COPPEAD/UFRJ

RELATÓRIO COPPEAD Nº 276

# ON THE USE OF DESCRIPTIVE SAMPLING TO ESTIMATE PROBABILITY DISTRIBUTIONS: AN APPLICATION TO A RISK ANALYSIS PROBLEM

Eduardo Saliby*

Maio 1993

* Professor da área de Métodos Quantitativos da COPPEAD/UFRJ.

**ABSTRACT**

Descriptive Sampling (DS) is a new sampling approach in Monte Carlo simulation, based on a deterministic selection of the input sample values and their random permutation. Although many empirical tests were already carried out to confirm the usefulness of DS, most of the attention has been given to the estimation of the mean and the standard deviation of the responses under study. A more complete validation of DS is possible by estimating the cumulative distribution function of the response variables. An application of this methodology and the results showing the benefits from using DS are reported.

## INTRODUCTION

Recently proposed (Saliby,1990), Descriptive Sampling (DS) is a new approach to sampling in Monte Carlo simulation. Apart from the benefits in terms of variance reduction it produces, Descriptive Sampling represents a considerable change to basic simulation ideas. This follows because the random selection of the sample values is replaced by a fully deterministic selection of such values. The only random element in a descriptive sample comes from the random permutation of the input values. Up to now, the empirical validation of DS has been concentrated on a few parameters of the simulation response variables, mainly the mean and standard deviation (Saliby, 1989 and 1990). However, given that DS was proposed as a general approach to any Monte Carlo application, it is important to show that its' benefits extend to the whole response distribution, not only a few parameters. This is carried out focusing the attention to the empirical cumulative distribution function (EDF) of the response variables and estimating a whole set of percentiles using both simple random sampling (SRS) and descriptive sampling (DS). As will be shown, the use of DS improves the quality of simulation estimates for the entire response variable distribution scope.

## DESCRIPTIVE SAMPLING

This section presents a basic introduction to the idea of DS. A more extensive description is given in papers by Saliby(1990) and Saliby and Paul(1993). Descriptive sampling was proposed in order to avoid the set variability in simulation studies. When using the standard Simple Random Sampling (SRS) or Monte Carlo approach, two kinds of variation are present in a randomly generated sample - one related to the set of values and the other to their sequence. But, of these two kinds of variability, only the sequence variability is really inevitable, while the set variability is in fact unnecessary. Symbolically, the two sampling methods can be represented as

Simple random sampling = random set x random sequence, whilst

Descriptive sampling = deterministic set x random sequence.

The only additional requirement to use DS instead of SRS is to know, in advance, the input sample size, which, as stressed in Saliby and Paul(1993), must be related to a full simulation run. Once the sample size is known, at least approximately, the set values are defined using the inverse transform method, also used for SRS, and given by

$$XD_i = F^{-1}[(i-0.5)/n] \quad , \quad i=1,..,n \; ,$$

where

$$F^{-1}(R) \quad , \; R \in (0,1)$$

is the inverse transform for the particular input distribution.

When the inverse distribution is not available, numerical or functional approximations can be used. One such case is the LogNormal distribution, for which the Ramberg and Schmeiser (1972) approximation for the parent Normal distribution works fairly well.

Completing the DS generation process, each set of input values is used in a random sequence in each simulation run. Now, unlike with SRS, set values are the same for all replicated runs in a simulation experiment. This random shuffling process is easily accomplished by sampling without replacement the descriptive set of values (Saliby,1990).

Apart from a higher memory requirement than SRS, DS generation is usually as fast as SRS generation. It can even be faster than SRS when multiple runs are carried out and the random generation of the input values is too time-consuming. Like this, other issues concerning DS generation deserve attention, but they are outside the scope of the present work. Here, our attention is directed to the quality of the results that DS produces.

**EXTENDING THE EMPIRICAL VALIDATION**

Assuming for simplicity, but without loss of generality, a terminating simulation with a single response Z, represented as a generic function of the input variables $(X_1,...,X_k)$, so that

$$Z = F(X_1,...,X_k),$$

a simulation run is defined by $N\geq 1$ trials, each trial producing one observed value for the response variable. Based on the sample of response values

$$Z_j, \ j=1,..,N \ ,$$

we can estimate parameters related to the response variable distribution. In this case, a generic parameter $\Theta_j$ may be estimated as

$$Y_j = H(Z_1,...,Z_N),$$

for which unbiasedness, so that $E[Y_j]=\Theta_j$ as $N\to\infty$, and consistency, so that $Var[Y_j]\to 0$ as $N\to\infty$, are two asymptotic properties required.

Up to the present, the empirical validation of DS has been concentrated on the mean and standard deviation estimates of the response variables. However, it is the aim of the present work to show that the benefits from using DS extend to the whole response variable distribution. This is achieved by comparing the performance of both sampling methods in estimating the response variable cumulative distribution function.

Given a sample of response values $Z_j$, $j=1,..,N$, the cumulative distribution function (CDF) is estimated by the empirical cumulative distribution function (EDF), defined by

$$EDF(z) = \#(Z_1,...,Z_N \leq z)/N, \quad z \in R ,$$

where $\#(Z_1,...,Z_N \leq z)$ is the number of cases for which $Z_j <= z$, $j=1,..,N$. As such, once z is fixed, EDF(z) can be seen as a random variable.

Let EDFR(z) be the empirical cumulative distribution function for the response variable Z when using SRS, and let EDFD(z) be the corresponding empirical cumulative distribution function when using DS. To give a full empirical validation for the DS proposal, we want to show that

$$E[EDFR(z)] \simeq E[EDFD(z)] ,$$

and that

$$Var[EDFD(z)] \leq Var[EDFR(z)] ,$$

for any $z \in R$. However, instead of fixing the z values and letting EDF(z) vary, we fixed EDF(z) to a set of $\alpha_t$, t=1,..k values of interest, thus letting the corresponding fractile $Z_\alpha$ vary between simulation runs. Defining an experiment by $M \geq 1$ replicated runs, we can estimate $E[Z_\alpha]$ and $Var[Z_\alpha]$ for both sampling methods and compare them. If $ZR_\alpha$ and $ZD_\alpha$ are, respectively, estimated fractiles using SRS and DS, we want to show that

$$E[ZR_\alpha] \simeq E[ZD_\alpha] ,$$

and that

$$Var[ZR_\alpha] \leq Var[ZD_\alpha] ,$$

for the whole set of $\alpha_t$, t=1,..k values of interest.

**METHODOLOGY**

A common but simple situation in a risk analysis problem is when two independent input random variables, like

$X$ = total market demand,

and

$Y$ = market share,

produce a response

$R = XY$,

that can be viewed as the revenue. For instance, the net present value of a project (NPV), a standard response in a risk analysis problem, is usually a linear combination of such products.

Let X and Y be the two input random variables and R the response random variable when using simple random sampling, and let XD, YD and RD be the corresponding random variables when using descriptive sampling. No matter the sampling method in use (SRS or DS), due to the independence of the two input variables, it follows that

$$E[R] = E[X].E[Y] \ ,$$
$$E[RD] = E[XD].E[YD] \ ,$$
$$Var[R]=[E(X)]^2.Var(Y)+[E(Y)]^2.Var(X)+Var(X).Var(Y) \ ,$$

and

$$Var[RD]=[E(XD)]^2.Var(YD)+[E(YD)]^2.Var(XD)+Var(XD).Var(YD)$$

But, since DS does not alter the sampled distribution, apart from inconsequential rounding errors, it follows that

$$E[X] = E[XD] \quad \text{and} \quad E[Y] = E[YD] \ ,$$

and that

$$Var[X] = Var[XD] \quad \text{and} \quad Var[Y] = Var[YD] \ ,$$

and, hence,

$$E[R] = E[RD], \text{ and}$$
$$Var[R] = Var[RD] \ .$$

Therefore, as far as the first two sample moments are concerned, DS produces unbiased estimates for the product of two independent random variables. Following a similar approach we could extend this proof to any moment related to the response variable distribution, thus showing that the DS response distribution is equal to the SRS response distribution. In this paper, however, we present an empirical verification of this property, which can be used to test this property for any kind of simulation and response variable, not just the product of two independent random variables.

Here, to study this product, the following input variable distributions were considered:

(a) Negative exponential,
(b) Normal,
(c) LogNormal,
(d) Triangular, and
(e) Uniform.

