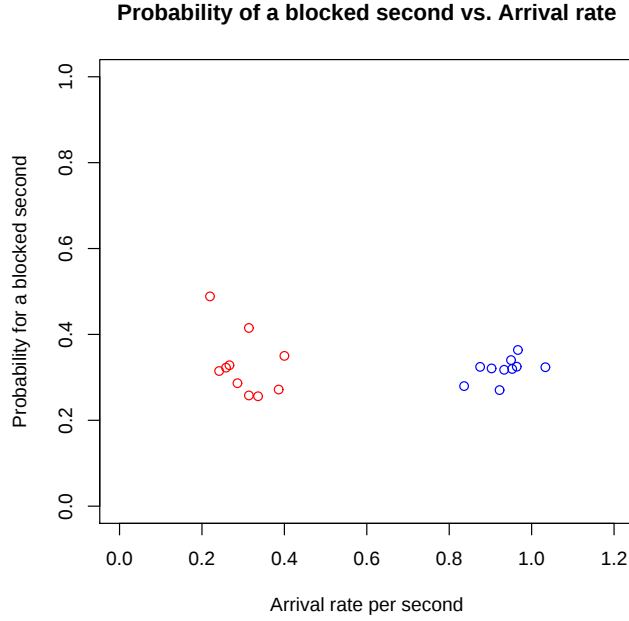


Figure 4: Probability of a blocked second as a function of the arrival rate (per second). Blue points correspond to intervals between 12:00-13:00 while red points intervals between 7:00-8:00

seconds, rather than 1 second, the difference between the Poisson CV and the Gatekeeper CV is small for  $p \approx 0.3$ .

## References

- [1] Lawrence D. Brown, Noah Gans, Avishai Mandelbaum, Anat Sakov, Haipeng Shen, Sergey Zeltyn, and Linda Zhao. Statistical analysis of a telephone call center: A queueing-science perspective. *JASA*, 100(469):36–55, March 2005.
- [2] Haipeng Shen and Jianhua Z. Huan. Analysis of call centre arrival data using singular value decomposition. *Applied Stochastic Models Bus.Ind.*, (21):251–263, 2005.
- [3] Yen Chu Cheng supervised by Lawrence D.Brown. Analysis of call center data. April 2004.
- [4] Valery Trofimov, Paul Feigin, Avishai Mandelbaum, and Eva Ishay. DataMOCCA – Data Model for Call Center Analysis. Technical report, Technion, Israel, 2003.
- [5] Jonathan Weinberg, Lawrence D. Brown, and Jonathan R. Stroud. Bayesian forecasting of an inhomogeneous Poisson process with applications to call center data. May 2005.

## Appendix Proof Outlines

In this appendix we will show how to calculate the coefficient of variation of  $\sum_{i=1}^n T_i$  and prove Equation (29). First we will prove Equation (28). Calculating the term  $(\text{Var } \sum_{i=1}^n T_i)$  requires a few steps.

**Step 1** *Proving the next three propositions:*

**Proposition 1** *For all  $i > j$  the following equation holds:*

$$E(B_j \cdot X_{j-1} \cdot X_{i-1}) = p^{i-j} \cdot E(X_{j-1}^2) + \left[ \frac{\lambda p^2}{1-p} - \frac{\lambda p^{i-j+1}}{1-p} + \lambda p^{i-j} \right] \cdot E(X_{j-1})$$

**Proposition 2** *For all  $i > j$  the following equation holds:*

$$E(Y_j \cdot X_{i-1}) = E(X_{j-1}) \cdot \lambda p^{i-j} + (\lambda^2 + \lambda) p^{i-j} + \frac{\lambda^2 p}{1-p} \cdot (1 - p^{i-j-1})$$

**Proposition 3** *For all  $i > j$  the following equation holds:*

$$E(B_j \cdot Y_j \cdot X_{i-1}) = E(X_{j-1}) \cdot \lambda p^{i-j} + (\lambda^2 + \lambda) p^{i-j} + \frac{\lambda^2 p^2}{1-p} \cdot (1 - p^{i-j-1})$$

**Step 2** *Calculating the term  $E(T_i \cdot T_j)$  for  $i > j$*

$$\begin{aligned} E(T_i \cdot T_j) &= E[(1 - B_i)(X_{i-1} + Y_i)(1 - B_j)(X_{j-1} + Y_j)] \\ &= (1 - p)[E(X_{i-1}X_{j-1}) - E(B_jX_{j-1}X_{i-1}) + E(X_{i-1}Y_j) - E(B_jY_jX_{i-1})] \\ &= (1 - p)[E(X_{j-1})(\lambda \cdot (1 - p^{i-j})) + \lambda^2 p \cdot (1 - p^{i-j-1}) + \lambda^2(1 - p)] \end{aligned}$$

**Step 3** *Calculating the term  $\text{Cov}(T_i, T_j)$  for  $i > j$*

$$\begin{aligned} \text{Cov}(T_i, T_j) &= E(T_i \cdot T_j) - E(T_i) \cdot E(T_j) \\ &= -\lambda^2 p^{i-j} \end{aligned}$$

**Step 4** *Calculating the term  $\text{Var}(\sum_{i=1}^n T_i)$*

$$\begin{aligned}
\text{Var}\left(\sum_{i=1}^n T_i\right) &= \sum_{i=1}^n \text{Var}(T_i) + 2 \sum_{i=2}^n \sum_{j=1}^{i-1} \text{Cov}(T_i, T_j) \\
&= n\lambda \cdot \left(1 + \frac{2\lambda p}{1-p}\right) - 2\lambda^2 \sum_{i=2}^n \sum_{j=1}^{i-1} p^{i-j} \\
&= n\lambda + 2\lambda^2 \frac{p(1-p^n)}{(1-p)^2}
\end{aligned}$$

Using the above results, it is easy to prove Equation 29 in the following manner:

$$CV\left(\sum_{i=1}^n T_i\right) = \frac{\sqrt{n\lambda + 2\lambda^2 \frac{p(1-p^n)}{(1-p)^2}}}{n\lambda} = \sqrt{\frac{1}{n\lambda} + 2p \frac{1-p^n}{(1-p)^2 n^2}}$$