1981

Error Analysis of Homogeneous Mean Queue and Response Time Estimators

Jeffrey A. Brumfield

Peter J. Denning

Report Number:
81-393

ERROR ANALYSIS OF HOMOGENEOUS MEAN QUEUE AND RESPONSE TIME ESTIMATORS*

Jeffrey A. Brumfield
Peter J. Denning

Department of Computer Sciences
Purdue University
West Lafayette, Indiana 47907

CSD-TR-393

January 1982

Abstract. Flow balance and homogeneity assumptions are needed to derive operational counterparts of M/M/1 queue length and response time formulas. This paper presents relationships between the assumption errors and the errors in the queue length and response time estimates. A simpler set of assumption error measures is used to derive bounds on the error in the response time estimate.

*This work was supported in part by National Science Foundation grant MCS78-01720 at Purdue University.

ERROR ANALYSIS OF HOMOGENEOUS MEAN QUEUE AND RESPONSE TIME ESTIMATORS

Jeffrey A. Brumfield
Peter J. Denning

It has been shown previously that the formulas for mean queue length ($\bar{n}$) and response time (R),

$$\bar{n} = \frac{U}{1-U} \quad \text{and} \quad R = \frac{S}{1-U} ,$$

are exact for flow balanced behavior sequences of single server queues in which arrivals and services are homogeneous [1,2,3,4]. These formulas are only estimates of the true values when applied to behavior sequences that do not satisfy the assumptions. Our goal in this paper is to show the relationship between errors in the assumptions and errors in the estimates of $\bar{n}$ and R. We will derive expressions for the exact error and for error bounds, the latter being inspired by Kowalk [5]. Table I summarizes the notation [from 2,3].

TABLE I. OPERATIONAL NOTATION

Symbol	Definition	Comments
T		Length of observation period
$n(t)$		Number in system at time t
N		Maximum number observed in system, $0 \le n(t) \le N$
$A(n)$		Number of arrivals observing $n(t)=n$; $A(N)=0$
A	$\sum_{n=0}^{N-1} A(n)$	Total number of arrivals
$C(n)$		Number of completions when $n(t)=n$; $C(0)=0$
C	$\sum_{n=1}^{N} C(n)$	Total number of completions
$T(n)$		Total time during which $n(t)=n$ $(T = \sum_{n=0}^{N} T(n))$
$p_A(n)$	$A(n)/A$	Arriver's queue length distribution
$\bar{n}_A$	$\sum_{n=1}^{N-1} np_A(n)$	Mean queue length observed by arrivers
$p(n)$	$T(n)/T$	Outside observer's queue length distribution
$\bar{n}$	$\sum_{n=1}^{N} np(n)$	Mean outside observer's queue length
W	$\sum_{n=1}^{N} nT(n)$	Accumulated waiting time (note $\bar{n} = W/T$)
R	W/C	Mean response time
X	C/T	Completion rate
U	$1-p(0)$	Utilization
λ	A/T	Arrival rate
α	$A/(T-T(N))$	Restricted arrival rate
$\lambda(n)$	$A(n)/T(n)$	Arrival rate conditioned on queue length
$S(n)$	$T(n)/C(n)$	Service function
S	$\sum_{n=1}^{N} T(n)/C$	Mean service time per completion

1. Measuring the Errors in Assumptions

Errors are measured as weighted aggregates of relative errors. Following standard practice in statistical estimation, we define relative error with respect to the assumed value rather than the actual value. In this section we define the error measures used in the next section to characterize the error in response time estimates.

Our analysis rests on the assumptions summarized in Table II. Flow balance (FB) asserts that as many jobs finish as arrive; this implies that $A=C$ and $A(n) = C(n+1)$ for $0 \le n < N$. The error, e_B, is measured as the difference between arrivals and completions relative to completions.