Different combinations of the two input variables were also considered, so that a more general conclusion could be drawn from the study. A simulation run was defined by N=1000 trials (products) so that in this case the input sample size was also N=1000 for both input variables. Based on the N response values, each run produced a set of 41 estimates:

(a) the mean of the N response values,
(b) the standard deviation of the N response values, and
(c) A set of k=39 fractiles:

$$Z_\alpha, \quad \alpha = 0.025,\ 0.050,\ \ldots,\ 0.950,\ 0.975\ .$$

For each combination of input variables under study, a simulation experiment, defined by M=1000 replicated runs using SRS and another M=1000 replicated runs using DS, was carried out. Each experiment produced the following results (Tables 2, 3 and 4 present such results for 3 experiments):

(a) Mean and standard deviation of the M mean estimates from the M runs, for both SRS and DS (First line of the bottom part of the table of results);

(b) Mean and standard deviation of the M standard deviation estimates from the M runs, for both SRS and DS (Second line of the bottom part of the table of results);

(c) For each of the K=39 fractiles under study, identified as $Z_\alpha$, $\alpha$=0.025,...,0.975, the mean and standard deviation of the M fractile values from the M runs;

(d) For each of the (K+2) estimates above, the mean ratio defined by (DS overall mean)/(SRS overall mean) and the standard deviation ratio defined by (DS overall standard deviation)/(SRS overall standard deviation). The mean ratio, if close to unity, will show that DS estimates are unbiased, whilst the standard deviation ratio will evaluate the standard error reduction achieved with DS.

Set values generation and fractile computation

Using the inverse transform method, descriptive set values $XD_i$, i=1,..,N , were generated in the following way:

(a) Uniform distribution.
**Uniform(a:lower limit, b:upper limit):**

$$XD_i = a + (b-a).(i-0.5)/n \ .$$

(b) Negative exponential distribution.
**Negexp(a:mean):**

$$XD_i = -a.Ln[(i-0.5)/n] \ .$$

(c) Normal distribution.

**Normal(m:mean, s:standard deviation):**

Using the Ramberg and Schmeiser (1972) approximation, a descriptive value for the standard Normal distribution is given by

$$ZD_i = \{[(i-0.5)/n]^{0.1349} + [1-(i-0.5)/n]^{0.1349}\}/0.1975$$

and

$$XD_i = m + s.ZD_i \ .$$

(d) LogNormal distribution.

**LogNormal(m:mean, s:standard deviation):**

Using the Ramberg and Schmeiser approximation above to generate descriptive values for the standard Normal, and then using

$$XD_i = Exp(ma + sa.ZD_i) \ ,$$

where

$$ma = Ln(m) - Ln[1+ (s/m)^2]/2 \ , \text{ and}$$

$$sa = \{Ln[1+ (s/m)^2]\}^{1/2}$$

are the mean and standard deviation of the Normal random variable that produces the LogNormal variable with mean **m** and standard deviation **s** (Law and Kelton, 1982).

(e) Triangular distribution.

**Triangular(a:lower limit, b:upper limit, c:most likely):**

Using the inverse transform (Law and Kelton, 1982), generate a value for the triangular(0,1,(c-a)/(b-a)), given by

$$YD_i = \{[(c-a)/(b-a)].[(i-0.5)/n]\}^{1/2} \ , \quad \text{if } [(i-0.5)/n \leq (c-a)/(b-a)] \ ,$$

or

$$YD_i = 1 - \{[1-(c-a)/(b-a)].[1-(i-0.5)/n]\}^{1/2} \ , \quad \text{if } [(i-0.5)/n > (c-a)/(b-a)] \ .$$

A descriptive value for the triangular(a,b,c) is given by

$$XD_i = a + (b-a).YD_i \ .$$

Simple random sampling values were also generated using the inverse transform, thus based on the same algorithms above. The only difference was the argument of the inverse transformation which, instead of a descriptive value in the unit interval defined by [(i-0.5)/n], was defined as a random uniform value.

For the fractile estimation, the N=1000 response values produced in each simulation run were first sorted into ascending order (we used the quicksort algorithm). Based on this set of ordered values $z_{(i)}$, i=1,..,N, the $\alpha^{th}$ fractile was given by

$$z_\alpha = z_{(k)} + [N\alpha-(k-0.5)].(z_{(k+1)}-z_{(k)}) ,$$

where

k is such that $k/N \leq \alpha < (k+1)/N$ .

For example, given that N=1000, the $\alpha$=0.500 fractile (the sample median) was given by the average

$$Z_{0.500} = (z_{(500)}+z_{(501)})/2 ,$$

with a similar result applying to any other fractile under study. Also based on the N=1000 response values from each run, their mean and standard deviation were computed in the usual way.

Finally, based on the results from M=1000 runs for each sampling method, the mean and standard deviation of the 41 estimates under study were computed. This procedure defined a simulation experiment, which was the same for all problems here studied. This experimental procedure was programmed in TURBO-PASCAL (source code available from the author) and run on a 486 microcomputer. Processing times ranged from 5 to 10 minutes for each experiment.

**RESULTS**

Table 1 presents a list of the problems used in our tests. A notable finding was that, for all problems and for all estimates, DS performed better than SRS. As a whole, precision gains, which varied between problems, were very high for the mean and lower, but still substantial, for the standard deviation. Concerning the fractiles, standard error ratios typically ranged from 0.5 to 0.9, with lower ratios (greater gains) near the distribution centre (median) and higher ratios (lower gains) near the distribution extremes.

As an example, Table 2 presents the results from one of such experiments, for problem 22. This problem concerns the product of two LogNormal random variables and was previously studied by Zaino and D'Errico (1989) in the context of decision and risk analysis. As seen, the mean ratios were very close to one, therefore showing that DS produces unbiased estimates. On the other hand, the standard deviation ratios, being lower than one, confirm that DS estimates are more precise. Those properties are better seen in figure 1, where the mean and standard deviation ratios are plotted for the whole set of fractiles under study.

To give further evidence in favour of DS, the results and corresponding graphics for two other problems are displayed: Table 3 and figure 2, which refer to problem 1 (Product of two negative exponential random variables); and Table 4 and figure 3, which refer to problem 11 (Product of an uniform and a triangular random variable).

**Table 1. List of problems for which Z=X*Y was studied**

| Problem number | X Variable | Y Variable |
|---|---|---|
| 1 | Negexp(4) | Negexp(1) |
| 2 | Negexp(1) | Negexp(1) |
| 3 | LogNormal(3,1) | LogNormal(1.5,0.3) |
| 4 | LogNormal(1.5,0.3) | LogNormal(1.5,0.3) |
| 5 | Normal(3.4,0.15) | Normal(-1.5,0.3) |
| 6 | Normal(3.4,0.15) | Normal(3.4,0.15) |
| 7 | Uniform(2,6) | Uniform(-1,4) |
| 8 | Uniform(-1,4) | Uniform(-1,4) |
| 9 | Triangular(0,1,0.5) | Triangular(2,4,3) |
| 10 | Triangular(0,1,0.5) | Triangular(0,1,0.5) |
| 11 | Uniform(2,6) | Triangular(2,6,4) |
| 12 | Uniform(0,1) | LogNormal(1.5,0.3) |
| 13 | Uniform(2,4) | Negexp(3) |
| 14 | Uniform(-1,4) | Normal(3.4,0.3) |
| 15 | Triangular(1,5,3) | Negexp(2) |
| 16 | Triangular(1,6,4) | LogNormal(2.5,0.2) |
| 17 | Triangular(0,1,0.1) | Normal(3.4,0.3) |
| 18 | Negexp(2) | LogNormal(1.5,0.3) |
| 19 | Negexp(1) | Normal(3.4,0.15) |
| 20 | LogNormal(1.5,0.3) | Normal(3.4,0.15) |
| 21 | Triangular(1,5,3) | Negexp(3) |
| 22 | LogNormal(30.3,4.571) | LogNormal(0.23,0.072) |

