ANALYSIS OF VARIANCE OF CUSTOMER BALANCES FOR A FAMILY OF STOCHASTIC SERVICE NETWORKS

Richard L. Patterson
School of Natural Resources, The University of Michigan, Ann Arbor, Michigan 48109-1115, USA

Abstract. The coefficient of variation of counts of customers in nodes of all members of an equivalence class of stochastic service networks is computed for three classes of arrival processes: i) Poisson arrivals of individual customers, ii) Poisson arrivals of fixed and random sized batches of customers, and iii) fixed times of arrivals of constant and random sized batches of customers. Time-variable expressions of means, variances, and coefficients of variation are computed in terms of arrival process parameters, nodal linkages within networks, and residence time distributions of customers in nodes. Coefficients of variation are compared and indices of traffic congestion are computed for member networks within an equivalence class. Use of these indices are an efficient means of rapidly evaluating design parameter changes on performance of networks within an equivalence class.

Key Words. network; stochastic; dynamic; queue

Running title. ANOVA of counts in networks.

Let \( \mathcal{F} = (\mathcal{A}, \mathcal{D}, \ldots, \mathcal{D}_N, \ldots) \) denote a family of equivalence classes of stochastic service networks in which each member network \( S_{\mathcal{D}_N} \) of each class \( \mathcal{D}_N \) contains a finite number \( N \) of linked nodes \( a_1, \ldots, a_N \) \((N=2, 3, \ldots)\). Each node \( a_j \) contains a finite number of servers each of unlimited capacity among which customers are routed according to a fixed probabilistic protocol prior to departing \( a_j \) for \( a_k \). Each member \( S_{\mathcal{D}_N} \) for all \( N \) is characterized by the following conditions:

i) For each initial node \( a_i \) of entry into \( S \) by customers from outside the network all other nodes \( a_j (j \neq i) \) are accessible from \( a_i \).

ii) Every customer entering a nonabsorbing node \( a_j \), following a common probabilistic tour among servers within \( a_j \) is retained in residence at each server a random length of time such that total nodal residence time is a random variable with conditional cdf \( w_j(t) \) whose first two moments are finite and for which \( w_j(t) \) is strictly positive for positive \( t \). Residence times in nodes for customers are iid random variables with cdf \( w_j(t) \). For absorbing nodes \( a_j \) \( w_j(t) \) is defined to be zero for finite \( t \) and one for \( t \) equal to infinity.

iii) Movements of customers among nodes of a member network \( S_{\mathcal{D}_N} \) are governed by a unique Markov renewal (MR) process, discrete state-continuous time, with \( N \)-dimensional conditional residence time distribution function matrix \( W = (w_{jk}(t)) \) where \( w_{jk}(t) \) satisfies conditions given in ii) above. States of (MR) are in one-to-one correspondence with nodes of \( S_{\mathcal{D}_N} \).

Comments. No network contains fewer than two nodes. At least one node of any member \( S_{\mathcal{D}_N} \) is absorbing which serves purposes of counting departing customers and maintaining records of traffic flow through the network. The requirement of finite mean and variance of \( w_{jk}(t) \) rules out absorbing servers within a node which is not an absorbing node. Uniqueness of nodal cdf \( w_{jk}(t) \) requires that all customer conditional residence times in node \( a_j \) are governed by \( w_{jk}(t) \) which, in turn, implies that \( w_{jk}(t) \) is independent of a customers point of entry into \( a_j \) from a prior node and its point of departure from within \( a_j \) to a given destination node \( a_k \).

The Markov renewal process defined by the pair \( P, W \) regulating traffic flow within a network generates a semi-Markov process \( (X(t): t \geq 0) \) describing locations of customers at their most recent changes of nodes. \( \Pr(X(t)=a_j | X(0)=a_i) \)
is the conditional probability that the process (a customer) is in state $j$ (node $a_j$) at time $t$, given that it was in state $i$ (initial node $a_i$) at time $t=0^+$. Conditional probabilities $f_{ij}(t) = Pr(X(t)=a_j/X(0^+)=a_i)$ define an interval transition probability function matrix $F=(f_{ij}(t))$ that is stochastic whenever $P$ is stochastic. Elements of $F$ are computed directly from elements of $P$ and $W$ by conditioning upon the number of changes of state of the process prior to time $t$. Although pure delays may exist at certain servers within a node alternate routes around such servers exist by assumption so that nodal residence time cdf's are positive for positive $t$.