TABLE II. ASSUMPTIONS AND THEIR ERROR MEASURES

Name	Assertion	Error Measure
Flow Balance (FB)	$A=C$	$e_B = \frac{A-C}{C}$
Homogeneous services (HS)	$S(n)=S, \ 1 \le n \le N$	$e_S = \sum_{n=1}^{N} n \frac{S(n)-S}{S} \frac{C(n)}{C}$
Homogeneous arrivals (HA1)	$\lambda(n) = \begin{cases} a, & 0 \le n < N \\ 0, & n=N \end{cases}$	$e_{A1} = \sum_{n=1}^{N-1} np(n) \frac{\lambda(n)-a}{a}$
Homogeneous arrivals (HA2)	$\lambda(n) = \lambda, \ 0 \le n \le N$	$e_{A2} = \sum_{n=1}^{N} np(n) \frac{\lambda(n)-\lambda}{\lambda}$

Homogeneous services (HS) asserts that the service function, $S(n)$, is a constant, S, for $0 \le n \le N$. The constant is chosen so that the law $S = \sum_{n=1}^{N} S(n) \frac{C(n)}{C}$ is satisfied. The error, e_S, is measured, by analogy with this law, as the weighted relative deviation of the actual from the assumed value. The weight, however, is also proportional to n because errors are more important at longer queue lengths.

Homogeneous arrivals (HA) asserts that the arrival rate, $\lambda(n)$, is a constant for $0 \le n < N$. Version HA1 assumes the constant is a, the arrival rate restricted to intervals when the queue is nonfull -- i.e., when arrivals are actually observed. The constant a is chosen so that the law $\lambda = \sum_{n=0}^{N-1} \lambda(n) p(n)$ is satisfied. Version HA2 assumes the constant is λ, the overall arrival rate, and that $\lambda(N)=\lambda$. Because our definitions ensure that $\lambda(N) = 0$, no finite behavior sequence of a real queueing system can satisfy HA2. HA2 is interesting because it is simpler in form than HA1 and because we can characterize the error in an analysis depending on it.

There is nothing in these assumptions that requires the queue to be a single server. Only for a single-server queue, however, will S be the mean intrinsic service requirement of jobs.

The error e_{A2} can be expressed in terms of e_{A1} as follows:

$$
\begin{aligned}
e_{A2} &= \sum_{n=1}^{N} np(n) \frac{\lambda(n)-\lambda}{\lambda} \\
&= \frac{a}{\lambda} \left[\sum_{n=1}^{N-1} np(n) \frac{\lambda(n)-a}{a} - Np(N) + \sum_{n=1}^{N} np(n) \frac{a-\lambda}{a} \right] \\
&= \frac{a}{\lambda} \left[e_{A1} - Np(N) + \bar{n}\left(1-\frac{\lambda}{a}\right) \right] .
\end{aligned}
$$

On substituting $\frac{\lambda}{a} = 1-p(N)$, this simplifies to

$$
e_{A2} = \frac{e_{A1} - p(N)(N-\bar{n})}{1-p(N)} . \tag{1.1}
$$

Although there are infinitely many flow balanced behavior sequences that satisfy the homogeneity assumptions HA1 and HA2, these assumptions strongly constrain the form of the data that may be observed. Consider a behavior sequence in which $\lambda(n)=a$ for $n=0, \ldots, N-1$ [HA1] and $S(n)=S$ for $n=1, \ldots, N$ [HS]. For such a sequence,

$$C(n) = A(n-1) \qquad \text{[flow balance]}$$

$$T(n) = SC(n) \qquad \text{[definition of } S(n)\text{]}$$

$$\begin{aligned} A(n) &= aT(n) \qquad \text{[definition of } \lambda(n)\text{]} \\ &= aSA(n-1) . \end{aligned}$$

It follows that

$$A(n) = A(0)(aS)^n \qquad n=0, \ldots, N-1$$

$$C(n) = A(0)(aS)^{n-1} \qquad n=0, \ldots, N-1$$

$$T(n) = A(0)S(aS)^{n-1} \qquad n=0, \ldots, N \tag{1.2}$$

In other words, a behavior sequence satisfies HA1 and HS if and only if the values $\{A(n), C(n), T(n)\}$ are geometric series.†