**Table 2. Summary of results for problem 22: Z = X.Y**
**X ~ LogNormal(30.3031,4.5712), Y ~**
**LogNormal(0.2334,0.0716)**
**Sample size = 1000 , Number of runs = 1000**

| | ---- SRS ---- | | ---- DS ----- | | Mean ratio | St Dev ratio |
|---|---|---|---|---|---|---|
| Fractile | mean | st dev | mean | st dev | DS/SRS | DS/SRS |
| 0.025 | 3.4598 | 0.1007 | 3.4529 | 0.0708 | 0.9980 | 0.7029 |
| 0.050 | 3.8489 | 0.0892 | 3.8429 | 0.0577 | 0.9985 | 0.6470 |
| 0.075 | 4.1230 | 0.0843 | 4.1194 | 0.0518 | 0.9991 | 0.6147 |
| 0.100 | 4.3504 | 0.0824 | 4.3447 | 0.0504 | 0.9987 | 0.6121 |
| 0.125 | 4.5472 | 0.0793 | 4.5426 | 0.0497 | 0.9990 | 0.6268 |
| 0.150 | 4.7254 | 0.0773 | 4.7236 | 0.0485 | 0.9996 | 0.6275 |
| 0.175 | 4.8890 | 0.0766 | 4.8864 | 0.0463 | 0.9995 | 0.6048 |
| 0.200 | 5.0436 | 0.0767 | 5.0410 | 0.0458 | 0.9995 | 0.5979 |
| 0.225 | 5.1913 | 0.0787 | 5.1905 | 0.0442 | 0.9998 | 0.5619 |
| 0.250 | 5.3334 | 0.0782 | 5.3338 | 0.0434 | 1.0001 | 0.5546 |
| 0.275 | 5.4714 | 0.0798 | 5.4726 | 0.0434 | 1.0002 | 0.5438 |
| 0.300 | 5.6067 | 0.0816 | 5.6101 | 0.0425 | 1.0006 | 0.5216 |
| 0.325 | 5.7409 | 0.0812 | 5.7443 | 0.0412 | 1.0006 | 0.5074 |
| 0.350 | 5.8743 | 0.0826 | 5.8767 | 0.0420 | 1.0004 | 0.5083 |
| 0.375 | 6.0062 | 0.0843 | 6.0084 | 0.0431 | 1.0004 | 0.5111 |
| 0.400 | 6.1397 | 0.0846 | 6.1398 | 0.0417 | 1.0000 | 0.4934 |
| 0.425 | 6.2727 | 0.0857 | 6.2748 | 0.0416 | 1.0003 | 0.4852 |
| 0.450 | 6.4073 | 0.0872 | 6.4082 | 0.0436 | 1.0001 | 0.4997 |
| 0.475 | 6.5438 | 0.0886 | 6.5451 | 0.0455 | 1.0002 | 0.5133 |
| 0.500 | 6.6822 | 0.0913 | 6.6836 | 0.0459 | 1.0002 | 0.5023 |
| 0.525 | 6.8235 | 0.0919 | 6.8250 | 0.0472 | 1.0002 | 0.5138 |
| 0.550 | 6.9691 | 0.0933 | 6.9704 | 0.0502 | 1.0002 | 0.5386 |
| 0.575 | 7.1171 | 0.0939 | 7.1200 | 0.0504 | 1.0004 | 0.5367 |
| 0.600 | 7.2718 | 0.0988 | 7.2757 | 0.0525 | 1.0005 | 0.5318 |
| 0.625 | 7.4323 | 0.1031 | 7.4374 | 0.0541 | 1.0007 | 0.5244 |
| 0.650 | 7.5999 | 0.1063 | 7.6044 | 0.0573 | 1.0006 | 0.5389 |
| 0.675 | 7.7771 | 0.1098 | 7.7795 | 0.0611 | 1.0003 | 0.5559 |
| 0.700 | 7.9630 | 0.1125 | 7.9646 | 0.0638 | 1.0002 | 0.5669 |
| 0.725 | 8.1603 | 0.1183 | 8.1623 | 0.0659 | 1.0002 | 0.5574 |
| 0.750 | 8.3746 | 0.1232 | 8.3730 | 0.0668 | 0.9998 | 0.5424 |
| 0.775 | 8.6043 | 0.1300 | 8.5998 | 0.0700 | 0.9995 | 0.5385 |
| 0.800 | 8.8591 | 0.1382 | 8.8535 | 0.0742 | 0.9994 | 0.5365 |
| 0.825 | 9.1417 | 0.1487 | 9.1360 | 0.0820 | 0.9994 | 0.5517 |
| 0.850 | 9.4600 | 0.1617 | 9.4498 | 0.0920 | 0.9989 | 0.5693 |
| 0.875 | 9.8290 | 0.1703 | 9.8224 | 0.1001 | 0.9993 | 0.5876 |
| 0.900 | 10.2704 | 0.1888 | 10.2664 | 0.1173 | 0.9996 | 0.6216 |
| 0.925 | 10.8320 | 0.2205 | 10.8199 | 0.1362 | 0.9989 | 0.6177 |
| 0.950 | 11.5950 | 0.2611 | 11.5900 | 0.1814 | 0.9996 | 0.6949 |
| 0.975 | 12.8822 | 0.3610 | 12.8719 | 0.2595 | 0.9992 | 0.7189 |
| MEAN | 7.0735 | 0.0799 | 7.0716 | 0.0102 | 0.9997 | 0.1279 |
| ST DEV | 2.4395 | 0.0786 | 2.4355 | 0.0438 | 0.9984 | 0.5572 |

Figure 1. Mean and St Deviation ratios
(Problem 22)

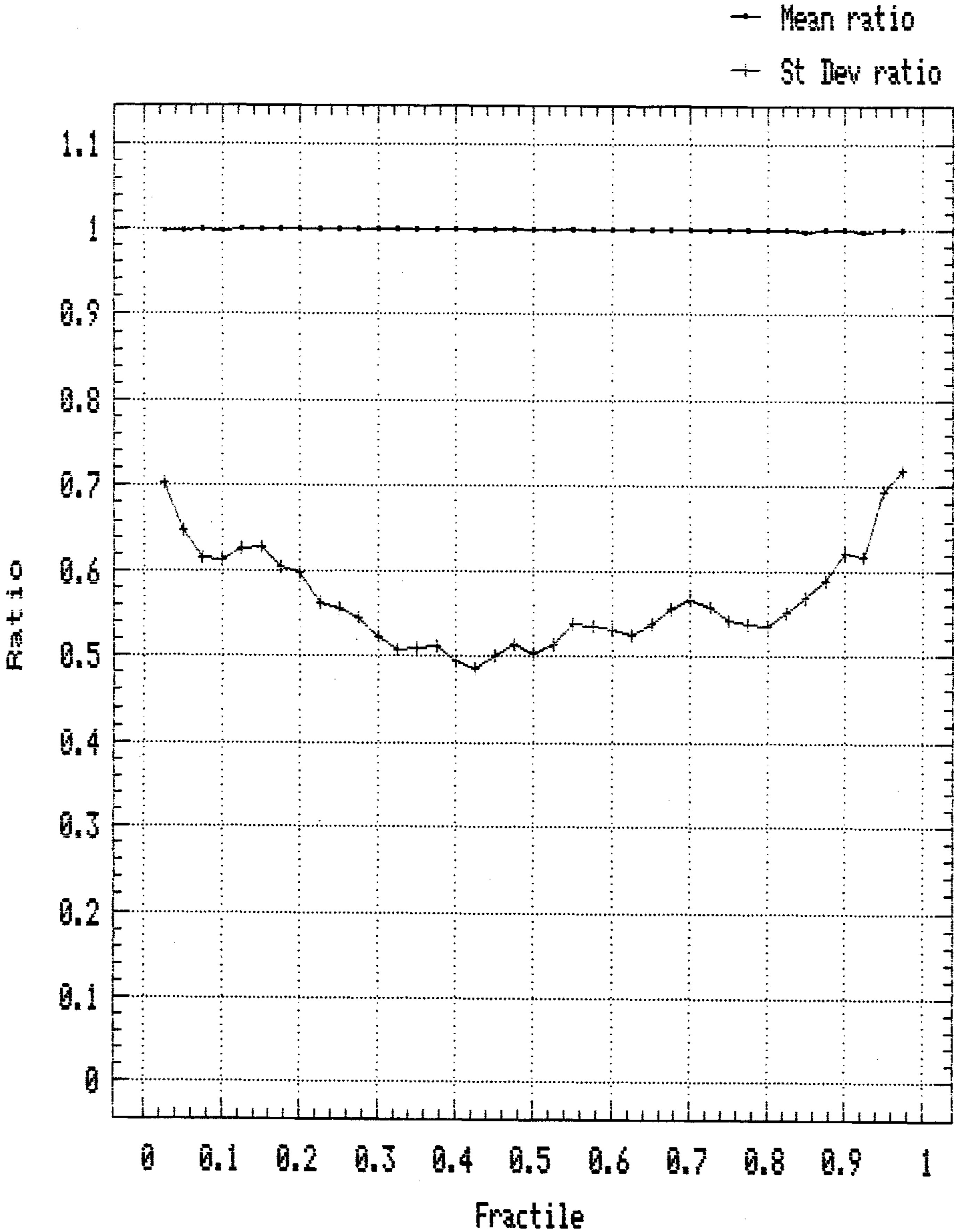


**Table 3. Summary of results for problem 1: Z = X.Y**
**X ~ Negexp(4), Y ~ Negexp(1)**
**Sample size = 1000 , Number of runs = 1000**