**Distributions of Customer Counts in Nodes**

Distributions of counts of customers in nodes are determined by arrival protocols to the network and processing behavior upon entering an initial node. Effects of processing behavior are studied separately by assuming $c_1, \ldots, c_L$ customers arrive at initial (non-absorbing) nodes $a_1, \ldots, a_L$ at instant $z=0$. Then at time $t>z$ the joint cdf of customers initially in node $a_j$ that are in nodes $a_1, \ldots, a_N$ is multinomial with parameters $c_j, f_{ij}(t-z), \ldots, f_{ijN}(t-z)$ ($f_{ijN}(t-z)>0$ for $k=1, \ldots, N$). When all $L$ initial nodes are accounted for the marginal cdf of customer count in node $a_k$ is that of the sum of $L$ independently distributed binomial random variables with mean $c_k f_{jk}(t-z)$ and variance $c_k f_{jk}(t-z)(1-f_{jk}(t-z))$. The cdf is approximated by a Poisson distribution with parameter $\lambda = \text{mean of the sum of the } L \text{ binomial random variables}$. The approximation error has been bounded (LeCam (1)). Traffic effects of initial conditions eventually die out although ripple effects through the network may be experienced for some period of time until customers finally enter absorbing nodes.

Three basic cases of arrival protocols are considered: i) Poisson arrivals of individual customers to an initial node; ii) Poisson arrivals of batches of customers to an initial node; iii) Arrivals of batches of customers at fixed intervals $0 < t_1 < t_2 < \ldots < t_k < t$. Case ii) contains three sub-cases: ii.1) batch sizes are iid random variables with common mean $m_B$ and variance $\nu_B$ where all customers are processed independently upon entering the initial node; ii.2) case ii.1 applies except all batches are processed as single units in which batch service times are modified according to some criterion; ii.3) case ii.1 applies except the cdf of batch size is a two point distribution, i.e., batches are of sizes $b_0$ and $b_1$ with respective probabilities $p$ and $1-p$.

**Case 1.** A network $S_1 \cdots S_N$ is initially empty and in an interval $(0,t)$ $n$ customers arrive and enter initial node $a_i$. The joint distribution of customer counts in the $N$ nodes of $S$ is:

$$Pr( C_1(t)=c_1, \ldots, C_N(t)=c_N ) =$$

$$\sum_{m_1=0}^{\infty} \cdots \sum_{m_N=0}^{\infty} Pr( C_1(t)=c_1, \ldots, C_N(t)=c_N | m_1, \ldots, m_N )$$

where:

$$m_j(t) = \int_0^t a_i(z) f_{ij}(t-z) dz \quad (j=1, \ldots, N)$$

and:

$$a_i(t)$$

is the intensity of the Poisson process of customer arrivals to initial node $a_i$.

Equation (1) was demonstrated by Kelly (2) using generating functions for a system similar to $S_1 \cdots S_N$. It was also demonstrated by Harrison and Lemoine (3) using a renewal argument for a system of servers all of which communicate. Equation (1) above is demonstrated simply by conditioning a multinomial probability function of counts in the $N$ nodes by total arrivals $n$ in the interval $(0,t)$ and averaging over $n$ using a Poisson cdf. Equation (1) shows counts in nodes to be mutually independent. A calculation shows that no count $C_j(t)$ is independent of the cumulative number of arrivals to initial node $a_i$ in $(0,t)$. It is only for this case that the joint cdf of counts in the $N$ nodes is
obtained, except for a sub-case in which all arriving batches are processed as single units, remaining intact throughout.

