1981 Winter Simulation Conference Proceedings
T.I. Oren, C.M. Delfosse, C.M. Shub (Eds.)

METHODS FOR STRATIFIED ESTIMATION IN NONREGENERATIVE SIMULATIONS

Andrew F. Seila
University of Georgia

John S. Carson
Georgia Institute of Technology

ABSTRACT

We discuss a general concept of natural stratification in simulation, and present an approach to estimating mean performance measures in stationary simulations which possess a natural stratification.

1. INTRODUCTION

Frequently the data produced by a simulation possesses a natural stratification, such that the simulator wishes to know not only the overall mean performance measure, but also the mean performance measure within each stratum. For example, in a priority queueing system, the investigator might wish to know the mean waiting time for each priority class and the overall mean waiting time.

Suppose that we have a stationary simulation with $s \geq 1$ distinct strata. Let

- r_j be the stationary mean of observations in stratum j;
- Π_j be the stationary proportion of observations in stratum j; and
- r be the overall stationary mean.

It is generally true that

$$(1.1) \quad r = \sum_{j=1}^{s} \Pi_j r_j ,$$

and the proportions $\Pi_1, \Pi_2, \ldots, \Pi_s$ can be computed in most simulations. Let $\hat{r}_j$ be an estimator of r_j, j=1,2,...,s. A stratified estimator, $\hat{r}$, of r is defined to be

$$(1.2) \quad \hat{r} = \sum_{j=1}^{s} \Pi_j \hat{r}_j .$$

2. STRATIFIED ESTIMATORS FOR REGENERATIVE SIMULATIONS

If, for the system being simulated, there are points in time such that the system "begins anew" in a probabilistic sense at each of these times, then the simulation is said to be regenerative. For example, in a queueing system in which all interarrival and waiting times are independent, when a customer arrives to find the system empty and idle, a regeneration occurs. See Fishman (1978) and references therein for further discussion of regenerative simulations. The important characteristic of a regenerative simulation is that observations can be grouped into groups of random size, called cycles, so that these cycles of data are independent and identically distributed. A cycle of data consists of the observations between two regeneration times.

Suppose that the system has $s > 1$ strata (e.g., classes of customers), and (1.1) specifies the relationship between the stationary mean waiting time r_i for customers in stratum i and the overall mean waiting time, r. We assume that the values of $\Pi_1, \Pi_2, \ldots, \Pi_s$ are known. Let

- X_{ij} be the sum of observations from stratum j during cycle i; and
- N_{ij} be the number of observations from stratum j during cycle i.

The regenerative estimator for r is

$$(2.1) \quad \hat{r}_j = \frac{\bar{X}_j}{\bar{N}_j} ,$$

where

$$(2.2) \quad \bar{X}_j = \frac{1}{n} \sum_{i=1}^{n} X_{ij}$$

and

$$(2.3) \quad \bar{N}_j = \frac{1}{n} \sum_{i=1}^{n} N_{ij} .$$

A $100(1-\alpha)$ - percent confidence interval for r_j is given by $\hat{r}_j \pm h_j(\alpha)$, where

$$(2.4)\quad h_j(\alpha) = Z_{\alpha/2} \frac{\hat{\sigma}_j}{\bar{N}_j\sqrt{n}} ;$$

$$(2.5)\quad \hat{\sigma}_j^2 = n^{-1}[s^2_{Xj} - 2\hat{r}_j s_{X_jN_j} + \hat{r}_j^2 s_N^2] ;$$

$$s^2_{X_j} = \sum_{j=1}^{n} X^2_{ij};\ s^2_{N_j} = \sum_{i=1}^{n} N^2_{ij};$$

$$s_{X_jN_j} = \sum_{j=1}^{n} X_{ij}N_{ij};\ \text{and}$$

$Z_{\alpha/2}$ is the $100(1-\alpha/2)^{th}$ percentile of the unit normal distribution.