| Fractile | ---- SRS ---- mean | st dev | ---- DS ----- mean | st dev | Mean ratio DS/SRS | St Dev ratio DS/SRS |
|---|---|---|---|---|---|---|
| 0.025 | 0.0191 | 0.0047 | 0.0192 | 0.0033 | 1.0075 | 0.7190 |
| 0.050 | 0.0461 | 0.0083 | 0.0458 | 0.0056 | 0.9927 | 0.6708 |
| 0.075 | 0.0784 | 0.0117 | 0.0779 | 0.0074 | 0.9933 | 0.6295 |
| 0.100 | 0.1155 | 0.0156 | 0.1151 | 0.0093 | 0.9969 | 0.5986 |
| 0.125 | 0.1577 | 0.0191 | 0.1571 | 0.0115 | 0.9961 | 0.6006 |
| 0.150 | 0.2045 | 0.0229 | 0.2043 | 0.0136 | 0.9991 | 0.5934 |
| 0.175 | 0.2567 | 0.0267 | 0.2564 | 0.0155 | 0.9990 | 0.5800 |
| 0.200 | 0.3137 | 0.0316 | 0.3137 | 0.0181 | 1.0000 | 0.5720 |
| 0.225 | 0.3771 | 0.0364 | 0.3764 | 0.0203 | 0.9983 | 0.5562 |
| 0.250 | 0.4460 | 0.0410 | 0.4455 | 0.0225 | 0.9990 | 0.5501 |
| 0.275 | 0.5214 | 0.0456 | 0.5203 | 0.0255 | 0.9978 | 0.5581 |
| 0.300 | 0.6030 | 0.0506 | 0.6011 | 0.0282 | 0.9968 | 0.5577 |
| 0.325 | 0.6919 | 0.0562 | 0.6900 | 0.0316 | 0.9973 | 0.5626 |
| 0.350 | 0.7885 | 0.0617 | 0.7873 | 0.0358 | 0.9985 | 0.5806 |
| 0.375 | 0.8945 | 0.0672 | 0.8917 | 0.0394 | 0.9968 | 0.5869 |
| 0.400 | 1.0094 | 0.0756 | 1.0057 | 0.0423 | 0.9964 | 0.5601 |
| 0.425 | 1.1344 | 0.0828 | 1.1300 | 0.0467 | 0.9961 | 0.5642 |
| 0.450 | 1.2715 | 0.0916 | 1.2660 | 0.0509 | 0.9957 | 0.5560 |
| 0.475 | 1.4190 | 0.1003 | 1.4135 | 0.0555 | 0.9961 | 0.5527 |
| 0.500 | 1.5817 | 0.1115 | 1.5767 | 0.0603 | 0.9968 | 0.5413 |
| 0.525 | 1.7600 | 0.1196 | 1.7534 | 0.0658 | 0.9962 | 0.5506 |
| 0.550 | 1.9563 | 0.1300 | 1.9482 | 0.0722 | 0.9958 | 0.5550 |
| 0.575 | 2.1697 | 0.1416 | 2.1640 | 0.0790 | 0.9973 | 0.5577 |
| 0.600 | 2.4089 | 0.1584 | 2.4000 | 0.0896 | 0.9963 | 0.5656 |
| 0.625 | 2.6702 | 0.1738 | 2.6649 | 0.0995 | 0.9980 | 0.5728 |
| 0.650 | 2.9638 | 0.1912 | 2.9594 | 0.1111 | 0.9985 | 0.5810 |
| 0.675 | 3.2938 | 0.2114 | 3.2879 | 0.1217 | 0.9982 | 0.5756 |
| 0.700 | 3.6618 | 0.2373 | 3.6562 | 0.1341 | 0.9985 | 0.5652 |
| 0.725 | 4.0845 | 0.2693 | 4.0753 | 0.1536 | 0.9977 | 0.5703 |
| 0.750 | 4.5708 | 0.3026 | 4.5614 | 0.1784 | 0.9979 | 0.5894 |
| 0.775 | 5.1330 | 0.3420 | 5.1170 | 0.2071 | 0.9969 | 0.6056 |
| 0.800 | 5.7872 | 0.3857 | 5.7755 | 0.2481 | 0.9980 | 0.6432 |
| 0.825 | 6.5647 | 0.4304 | 6.5602 | 0.2952 | 0.9993 | 0.6858 |
| 0.850 | 7.5257 | 0.4903 | 7.5259 | 0.3555 | 1.0000 | 0.7250 |
| 0.875 | 8.7161 | 0.5765 | 8.7454 | 0.4290 | 1.0034 | 0.7440 |
| 0.900 | 10.2833 | 0.7101 | 10.3277 | 0.5310 | 1.0043 | 0.7478 |
| 0.925 | 12.4900 | 0.9133 | 12.5240 | 0.6995 | 1.0027 | 0.7659 |
| 0.950 | 15.8791 | 1.2436 | 15.9242 | 0.9423 | 1.0028 | 0.7577 |
| 0.975 | 22.4282 | 2.0774 | 22.5074 | 1.7036 | 1.0035 | 0.8201 |
| MEAN | 3.9939 | 0.2278 | 3.9941 | 0.1267 | 1.0001 | 0.5560 |
| ST DEV | 6.8759 | 0.7608 | 6.8453 | 0.6044 | 0.9955 | 0.7944 |

Figure 2. Mean and St Deviation ratios
(Problem 1)