Case ii.1. When customers arrive at initial node \( a_i \) in batches of random size where the cdf of batch size has mean \( m_B \) and variance \( v_B \) (Poisson arrivals of batches) the counts of customers in nodes are no longer mutually independent. The cdf of the marginal count \( C_j(t) \) is compound Poisson with mean and variance:

\[
E(C_j(t)) = m_B \int_0^t a_i(z).f_{ij}(t-z)dz \quad (j=1,\ldots,N) \quad (2)
\]

and:

\[
V(C_j(t)) = m_B \int_0^t a_i(z).f_{ij}(t-z)(1-f_{ij}(t-z))dz + (m_B^2 + v_B) \int_0^t a_i(z).f_{ij}(t-z)dz \quad (j=1,\ldots,N) \quad (3)
\]

Assume a cdf of batch size such that \( b_k \) customers arrive in a batch with probability \( p_k \) (\( k=0,1,\ldots,M \)). The arrival process to initial node \( a_i \) of batches of size \( b_k \) only is Poisson with intensity \( p_k.a_i(t) \). The mean and variance of numbers of customers in node \( a_j \) is obtained from equations 2 and 3 above by setting \( m_B=b_k, v_B=0 \), and replacing \( a_i(z) \) by \( p_k.a_i(z) \).

Case ii.2. When all batches are processed as single units throughout every node with residence time cdfs \( w_{ijB}(t) \) the cdf of the number of batches resident in node \( a_j \) at time \( t \) is Poisson with mean:

\[
E(C_j(t)) = m_B \int_0^t a_i(z).f_{ijB}(t-z)dz \quad (j=1,\ldots,N) \quad (4)
\]

and:

\[
V(C_j(t)) = (m_B^2 + v_B) \int_0^t a_i(z).f_{ijB}(t-z)dz \quad (j=1,\ldots,N) \quad (5)
\]

If batches of fixed size \( b_k \) only are considered equations 4 and 5 are modified in a manner analogous to that shown in case ii.1. above.

Case ii.3. When the cdf of batch size is a two point distribution, a subcase of case ii.1. above, assume batches of sizes \( b_0 \) and \( b_1 \) arrive at initial node \( a_i \) with probabilities \( p_0 \) and \( p_1=1-p_0 \), respectively. The mean and variance of batch size are, respectively:

\[
b_0=p_0 \cdot b_0 + b_1 \cdot p_1 = m_B
\]

and:

\[
(b_0^2-m_B^2)p_0 + (b_1-m_B)^2p_1 = v_B
\]

For fixed \( p_0 \) batch sizes \( b_0 \) and \( b_1 \) can be computed which give predetermined mean and variance of batch size. In particular, the mean \( m_B \) can be small and the variance can be large or vice versa. The mean and variance of customer count \( C_j(t) \) in node \( a_j \) at time \( t \) are given by equations 2 and 3. If \( b_0=0 \) case ii.3. is equivalent to case ii.1. where \( a_i(t) \) is replaced by \( a_i(t)p_1, v_B=0, \) and \( m_B=b_1 \).

Case iii. Batches arrive at fixed instants \( t_1,\ldots,t_k \) (\( 0=t_1<\ldots<t_k < t \)) at initial node \( a_i \). For each arriving batch of size \( B_1,\ldots,B_k \) all customers are processed independently throughout. The mean and variance of the marginal cdf of \( C_j(t) \), customer count in node \( a_j \) at time \( t \), are:

\[
E(C_j(t)) = m_B \sum_{r=1}^k f_{ij}(t-t_r) \quad (j=1,\ldots,N) \quad (6)
\]

and:

\[
V(C_j(t)) = m_B^2 \sum_{r=1}^k f_{ij}(t-t_r)(1-f_{ij}(t-t_r)) + \sum_{r=1}^k v_B \cdot f_{ijB}(t-t_r) \quad (j=1,\ldots,N) \quad (7)
\]
The probability that a customer entering initial node $a_i$ at time $t > 0$ is in node $o_j$ at time $t' > t$ is $f_{ij}(t - z)$. The mean amount of time a customer in residence in nonabsorbing node $a_j$ prior to entering an absorbing node is therefore:

$$m_j = \int_0^\infty f_{ij}(t - z) \, dz$$  \hspace{1cm} (12)

The mean proportion of time in residence in nonabsorbing node $a_j$ relative to all nonabsorbing nodes is:

$$\frac{T_j}{\sum_{k=1}^K T_k}$$

where:

$K$ denotes indices of all nonabsorbing nodes of $S\cup N$.