† Given a table of values, one can construct a behavior sequence by tiers from level N downward. (Assume $n(0)=n(T)=0$.) The i^{th} tier consists of a sequence whose data are $\{A(n), C(n), T(n) \mid n \geq N-i+1\}$. This tier includes arrival and completion events acting as "stubs" connected to level $n=N-i$. The $i-1^{\text{st}}$ tier is expanded as follows: the $C(N-i)$ completions are distributed arbitrarily among the stubs, with one completion at the rightmost stub; arrivals are also distributed so that each completion matches a previous arrival, with one arrival at the leftmost stub; finally, nonzero intervals are inserted between arrivals and completions to consume all of $T(N-i)$. This process starts with $i=0$ and a single, empty stub. It terminates with $i=N$ and a behavior sequence whose data match those given.

2. Estimates of Response Time

We will extend derivations from [2,3] for R and $\bar{n}_A$ to show the propagation of error from each homogeneity assumption. The error from the flow balance assumption can be tracked by noting $C(n) = A(n-1) + d(n)$, where

$$d(n) = \begin{cases} -1, & n(0) < n \le n(T) \quad [A>C] \\ +1, & n(T) < n \le n(0) \quad [A<C] \\ 0, & \text{otherwise} \end{cases}$$

If flow is balanced, $d(n) = 0$ for all n. We will need the value of the expression $\sum_{n=1}^{N} n \frac{d(n)}{C}$ in the following analysis. Using the identities $\sum_{k=1}^{K} k = K(K+1)/2$ and $n(T) - n(0) = A - C$, it is not difficult to show

$$\sum_{n=1}^{N} n \frac{d(n)}{C} = -e_B \frac{n(0)+n(T)+1}{2} \tag{2.1}$$

where e_B is the flow balance error (Table II).

2.1 Homogeneous Services

The homogeneous services (HS) assumption asserts that $S(n) = S$ for $n = 1,\ldots,N$. This assumption leads to a simple relation between the response time, R, and the mean queue length seen by arrivers, $\bar{n}_A$. The response time satisfies the identity

$$R = \sum_{n=1}^{N} nS(n)\frac{C(n)}{C} .$$

By adding and subtracting terms, this can be expanded to

$$R = S\left[\sum_{n=1}^{N} n\frac{A(n-1)}{A} + \frac{A-C}{C}\sum_{n=1}^{N} n\frac{A(n-1)}{A} + \sum_{n=1}^{N} n\frac{d(n)}{C} + \sum_{n=1}^{N} n\frac{S(n)-S}{S}\frac{C(n)}{C}\right] .$$

The first two summations evaluate to $(\bar{n}_A+1)(1+e_B)$. The value of the third sum is given by Equation 2.1. The fourth sum is e_S, the error in the HS assumption (Table II). Collecting these terms we obtain

$$R = S\ (\bar{n}_A+1) + \left\{ S\left[e_S + (\ \bar{n}_A + 1 - \frac{n(0)+n(T)+1}{2})\ e_B \right]\right\} \quad . \tag{2.2}$$

The terms in braces are the error terms, which vanish for behavior sequences that satisfy assumptions FB and HS. That is, $R = S\ (\bar{n}_A + 1)$ if flow is balanced and services are homogeneous.

2.2 Homogeneous Arrivals

The homogeneous arrival assumptions (HA1 and HA2) lead to simple relations between the arriver's mean queue length, $\bar{n}_A$, and the outside observers mean queue length, $\bar{n}$. Appling the law $p_A(n) = p(n)\lambda(n)/\lambda$ to the definition of $\bar{n}_A$ gives

$$\bar{n}_A = \sum_{n=1}^{N-1} np(n)\frac{\lambda(n)}{n} \quad . \tag{2.3}$$

By adding and subtracting terms, this can be expanded to

$$\bar{n}_A = \frac{a}{\lambda}\sum_{n=1}^{N-1} np(n) + \frac{a}{\lambda}\sum_{n=1}^{N-1} np(n)\frac{\lambda(n)-a}{a} \quad .$$

The first sum evaluates to $\frac{a}{\lambda}(\bar{n}-Np(N))$. The second sum is $\frac{a}{\lambda}e_{A1}$, where e_{A1} is the

HA1 error measure (Table II). Note that $\lambda / \alpha = 1-p(N)$. Collecting these terms, we obtain

$$\bar{n}_A = \frac{\bar{n} - Np(N)}{1-p(N)} + \left\{ \frac{e_{A1}}{1-p(N)} \right\} . \qquad (2.4)$$

The error term, in braces, vanishes when HA1 is satisfied.