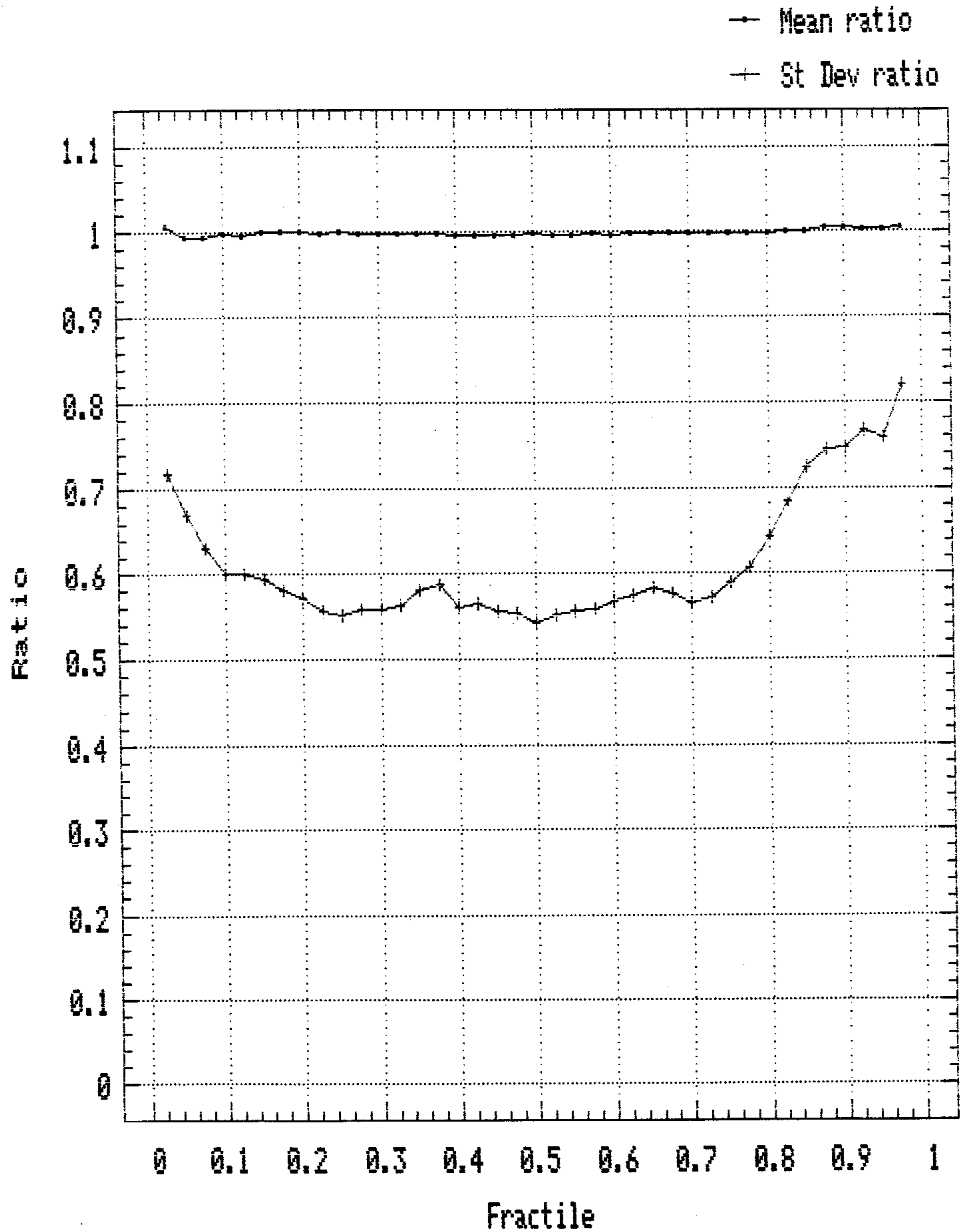
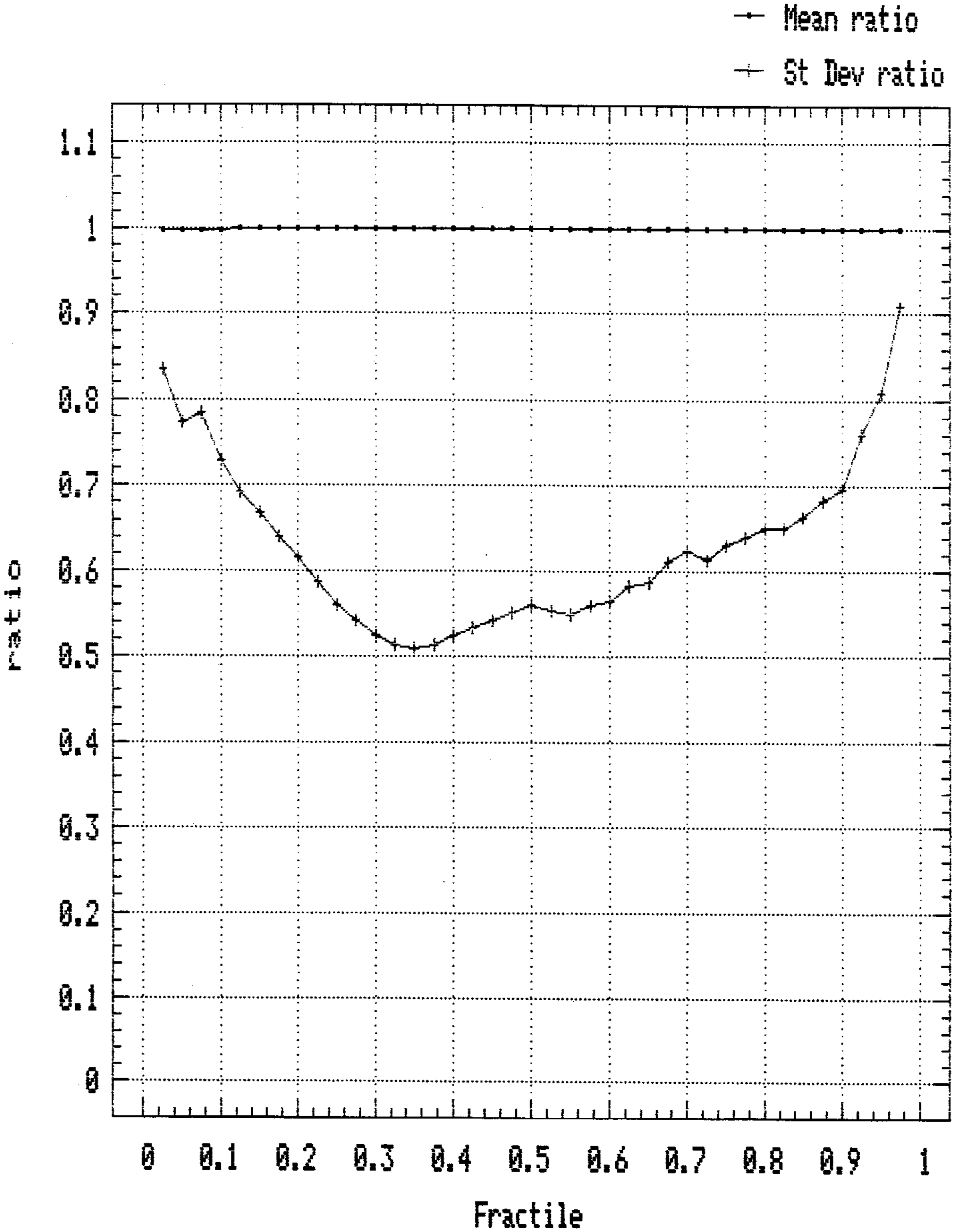


**Table 4. Summary of results for problem 11: Z = X.Y**
**X ~ Uniform(2,6), Y ~ Triangular(2,6,4)**
**Sample size = 1000 , Number of runs = 1000**

| | ---- SRS ---- | | ---- DS ----- | | Mean ratio | St Dev ratio |
|---|---|---|---|---|---|---|
| Fractile | mean | st dev | mean | st dev | DS/SRS | DS/SRS |
| 0.025 | 6.9591 | 0.2162 | 6.9477 | 0.1803 | 0.9984 | 0.8341 |
| 0.050 | 7.8113 | 0.1982 | 7.8014 | 0.1534 | 0.9987 | 0.7737 |
| 0.075 | 8.4354 | 0.1889 | 8.4237 | 0.1482 | 0.9986 | 0.7846 |
| 0.100 | 8.9582 | 0.1846 | 8.9471 | 0.1346 | 0.9988 | 0.7293 |
| 0.125 | 9.4225 | 0.1917 | 9.4187 | 0.1327 | 0.9996 | 0.6924 |
| 0.150 | 9.8614 | 0.1982 | 9.8591 | 0.1323 | 0.9998 | 0.6676 |
| 0.175 | 10.2760 | 0.2037 | 10.2742 | 0.1300 | 0.9998 | 0.6382 |
| 0.200 | 10.6802 | 0.2065 | 10.6782 | 0.1272 | 0.9998 | 0.6159 |
| 0.225 | 11.0762 | 0.2146 | 11.0718 | 0.1259 | 0.9996 | 0.5869 |
| 0.250 | 11.4627 | 0.2194 | 11.4566 | 0.1228 | 0.9995 | 0.5595 |
| 0.275 | 11.8415 | 0.2254 | 11.8393 | 0.1224 | 0.9998 | 0.5427 |
| 0.300 | 12.2228 | 0.2326 | 12.2230 | 0.1218 | 1.0000 | 0.5237 |
| 0.325 | 12.6080 | 0.2368 | 12.6066 | 0.1217 | 0.9999 | 0.5137 |
| 0.350 | 12.9959 | 0.2426 | 12.9942 | 0.1238 | 0.9999 | 0.5104 |
| 0.375 | 13.3780 | 0.2519 | 13.3808 | 0.1296 | 1.0002 | 0.5147 |
| 0.400 | 13.7675 | 0.2486 | 13.7697 | 0.1304 | 1.0002 | 0.5245 |
| 0.425 | 14.1605 | 0.2521 | 14.1621 | 0.1344 | 1.0001 | 0.5329 |
| 0.450 | 14.5545 | 0.2526 | 14.5590 | 0.1367 | 1.0003 | 0.5413 |
| 0.475 | 14.9501 | 0.2551 | 14.9539 | 0.1407 | 1.0002 | 0.5516 |
| 0.500 | 15.3612 | 0.2572 | 15.3605 | 0.1439 | 1.0000 | 0.5597 |
| 0.525 | 15.7705 | 0.2620 | 15.7683 | 0.1448 | 0.9999 | 0.5525 |
| 0.550 | 16.1852 | 0.2644 | 16.1865 | 0.1449 | 1.0001 | 0.5480 |
| 0.575 | 16.6122 | 0.2703 | 16.6132 | 0.1510 | 1.0001 | 0.5589 |
| 0.600 | 17.0466 | 0.2735 | 17.0498 | 0.1543 | 1.0002 | 0.5641 |
| 0.625 | 17.4933 | 0.2785 | 17.4976 | 0.1619 | 1.0002 | 0.5811 |
| 0.650 | 17.9540 | 0.2846 | 17.9553 | 0.1669 | 1.0001 | 0.5866 |
| 0.675 | 18.4291 | 0.2871 | 18.4253 | 0.1754 | 0.9998 | 0.6109 |
| 0.700 | 18.9131 | 0.2915 | 18.9160 | 0.1816 | 1.0002 | 0.6228 |
| 0.725 | 19.4248 | 0.2970 | 19.4206 | 0.1822 | 0.9998 | 0.6134 |
| 0.750 | 19.9544 | 0.2983 | 19.9445 | 0.1878 | 0.9995 | 0.6297 |
| 0.775 | 20.5108 | 0.3070 | 20.5036 | 0.1965 | 0.9997 | 0.6401 |
| 0.800 | 21.0996 | 0.3108 | 21.0886 | 0.2018 | 0.9995 | 0.6491 |
| 0.825 | 21.7202 | 0.3221 | 21.7136 | 0.2093 | 0.9997 | 0.6498 |
| 0.850 | 22.4066 | 0.3329 | 22.3924 | 0.2209 | 0.9994 | 0.6636 |
| 0.875 | 23.1603 | 0.3416 | 23.1490 | 0.2330 | 0.9995 | 0.6821 |
| 0.900 | 24.0165 | 0.3581 | 24.0116 | 0.2493 | 0.9998 | 0.6961 |
| 0.925 | 25.0320 | 0.3811 | 25.0400 | 0.2896 | 1.0003 | 0.7600 |
| 0.950 | 26.3510 | 0.4154 | 26.3344 | 0.3363 | 0.9994 | 0.8096 |
| 0.975 | 28.2242 | 0.4761 | 28.2023 | 0.4332 | 0.9992 | 0.9099 |
| MEAN | 16.0037 | 0.1880 | 15.9994 | 0.0301 | 0.9997 | 0.1603 |
| ST DEV | 5.7367 | 0.1156 | 5.7343 | 0.0884 | 0.9996 | 0.7647 |

Figure 3. Mean and St Deviation ratios
(Problem 11)



## CONCLUSIONS

The results showed that DS produces better estimates for the whole response variable distribution, not just the mean and standard deviation. This same methodology can also be applied to any other simulation problem. In fact, studying the project duration in a PERT network, we arrived at similar results as those reported here. Although not a common practice in simulation studies, the estimation of a response variable distribution is a valuable information for the decision maker. For instance, this is the case in a risk analysis, where a risk profile (the probability distribution of the net present value or the internal rate of return) is usually estimated. As such, Descriptive Sampling represents a highly promising improvement as far as risk analysis and other simulation studies are concerned.