The mean amount of time nonabsorbing node $a_j$ contains $n$ customers is:

$$C_{jn} = \sum_{m=0}^\infty \Pr(C_j(t) = m) \, dt$$  \hspace{1cm} (13)

The mean proportion of time nonabsorbing node $a_j$ contains $n$ customers is consequently:

$$C_{jn} / \sum_{k=1}^K C_{jk}$$

The mean time of residence of a customer in $S\cup N$ exclusive of absorbing nodes is:

$$\sum_{j \neq a}^{T_j} = \sum_{j=1}^J \int_0^\infty f_{ij}(t - z) \, dz$$  \hspace{1cm} (14)

where:

index $j$ ranges over all absorbing nodes in $S$.

Network residence times for customers and counts of customers in a network are related by the formula:

$$\Sigma a_j(z) \cdot f_k(t - z) \, dz = \Sigma_{j=1}^J \int_0^t f_{ij}(t - z) \, dz$$

where:

1) index $j$ ranges over all absorbing nodes;
2) index $j$ ranges over all absorbing nodes.

Indices for Network Design

Equations (12) - (16) may be used to compare different structural configurations of networks within a given class $\mathcal{N}$. Additionally, the squared C.V. and ratios of squared C.V.'s may be used for making the same kinds of comparisons. The above indices may be used to make comparisons at both a nodal and network wide level. Interval transition probability functions need not necessarily be computed exactly although a catalog of $f_{ij}(t)$ functions may be compiled for a wide variety of network configurations. Qualitative behavior of $f_{ij}(t)$'s may be used to make comparisons by assuming monotonicity, convexity or concavity, maximum values, and zeros of the first time derivative of $f_{ij}(t)$ within an interval $(0,T)$. An advantage of reducing a network to membership in an equivalence class of minimal dimension prior to evaluation lies in efficiency of analysis and an ability to compare networks which may contain widely differing numbers of component service systems, each of infinite capacity.

Networks containing $N$ nodes some or all of which have finite capacity behave the same as
Continuous approximations to equations 6 and 7 are:

\[ E(C_j(t)) = \int_0^t m_B \cdot f_{ij}(t-z) dz \quad (j=1, \ldots, N) \quad (8) \]

and:

\[ V(C_j(t)) = \int_0^t \int_0^t f_{ij}(t-z)(1-f_{ij}(t-z)) dz + \int_0^t \frac{1}{m_B} \cdot f_{ij}^2(t-z) dz \quad (j=1, \ldots, N) \quad (9) \]

**Coefficients of Variation of \( C_j(t) \)**

The squared coefficients of variation of \( C_j(t) \) are used to compare variability among cases. Let

\[ I_1 = \int_0^t f_{ij}(t-z) \cdot f_{ij}(t-z) dz \quad \text{and} \quad I_2 = \int_0^t f_{ij}^2(t-z) dz \]

The squared C.V. is shown in Table 1 for all cases. By dividing the squared C.V. in each case except ii.2 by \( I_1^{-1} \), a comparison of the coefficient of variation relative to that of the standard case of individual Poisson arrivals is obtained and shown in Table 2. Adequacy of approximation of models of cases ii.1, ii.3, and iii. by the standard model of Poisson arrivals of single customers is shown in terms of magnitudes of ratios of squared C.V.'s. Effects of variability of times between successive arrivals of batches are estimated by comparing cases ii.1, ii.2, and ii.3. In either Tables 1 or 2, representing completely random arrivals and completely determined times of arrivals. As shown in Table 2 the difference in ratios is approximately \( I_2 - I_1^{-1} \), a fraction always less than unity. The corresponding difference computed from Table 1 is approximately \( I_2 - I_1^{-1} \), a quantity that does not exceed \( I_1^{-1} \) in magnitude, the reciprocal of the mean number of customers in node \( a_j \) arriving from initial node \( d_i \) in \((0,t)\). Processes describing batch arrivals to initial node \( a_j \) for which the variance of independent inter-arrival times is less than that of the Poisson process with intensity \( a_j \) will yield squared C.V.'s of \( C_j(t) \) which differ from either case ii.1 or case iii. by at most \( I_1^{-1} \).