The upper limit of summation in Equation 2.3 can be extended to N since $\lambda(N) = 0$. Equation 2.3 can then be expanded to

$$\bar{n}_A = \sum_{n=1}^{N} np(n) + \sum_{n=1}^{N} np(n) \frac{\lambda(n)-\lambda}{\lambda} .$$

The first sum is the definition of $\bar{n}$; the second is the definition of e_{A2} (Table II). Thus,

$$\bar{n}_A = \bar{n} + e_{A2} \qquad (2.5)$$

Equation 2.5 shows that the HA2 assumption allows us to approximate $\bar{n}_A$ by $\bar{n}$. But, because HA2 can never be satisfied exactly, $\bar{n}_A$ would never be exactly equal to $\bar{n}$.

2.3 Error Formulas

Table III gives two formulas for estimating mean response time corresponding to the two arrival assumptions, together with relative errors. For a formula $R = expr$, the relative error is $\frac{R - expr}{expr}$. The bottom half of the table displays the simpler formulas that hold when flow is balanced.

TABLE III. ERRORS OF RESPONSE TIME ESTIMATES

Estimator	Relative Error
[HA1] $R = \frac{S}{1-U-p(N)}(1-(N+1)p(N))$	$\frac{1-p(N)}{1-(N+1)p(N)}\left[\frac{e_{A1}}{1-p(N)} + e_S + (\bar{n}_A+1-\frac{n(0)+n(T)+1}{2})\, e_B\right]$
[HA2] $R = \frac{S}{1-U}$	$e_{A2} + e_S + (\bar{n}_A + 1 - \frac{n(0)+n(T)+1}{2})\, e_B$
For flow balance ($e_B = 0$):	
[HA1] $R = \frac{S}{1-U-p(N)}(1-(N+1)p(N))$	$\frac{1-p(N)}{1-(N+1)p(N)}\left[\frac{e_{A1}}{1-p(N)} + e_S\right]$
[HA2] $R = \frac{S}{1-U}$	$e_{A2} + e_S$

The error for the response time formula relying on HA1 is derived as follows. If $R = S(\bar{n}_A+1) + \varepsilon_1$ (as in Equation 2.2) and $\bar{n}_A = \frac{\bar{n} - Np(N)}{1-p(N)} + \varepsilon_2$ (as in Equation 2.4), then

$$R = \frac{S\bar{n}}{1-p(N)} + \frac{S\,(1-(N+1)p(N))}{1-p(N)} + \{\, \varepsilon_1 + S\,\varepsilon_2 \,\} \quad .$$

Now, $S\bar{n} = S\lambda R = UR$ by Little's law and the utilization law. Making this substitution and solving,

$$R = \frac{S}{1-U-p(N)}(1-(N+1)p(N)) + \left\{ \frac{1-p(N)}{1-U-p(N)}(\varepsilon_1 + S\varepsilon_2) \right\} \quad . \qquad (2.6)$$

The term in braces is the absolute error. On dividing the absolute error by the formula for R, we obtain the relative error

$$\frac{1-p(N)}{1-(N+1)p(N)} \left(\frac{\varepsilon_1}{S} + \varepsilon_2 \right) \quad .$$

The formula in Table III results when ε_1 and ε_2 are substituted from Equations 2.2 and 2.4, respectively.