The stratified regenerative estimator is given by

$$(2.6)\quad \tilde{r} = \sum_{j=1}^{s} \Pi_j \hat{r}_j .$$

This estimator is strongly consistent and asymptotically normal. An approximate large sample confidence interval for r is given by Seila (1980).

$$(2.7)\quad (\tilde{r} - Z_{\alpha/2}\frac{\tilde{\sigma}}{\sqrt{n}},\ \tilde{r} + Z_{\alpha/2}\frac{\tilde{\sigma}}{\sqrt{n}}),$$

$$(2.8)\quad \tilde{\sigma}^2 = \sum_{j=1}^{s}\sum_{k=1}^{s} \frac{\Pi_j\Pi_k}{\bar{N}_j\bar{N}_k}\hat{\sigma}_{jk} ,$$

$$(2.9)\quad \hat{\sigma}_{jk} = \frac{1}{n}\sum_{i=1}^{n} (X_{ij} - \hat{r}_j N_{ij})(X_{ik} - \hat{r}_k N_{ik})$$

and $Z_{\alpha/2}$ is the $100(1-\alpha/2)^{th}$ percentile of the unit normal distribution.

3. STRATIFIED ESTIMATION IN NONREGENERATIVE SIMULATIONS USING BATCH MEANS

Frequently, the regenerative approach cannot be used, either because the system is not regenerative, or because the system is regenerative but regenerations occur so infrequently that excessive run lengths are required in order to get a sufficient number of cycles of data. A number of alternative approaches have been developed for estimating the stationary mean performance measure, including time series methods, independent replications and batch means.

Suppose that the simulation is stationary but not regenerative and that each observation comes from one of s mutually exclusive strata. The simulation run consists of n consecutive batches of m waiting times each, taken while the system is in steady-state. Let:

X_{ij} be the sum of observations from stratum j in batch i,

N_{ij} be the number of observations from stratum j in batch i.

Since the batch size is fixed at m, we have

$$\sum_{j=1}^{s} N_{ij} = m$$

for each batch i=1,2,...,m. However, the N_{ij}'s are random variables since the number of observations (customers, in this case) from each stratum is not fixed for each batch. We will assume that batches are large enough that the vectors $\{(X_{i1}, N_{i1}), (X_{i2}, N_{i2}), \ldots, (X_{is}, N_{is})\}$ are approximately uncorrelated between batches.

If we treat each batch as a cycle, we can apply the stratified regenerative method. The estimators for mean waiting time within each stratum are given by (2.1) through (2.3), and the confidence intervals for mean waiting time within each stratum are given by (2.4) through (2.5). The point estimator and confidence interval for the overall mean is given by (2.6) through (2.9).

Sometimes the strata in the model are defined so that observations are independent with respect to the strata they represent. This is the case for example, in a queueing system with s priority classes, when each customer's priority class is selected by making an independent random drawing from the discrete distribution $\{\Pi_1, \Pi_2, \ldots, \Pi_s\}$. Frequently, in queueing, service or production system models, customer or job attributes are determined upon arrival by an independent random drawing from an appropriate probability distribution. If these attributes are the basis for defining strata, then the observations are independent with respect to their strata.

If the observations are independent with respect to their strata, and Π_j is the probability that an observation falls in stratum j, then the vector of strata occupancy $\{N_{i1}, N_{i2}, \ldots, N_{is}\}$ has a multinomial distribution with parameters m and $\Pi_1, \Pi_2, \ldots, \Pi_s$. In this special case, each N_{ij} has a binomial distribution with parameters m and Π_j, and

$$E(N_{ij}) = m\Pi_j,$$

$$E(N_{ij}^2) = m\Pi_j\,(1 + (m-1)\Pi_j),$$

and

$$E(N_{ij}N_{ik}) = m(m-1)\Pi_j\Pi_k \text{ if } j \neq k$$

$$= m(m-1)\Pi_j^2 + m\Pi_j \text{ if } j = k.$$

When these expressions are used for analogous sample quantities in (2.4)-(2.9), we obtain estimators, $\tilde{r}_j$, j=1,2,...,s, and $\tilde{r}$ which are unbiased but have larger variance than the ratio estimators.

REFERENCES

Fishman, G. S. (1978), Principles of Discrete Event Simulation, John Wiley, New York.

Seila, A. F. (1980), A Stratified Regenerative Method for Estimating Mean Waiting Times in Queues with Multiple Renewal Input Streams, Technical

Report, Department of Quantitative Business
Analysis, University of Georgia, Athens,
Georgia.