**REFERENCES**


LAW, A.M.; KELTON, W.D. Simulation modeling and analysis. New York: McGraw-Hill, 1992.

RAMBERG, B.W.; SCHMEISER, J.S. An approximate method for generating symmetric random variables. Comm of the ACM, v.1,p.987-990, 1972.

SALIBY, E. Descriptive sampling: a better approach to Monte Carlo simulation. J. Opl Res. Soc., v.41, p.1133-1142, 1990.

________. Rethinking simulation: descriptive sampling. São Paulo: Atlas/EDUFRJ, 1989. (In Portuguese).

________.; PAUL, R.J. Implementing descriptive sampling in three-phase discrete event simulation models. J. Opl Res. Soc., v. 44, p.147-160, 1993.

ZAINO, N.A.; D'ERRICO, J. Optimal discrete approximations for continuous outcomes with applications in decision and risk analysis. J. Opl Res. Soc., v.40, p.379-388, 1989.

Figure 1. Mean and St Deviation ratios
(Problem 22)

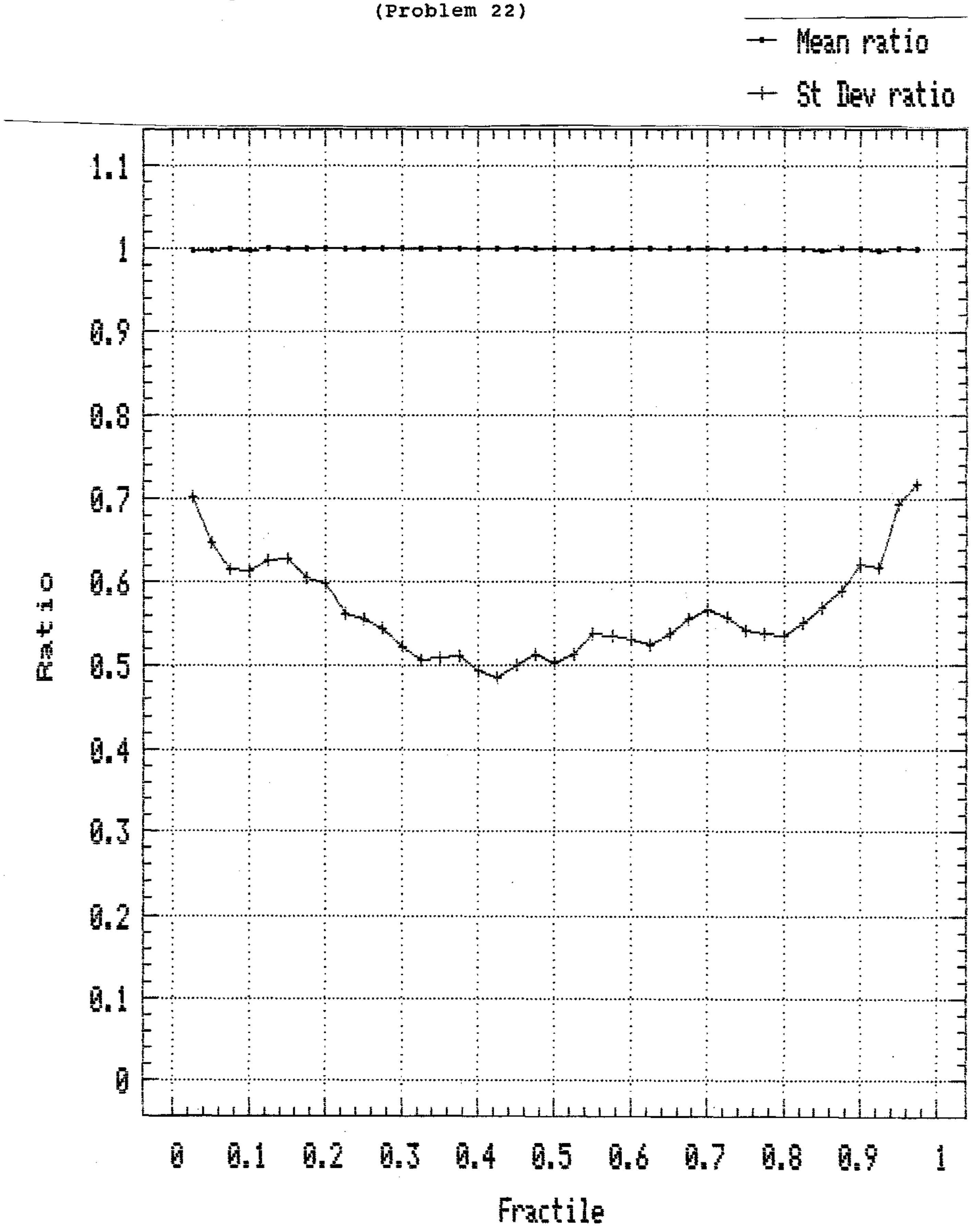


**Table 3. Summary of results for problem 1: Z = X.Y**
**X ~ Negexp(4), Y ~ Negexp(1)**
**Sample size = 1000 , Number of runs = 1000**