The error for the response time formula relying on HA2 is derived as follows. If $R = S(\bar{n}_A + 1) + \varepsilon_1$ (as in Equation 2.2) and $\bar{n}_A = \bar{n} + \varepsilon_3$ (as in Equation 2.5), then

$$R = S\bar{n} + S + \{ \varepsilon_1 + S\varepsilon_3 \} \quad .$$

Using $S\bar{n} = UR$ and solving for R,

$$R = \frac{S}{1-U} + \frac{\varepsilon_1 + S\,\varepsilon_3}{1-U} \quad . \qquad (2.7)$$

Dividing the error term by the formula for R, we obtain the relative error

$$\frac{\varepsilon_1}{S} + \varepsilon_3 \quad .$$

Substituting the errors from Equations 2.2 and 2.5, respectively, this reduces to the form shown in Table III.

Formulas for $\bar{n}$ are obtained from Little's law, $\bar{n} = \lambda R$. The relative errors in these estimates are the same as in the estimates for R.

2.4 Discussion

The formulas in Table III express the relative errors in the response time estimates in terms of the assumption errors and the quantities N, $p(N)$, $n(0)$, $n(T)$, and $\bar{n}_A$. Although the definitions insure that $p(N)$ is never identically 0, $p(N)$ may be negligible for many behavior sequences. As $p(N)$ approaches 0, assumptions HA1 and HA2 become identical. The two estimates and their errors also become identical.

The denominators of both the HA1 estimate and its absolute error (from Equation 2.6) are zero when $U = 1-p(N)$. The same is true of the HA2 estimate and its absolute error (from Equation 2.7) when $U = 1$. It has been shown [3] that when $U = 1-p(N)$ and assumptions HA1, HS, and FB are satisfied, then $p(n) = 1/(N+1)$ for $n = 0,\ldots,N$ and $R = SN/2U$.† On the other hand, when $U = 1$ and assumptions HA2, HS, and FB are satisfied, the $p(n)$ and R are undefined.

The behavior of the relative errors in these special cases is more subtile than the formulas in Table III suggest. Consider the estimate for R under assumptions FB and HA2, for which the relative error, $e_{A2}+e_S$, does not explicitly depend on U. In this case, the formula

$$R = \frac{S}{1-U} + \frac{S}{1-U}(e_{A2} + e_S)$$

is exact for all $U \neq 1$. As U approaches 1, $S/(1-U)$ approaches $+\infty$. But, because the

† Equation 1.2 implies $p(n) = (1-aS)(aS)^n/(1-(aS)^{N+1})$. In this case, the limit of $p(n)$ is $1/(N+1)$ and $U = 1-p(N)$.

response time is always bounded ($0 \leq R \leq NT/C$), the quantity $e_{A2}+e_S$ must approach $-\infty$. In other words, the assumption errors are likely to be unbounded as U approaches 1, causing the relative error in the estimates to be unbounded.

Because the HA2 assumption requires that $\lambda(N) \neq 0$, the HA2 estimate contains an irremovable error. The only observable constant arrival function is $\lambda(n)=a$ for $0 \leq n < N$, in which case Equation 1.1 shows that $e_{A2} = -(N-\bar{n})p(n)/(1-p(N))$. This would be the relative error in the HA2 estimate for a flow balanced behavior sequence satisfying HA1 and HS.

3. Response Time Error Bounds

The error measures in Table II allow an exact formulation of the errors in the response time estimates. Unfortunately, the computation of these error measures requires detailed knowledge of the arrival function, the service function, and the overall queueing distribution. In this section, we use a simpler set of error measures to derive bounds on the errors in the response time estimates.

3.1 Alternate Error Measures

Table IV shows a different set of error measures for the four assumptions. The quantities E_S, E_{A1}, and E_{A2} are the maximum relative errors between the actual value of the function (service or arrival) and the assumed constant value. Unlike the error measures of the previous section, those in Table IV do not reflect the individual errors at each queue length; they can only lead to bounds on the errors in the response time estimates.

Table V shows the relations between the error measures of the previous section and the error measures of Table IV. These relations are derived by taking magnitudes of the measures in Table II, applying the triangle inequality, and expressing the results in terms of the Table IV measures. These relations then give bounds on the magnitudes of the error formulas in Table III. Table VI summarizes these bounds. As before, relative error is defined with respect to the estimated value.