**Variability of customer loading in a node \( a_j \) at time \( t \) expressed as a percentage of the mean may be maximized in some cases.**

As shown in Table 1 for cases ii.1, ii.3, and iii. the squared C.V. is maximized for a fixed \( t \) by setting the first derivative of squared C.V. with respect to \( I_1^{-1} \) equal to zero and solving for \( I_1^{-1} \). For cases ii.1 and ii.3, maximization occurs when:

\[ I_1 = 2I_2 \cdot (1-\frac{m_B}{m_B}) \quad (10) \]

For case iii. maximization occurs when:

\[ I_1 = 2I_2 \cdot (1-\frac{m_B}{m_B}) \quad (11) \]

For given arrival intensity \( a_j \) equations 10 and 11 are integral equations for the unknown interval transition probability function \( f_{ij}(z) \) defined on \((0,z)\).

**Effects of Dependencies in Customer Routing**

Without specification of service times of batches considered as single units a direct comparison of cases ii.1 and ii.2 cannot be made. In the special case where service time of a batch of any size is the same as service time of a customer the difference between squared C.V. for those cases is:

\[ \frac{1}{m_B} \cdot I_1^{-1} \cdot (1-I_2^{-1}) \]

from which bounds on the difference are:

\[ 0 \leq \frac{1}{m_B} \cdot I_1^{-1} \cdot (1-I_2^{-1}) \leq \frac{1}{m_B} \cdot I_1^{-1} \]

Cases ii.1 and ii.2 represent extremes of dependencies among customer routings among nodes of \( S_{DN} \). Squared C.V. for case ii.1 is never smaller than that for ii.2 and it is only when \( f_{ij}(z) \) is identically 1 over the interval \((0,t)\) that they are equal.

**Bounds on \( C_j(t) \)**

Hoeffding (4) developed bounds on probabilities of sums of independently but not identically distributed Bernoulli random variables, appli-
so long as no customer is interrupted during its normal processing sequence through the infinite capacity network. Therefore, if interruptions due to blockages, waits for service, etc., occur only occasionally a finite capacity network may also be evaluated in an approximate sense using the indices developed above. For the majority of customers that encounter no interruptions the above indices apply exactly.

References


TABLE 1 Squared Coefficient of Variation of $C_y(t)$ for Five Cases

Case Number | Squared C.V.
---|---

i. | $1$ 

ii.1. | $\frac{1}{m_B} \cdot \gamma_1^{-1} + (1 + \frac{V_B}{m_B^2}) \cdot I_2 \cdot \gamma_1^{-2}$

ii.2. | $(1 + \frac{V_B}{m_B^2} + 0) \cdot I_2 \cdot \gamma_1^{-2}$

ii.3. | $\frac{1}{m_B} \cdot \gamma_1^{-1} + (1 + \frac{V_B}{m_B^2}) \cdot I_2 \cdot \gamma_1^{-2}$

TABLE 2 Ratio of Squared Coefficients of Variation Using Case i. as a Standard of Comparison

Case Number | Ratio of Squared C.V.
---|---

i. | $1$

ii.1. | $\frac{1}{m_B} \cdot \gamma_1^{-1} + (1 + \frac{V_B}{m_B^2} - \frac{1}{m_B}) \cdot I_2 \cdot \gamma_1^{-2}$

ii.2. | $\frac{1}{m_B} \cdot \gamma_1^{-1} + (1 + \frac{V_B}{m_B^2} - \frac{1}{m_B}) \cdot I_2 \cdot \gamma_1^{-2}$

ii.3. | $\frac{1}{m_B} \cdot \gamma_1^{-1} + (0 + \frac{V_B}{m_B^2} - \frac{1}{m_B}) \cdot I_2 \cdot \gamma_1^{-2}$