| Fractile | SRS mean | SRS st dev | DS mean | DS st dev | Mean ratio DS/SRS | St Dev ratio DS/SRS |
|---|---|---|---|---|---|---|
| 0.025 | 0.0191 | 0.0047 | 0.0192 | 0.0033 | 1.0075 | 0.7190 |
| 0.050 | 0.0461 | 0.0083 | 0.0458 | 0.0056 | 0.9927 | 0.6708 |
| 0.075 | 0.0784 | 0.0117 | 0.0779 | 0.0074 | 0.9933 | 0.6295 |
| 0.100 | 0.1155 | 0.0156 | 0.1151 | 0.0093 | 0.9969 | 0.5986 |
| 0.125 | 0.1577 | 0.0191 | 0.1571 | 0.0115 | 0.9961 | 0.6006 |
| 0.150 | 0.2045 | 0.0229 | 0.2043 | 0.0136 | 0.9991 | 0.5934 |
| 0.175 | 0.2567 | 0.0267 | 0.2564 | 0.0155 | 0.9990 | 0.5800 |
| 0.200 | 0.3137 | 0.0316 | 0.3137 | 0.0181 | 1.0000 | 0.5720 |
| 0.225 | 0.3771 | 0.0364 | 0.3764 | 0.0203 | 0.9983 | 0.5562 |
| 0.250 | 0.4460 | 0.0410 | 0.4455 | 0.0225 | 0.9990 | 0.5501 |
| 0.275 | 0.5214 | 0.0456 | 0.5203 | 0.0255 | 0.9978 | 0.5581 |
| 0.300 | 0.6030 | 0.0506 | 0.6011 | 0.0282 | 0.9968 | 0.5577 |
| 0.325 | 0.6919 | 0.0562 | 0.6900 | 0.0316 | 0.9973 | 0.5626 |
| 0.350 | 0.7885 | 0.0617 | 0.7873 | 0.0358 | 0.9985 | 0.5806 |
| 0.375 | 0.8945 | 0.0672 | 0.8917 | 0.0394 | 0.9968 | 0.5869 |
| 0.400 | 1.0094 | 0.0756 | 1.0057 | 0.0423 | 0.9964 | 0.5601 |
| 0.425 | 1.1344 | 0.0828 | 1.1300 | 0.0467 | 0.9961 | 0.5642 |
| 0.450 | 1.2715 | 0.0916 | 1.2660 | 0.0509 | 0.9957 | 0.5560 |
| 0.475 | 1.4190 | 0.1003 | 1.4135 | 0.0555 | 0.9961 | 0.5527 |
| 0.500 | 1.5817 | 0.1115 | 1.5767 | 0.0603 | 0.9968 | 0.5413 |
| 0.525 | 1.7600 | 0.1196 | 1.7534 | 0.0658 | 0.9962 | 0.5506 |
| 0.550 | 1.9563 | 0.1300 | 1.9482 | 0.0722 | 0.9958 | 0.5550 |
| 0.575 | 2.1697 | 0.1416 | 2.1640 | 0.0790 | 0.9973 | 0.5577 |
| 0.600 | 2.4089 | 0.1584 | 2.4000 | 0.0896 | 0.9963 | 0.5656 |
| 0.625 | 2.6702 | 0.1738 | 2.6649 | 0.0995 | 0.9980 | 0.5728 |
| 0.650 | 2.9638 | 0.1912 | 2.9594 | 0.1111 | 0.9985 | 0.5810 |
| 0.675 | 3.2938 | 0.2114 | 3.2879 | 0.1217 | 0.9982 | 0.5756 |
| 0.700 | 3.6618 | 0.2373 | 3.6562 | 0.1341 | 0.9985 | 0.5652 |
| 0.725 | 4.0845 | 0.2693 | 4.0753 | 0.1536 | 0.9977 | 0.5703 |
| 0.750 | 4.5708 | 0.3026 | 4.5614 | 0.1784 | 0.9979 | 0.5894 |
| 0.775 | 5.1330 | 0.3420 | 5.1170 | 0.2071 | 0.9969 | 0.6056 |
| 0.800 | 5.7872 | 0.3857 | 5.7755 | 0.2481 | 0.9980 | 0.6432 |
| 0.825 | 6.5647 | 0.4304 | 6.5602 | 0.2952 | 0.9993 | 0.6858 |
| 0.850 | 7.5257 | 0.4903 | 7.5259 | 0.3555 | 1.0000 | 0.7250 |
| 0.875 | 8.7161 | 0.5765 | 8.7454 | 0.4290 | 1.0034 | 0.7440 |
| 0.900 | 10.2833 | 0.7101 | 10.3277 | 0.5310 | 1.0043 | 0.7478 |
| 0.925 | 12.4900 | 0.9133 | 12.5240 | 0.6995 | 1.0027 | 0.7659 |
| 0.950 | 15.8791 | 1.2436 | 15.9242 | 0.9423 | 1.0028 | 0.7577 |
| 0.975 | 22.4282 | 2.0774 | 22.5074 | 1.7036 | 1.0035 | 0.8201 |
| MEAN | 3.9939 | 0.2278 | 3.9941 | 0.1267 | 1.0001 | 0.5560 |
| ST DEV | 6.8759 | 0.7608 | 6.8453 | 0.6044 | 0.9955 | 0.7944 |

Figure 2. Mean and St Deviation ratios
(Problem 1)

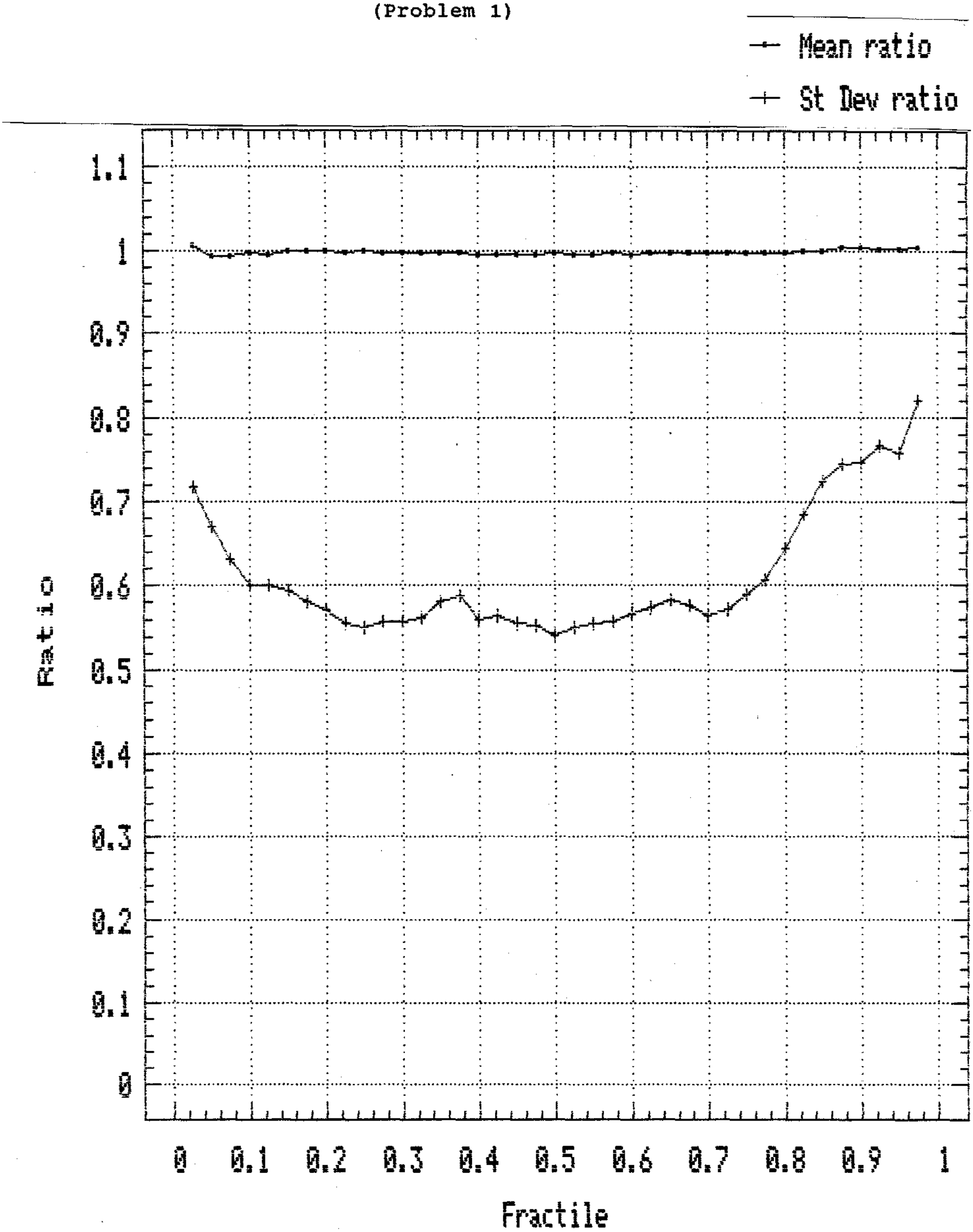
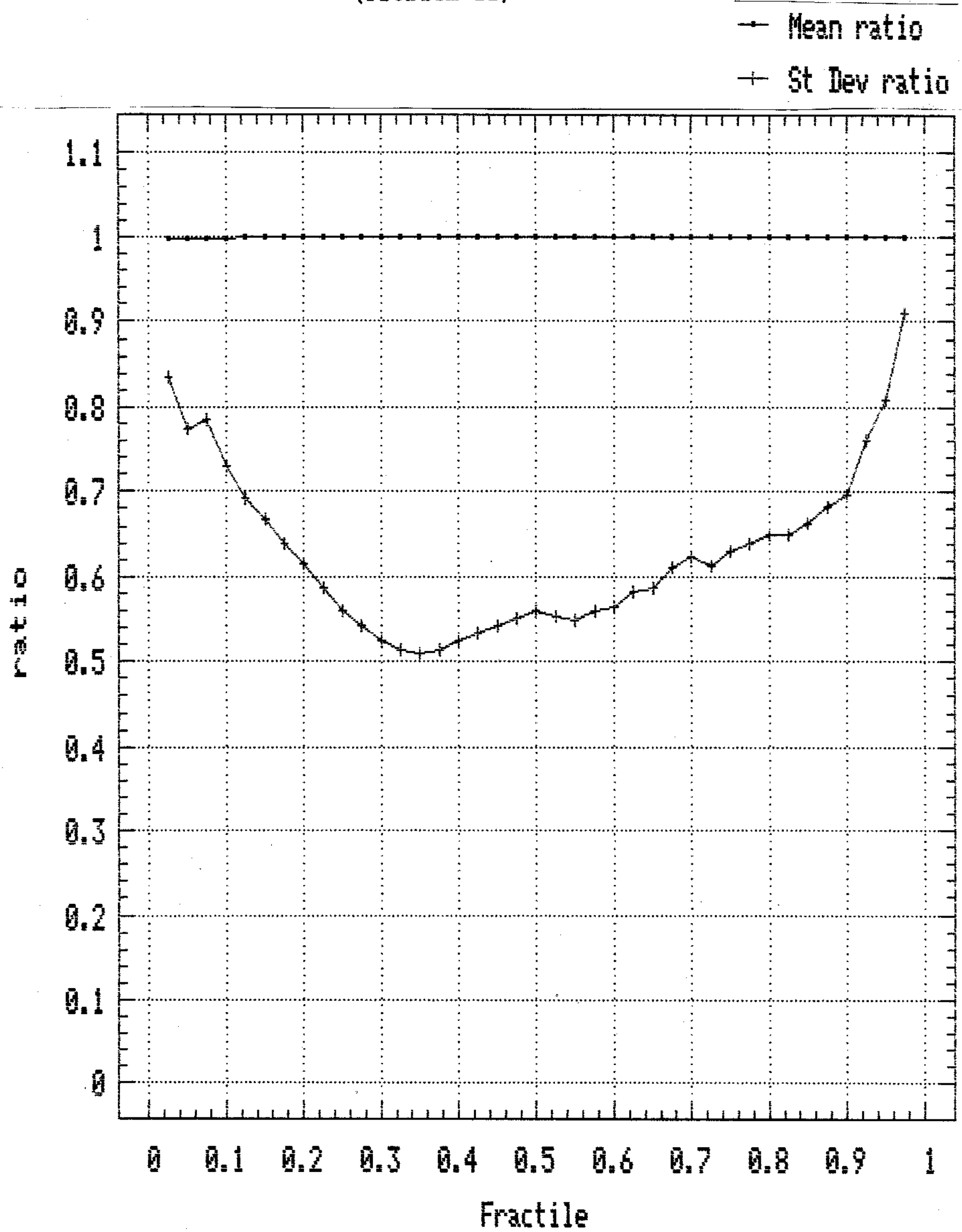