TABLE IV. ALTERNATE ERROR MEASURES FOR ASSUMPTIONS

Assumption	Error Measure
Flow Balance (FB)	$E_B = \left\lvert \frac{A-C}{C} \right\rvert$
Homogeneous Services (HS)	$E_S = \max\limits_{1 \le n \le N} \left\lvert \frac{S(n)-S}{S} \right\rvert$
Homogeneous Arrivals (HA1)	$E_{A1} = \max\limits_{0 \le n < N} \left\lvert \frac{\lambda(n)-\alpha}{\alpha} \right\rvert$
Homogeneous Arrivals (HA2)	$E_{A2} = \max\limits_{0 \le n < N} \left\lvert \frac{\lambda(n)-\lambda}{\lambda} \right\rvert$

The formulas in Table VI can also be derived by a method similar to Kowalk's [5], which annotates each step of a derivation with a bound on the error present at that step. Because the algebra is similar to that appearing in the derivations for exact errors (Table III), we did not use Kowalk's approach in writing this paper. We have, however, verified that his approach leads to the same results.

TABLE V. RELATIONS BETWEEN ERROR MEASURES

$\lvert e_B \rvert = E_B$
$\lvert e_S \rvert \le E_S(\bar{n}_A+1) + E_B E_S\,(\bar{n}_A+1+\frac{n(0)+n(T)+1}{2})$
$\lvert e_{A1} \rvert \le E_{A1}\,(\bar{n}-Np(N))$
$\lvert e_{A2} \rvert \le E_{A2}\,(\bar{n}-Np(N)) + Np(N)$

TABLE VI. RESPONSE TIME ERROR BOUNDS

Estimator	Relative Error Bound
[HA1] $R = \frac{S}{1-U-p(N)}(1-(N+1)p(N))$	$\left\lvert \frac{1-p(N)}{1-(N+1)p(N)} \right\rvert \left[E_{A1}\frac{\bar{n}-Np(N)}{1-p(N)} + E_S(\bar{n}_A+1) + E_B(1+E_S)(\bar{n}_A+1+\frac{n(0)+n(T)+1}{2}) \right]$
[HA2] $R = \frac{S}{1-U}$	$E_{A2}(\bar{n}-Np(N)) + Np(N) + E_S(\bar{n}_A+1) + E_B(1+E_S)(\bar{n}_A+1+\frac{n(0)+n(T)+1}{2})$

3.2 An Empirical Study

To assess the differences between the exact errors and the error bounds, we conducted a small simulation study of behavior sequences of M/M/1 ensembles.

Ten flow balanced behavior sequences consisting of 2000 arrivals/completions were randomly generated. Interarrival times were drawn from an exponential distribution having mean 1.0; service times were drawn from an exponential distribution having means ranging from 0.1 to 0.95. This produced behavior sequences with a wide range of utilizations.

For each behavior sequence, all quantities needed to compute the response time estimates and their error bounds were measured. For each estimator, the actual relative error and its bound were computed. Table VII summarizes the results; each row represents a single behavior sequence. These data confirm the tendency for the assumption errors and result errors to grow in magnitude as U approaches 1.

Because the values of $p(N)$ were always less than 0.001 in our experiments, the two estimators were similar. (We did not study behavior sequences in which $p(N)$ is significant, for example, the bottleneck of a closed system.) While the actual errors in the estimates were mostly under 5% and always less than 37%, the error bounds were 20-100 (or more) times the magnitudes of the actual errors. Error bounds denoting 100% confidence intervals around the estimated value of R may be too loose for practical application of error analysis.