**Table 4. Summary of results for problem 11: Z = X.Y**
**X ~ Uniform(2,6), Y ~ Triangular(2,6,4)**
**Sample size = 1000 , Number of runs = 1000**

| | ---- SRS ---- | | ---- DS ----- | | Mean ratio | St Dev ratio |
|---|---|---|---|---|---|---|
| Fractile | mean | st dev | mean | st dev | DS/SRS | DS/SRS |
| 0.025 | 6.9591 | 0.2162 | 6.9477 | 0.1803 | 0.9984 | 0.8341 |
| 0.050 | 7.8113 | 0.1982 | 7.8014 | 0.1534 | 0.9987 | 0.7737 |
| 0.075 | 8.4354 | 0.1889 | 8.4237 | 0.1482 | 0.9986 | 0.7846 |
| 0.100 | 8.9582 | 0.1846 | 8.9471 | 0.1346 | 0.9988 | 0.7293 |
| 0.125 | 9.4225 | 0.1917 | 9.4187 | 0.1327 | 0.9996 | 0.6924 |
| 0.150 | 9.8614 | 0.1982 | 9.8591 | 0.1323 | 0.9998 | 0.6676 |
| 0.175 | 10.2760 | 0.2037 | 10.2742 | 0.1300 | 0.9998 | 0.6382 |
| 0.200 | 10.6802 | 0.2065 | 10.6782 | 0.1272 | 0.9998 | 0.6159 |
| 0.225 | 11.0762 | 0.2146 | 11.0718 | 0.1259 | 0.9996 | 0.5869 |
| 0.250 | 11.4627 | 0.2194 | 11.4566 | 0.1228 | 0.9995 | 0.5595 |
| 0.275 | 11.8415 | 0.2254 | 11.8393 | 0.1224 | 0.9998 | 0.5427 |
| 0.300 | 12.2228 | 0.2326 | 12.2230 | 0.1218 | 1.0000 | 0.5237 |
| 0.325 | 12.6080 | 0.2368 | 12.6066 | 0.1217 | 0.9999 | 0.5137 |
| 0.350 | 12.9959 | 0.2426 | 12.9942 | 0.1238 | 0.9999 | 0.5104 |
| 0.375 | 13.3780 | 0.2519 | 13.3808 | 0.1296 | 1.0002 | 0.5147 |
| 0.400 | 13.7675 | 0.2486 | 13.7697 | 0.1304 | 1.0002 | 0.5245 |
| 0.425 | 14.1605 | 0.2521 | 14.1621 | 0.1344 | 1.0001 | 0.5329 |
| 0.450 | 14.5545 | 0.2526 | 14.5590 | 0.1367 | 1.0003 | 0.5413 |
| 0.475 | 14.9501 | 0.2551 | 14.9539 | 0.1407 | 1.0002 | 0.5516 |
| 0.500 | 15.3612 | 0.2572 | 15.3605 | 0.1439 | 1.0000 | 0.5597 |
| 0.525 | 15.7705 | 0.2620 | 15.7683 | 0.1448 | 0.9999 | 0.5525 |
| 0.550 | 16.1852 | 0.2644 | 16.1865 | 0.1449 | 1.0001 | 0.5480 |
| 0.575 | 16.6122 | 0.2703 | 16.6132 | 0.1510 | 1.0001 | 0.5589 |
| 0.600 | 17.0466 | 0.2735 | 17.0498 | 0.1543 | 1.0002 | 0.5641 |
| 0.625 | 17.4933 | 0.2785 | 17.4976 | 0.1619 | 1.0002 | 0.5811 |
| 0.650 | 17.9540 | 0.2846 | 17.9553 | 0.1669 | 1.0001 | 0.5866 |
| 0.675 | 18.4291 | 0.2871 | 18.4253 | 0.1754 | 0.9998 | 0.6109 |
| 0.700 | 18.9131 | 0.2915 | 18.9160 | 0.1816 | 1.0002 | 0.6228 |
| 0.725 | 19.4248 | 0.2970 | 19.4206 | 0.1822 | 0.9998 | 0.6134 |
| 0.750 | 19.9544 | 0.2983 | 19.9445 | 0.1878 | 0.9995 | 0.6297 |
| 0.775 | 20.5108 | 0.3070 | 20.5036 | 0.1965 | 0.9997 | 0.6401 |
| 0.800 | 21.0996 | 0.3108 | 21.0886 | 0.2018 | 0.9995 | 0.6491 |
| 0.825 | 21.7202 | 0.3221 | 21.7136 | 0.2093 | 0.9997 | 0.6498 |
| 0.850 | 22.4066 | 0.3329 | 22.3924 | 0.2209 | 0.9994 | 0.6636 |
| 0.875 | 23.1603 | 0.3416 | 23.1490 | 0.2330 | 0.9995 | 0.6821 |
| 0.900 | 24.0165 | 0.3581 | 24.0116 | 0.2493 | 0.9998 | 0.6961 |
| 0.925 | 25.0320 | 0.3811 | 25.0400 | 0.2896 | 1.0003 | 0.7600 |
| 0.950 | 26.3510 | 0.4154 | 26.3344 | 0.3363 | 0.9994 | 0.8096 |
| 0.975 | 28.2242 | 0.4761 | 28.2023 | 0.4332 | 0.9992 | 0.9099 |
| MEAN | 16.0037 | 0.1880 | 15.9994 | 0.0301 | 0.9997 | 0.1603 |
| ST DEV | 5.7367 | 0.1156 | 5.7343 | 0.0884 | 0.9996 | 0.7647 |

Figure 3. Mean and St Deviation ratios
(Problem 11)



## CONCLUSIONS

The results showed that DS produces better estimates for the whole response variable distribution, not just the mean and standard deviation. This same methodology can also be applied to any other simulation problem. In fact, studying the project duration in a PERT network, we arrived at similar results as those reported here. Although not a common practice in simulation studies, the estimation of a response variable distribution is a valuable information for the decision maker. For instance, this is the case in a risk analysis, where a risk profile (the probability distribution of the net present value or the internal rate of return) is usually estimated. As such, Descriptive Sampling represents a highly promising improvement as far as risk analysis and other simulation studies are concerned.

**REFERENCES**

LAW, A.M.; KELTON, W.D. Simulation modeling and analysis. New York: McGraw-Hill, 1992.

RAMBERG, B.W.; SCHMEISER, J.S. An approximate method for generating symmetric random variables. Comm of the ACM, v.1,p.987-990, 1972.

SALIBY, E. Descriptive sampling: a better approach to Monte Carlo simulation. J. Opl Res. Soc., v.41, p.1133-1142, 1990.

________. Rethinking simulation: descriptive sampling. São Paulo: Atlas/EDUFRJ, 1989. (In Portuguese).

________.; PAUL, R.J. Implementing descriptive sampling in three-phase discrete event simulation models. J. Opl Res. Soc., v. 44, p.147-160, 1993.

ZAINO, N.A.; D'ERRICO, J. Optimal discrete approximations for continuous outcomes with applications in decision and risk analysis. J. Opl Res. Soc., v.40, p.379-388, 1989.