TABLE VII. SIMULATION RESULTS

	Assumption Errors						HA1 Estimate		HA2 Estimate	
U	e_S	e_{A1}	e_{A2}	E_S	E_{A1}	E_{A2}	error	bound	error	bound
.101	.003	.012	.011	.115	1.038	1.038	.015	.247	.015	.247
.206	.020	.009	.007	.513	.513	.513	.029	.791	.027	.792
.301	.003	.014	.013	.208	.608	.608	.017	.606	.016	.607
.388	.024	-.029	-.030	.283	.700	.700	-.006	.894	-.007	.894
.517	.009	-.012	-.015	.519	.746	.740	-.003	1.804	-.006	1.661
.601	-.005	-.044	-.045	.717	.264	.264	-.050	2.091	-.051	2.090
.690	.157	.036	.032	.694	.623	.623	.194	4.377	.189	4.301
.835	-.012	.060	.035	1.010	16.627	16.642	.050	105.426	.027	102.805
.883	-.238	.133	.112	.583	.741	.742	-.106	9.634	-.121	9.460
.978	-.150	-.215	-.224	.723	1.120	1.120	-.366	52.461	-.370	52.007

4. Conclusions

This paper illustrates an operational sensitivity analysis – the evaluation of the error between the value of the response time estimated from a formula such as $R = S/(1-U)$ (using observed values of S and U) and the actual value of response time observed in a given behavior sequence. The error analysis verifies that response time estimates tend to be less reliable for heavily-utilized systems than for lightly-loaded ones. The error analysis shows that the value of $p(N)$ is unimportant in a lightly-loaded system but may be important in a heavily-utilized one.

Several additional research directions are apparent. One is to extend this analysis for the response time formula to M/G/1 systems. Kowalk has worked out error bounds for this case [5] and we have sketched out an exact analysis [4].

Another direction is to generalize the confidence interval analysis to deal with this form of question: "Suppose $E(e_1, \ldots, e_k)$ is the result error when the errors of the k assumptions are $e_1, \ldots, e_k$, and suppose a fraction p of behavior sequences have assumption error $\leq e_i$, for all i; what is the largest fraction q of behavior sequences having result error $\leq E(e_1, \ldots, e_k)$?" If the k assumptions are independent, $q = p^k$. In general, however, errors of different assumptions may be correlated. This question requires a careful analysis. It could circumvent the difficulty, noted in the empirical study, that the $p = 100\%$ error bounds are too loose, by giving much tighter bounds for large subsets of behavior sequences.

A final direction is the error analysis of performance metrics for closed systems. In closed queueing networks, for example, errors in estimates of mean service times are attenuated as they propagate to standard metrics [6,7]. It remains to investigate whether more primitive assumptions, such as network homogeneity [3,8], are similarly attenuated.

Acknowledgements

We are grateful to Jeffrey Buzen for his suggestions on an early draft of this paper and to Wolfgang Kowalk for inspirations on error bounds. NSF supported part of this research at Purdue University through grant MCS78-01729.

References

1. P. J. Denning, "Operational estimates of mean queue length and response time of a single server queue," Purdue University, Computer Sciences Department, Memorandum PJD78.5 (August 1978).

2. J. P. Buzen and P. J. Denning, "Measuring and calculating queue length distributions," IEEE *Computer* (April 1980), 33-44.

3. J. P. Buzen and P. J. Denning, "Operational treatment of queue distributions and mean value analysis," *Computer Performance 1*, 1 (June 1981), 6-15.

4. J. Brumfield and P. Denning, "Operational analysis of queues with general service times," Purdue University, Computer Sciences Department, W. Lafayette, IN 47907, Report CSD-TR-357 (January 1981).

5. W. Kowalk, "Extensions of operational analysis," in *Messung, Modellierung und Bewertung von Rechensystemen*, Informatik-Fachberichte No. 41, Springer (February 1981).

6. A. C. Williams and R. A. Bhandiwad, "A generating function approach to queueing network analysis of multiprogrammed computers," *Networks*, Vol. 6, No. 1, 1976, pp. 1-22.

7. R. Suri, "Robustness of analytical formulae for performance prediction in certain nonclassical queueing networks," Technical Report No. 674, Division of Applied Sciences, Harvard University, August 1980.

8. P. J. Denning and J. P. Buzen, "The operational analysis of queueing network models," *Computing Surveys*, Vol. 10, No. 3, September 1978, pp. 225-261.