A major purpose of the Technical Information Center is to provide the broadest dissemination possible of information contained in DOE's Research and Development Reports to business, industry, the academic community, and federal, state and local governments.

Although a small portion of this report is not reproducible, it is being made available to expedite the availability of information on the research discussed herein.

1

NOTICE

PORTIONS OF THIS REPORT ARE ILLEGIBLE. It has been reproduced from the best available copy to permit the broadest possible availability.

LA-UR--84-1583

DE84 012438

Los Alamos National Laboratory is operated by the University of California for the United States Department of Energy under contract W-7405-ENG-36

TITLE MONTE CARLO ESTIMATION UNDER DIFFERENT DISTRIBUTIONS USING THE SAME SIMULATION

AUTHOR(S) R. J. Beckman, M. D. McKay

SUBMITTED TO Technometrics

Los Alamos Los Alamos National Laboratory Los Alamos, New Mexico 87545

FORM NO. 836 R4

MONTE CARLO ESTIMATION UNDER DIFFERENT DISTRIBUTIONS USING THE SAME SIMULATION

By

R. J. Beckman, M. D. McKay

Keywords: Computer Models, Sampling, Sensitivity Analysis, Importance Sampling

ABSTRACT

Two methods for reducing the computer time necessary to investigate changes in distribution of random inputs of large simulation computer codes are presented. The first method produces unbiased estimators of functions of the output variable under the new distribution of the inputs. The second method generates a subset of the original outputs which has a distribution corresponding to the new distribution of inputs. Efficiencies of the two methods are examined.

1. Introduction and Summary

Long running computer codes have been used in assessing the risks and benefits of such things nuclear power and hazardous waste disposal. Examples of these may be found in Dillon, Lantz and Pahwa (1978) Hirt and Romero (1975) and McKay, Conover and Beckman (1979). An additional study is found in the example given by Iman and Conover (1980), and such studies are suggested by Goodman and Koch (1982) and Levinson and Yeater (1983). Typically these codes require a number of input parameters whose value are not known with certainty because of lack of knowledge of the physical process being modeled or variations due to sampling distributions. This uncertainty is modeled by assigning probability distributions to the input parameters. The probability distributions, centered about nominal values in the parameter space and (in some way) reflecting knowledge the analyst has about the parameters, may represent either an estimate of the sampling distribution of estimators of the unknown parameters or a "degree of belief" in the values of the parameters.

When input parameters are considered random variables in computer codes, the output variables are also random variables, and the properties of the distributions of these variables are of interest. For example, one may wish to estimate the mean value of an output variable, or its distribution function These tasks are straight forward and can be carried out by the analyst with a minimum amount of work. Often, however, the analyst is faced with defending his choice of the distribution of the input variables and its influence on the conclusions he has drawn. Due to the mathematical complexity of the model it is usually impossible to analytically deduce the effects of changing the input distribution, and costs may make rerunning the code with new distributions prohibitive. The purpose of this paper is to give two methods which allow the analyst to change the distributions of the input variables without rerunning the computer code.

Let $\underset{\sim}{X}_1$, $\underset{\sim}{X}_2$, ... $\underset{\sim}{X}_n$ be independent vectors of input variables with density $f_1(\underset{\sim}{x})$, and let $Y_i = h(\underset{\sim}{X}_i)$ be the output variable for input $\underset{\sim}{X}_i$ where h represents the computer code. Also assume that $\theta_1 = E_{f_1}\{g(y)\}$ is the parameter to be estimated, and suppose that the analyst would like to study $h(\underset{\sim}{X})$ where $\underset{\sim}{X}$ comes from a different density $f_2(\underset{\sim}{X})$. Two methods for changing the distribution of the inputs are given. The first method which is a weighting scheme simular to importance sampling see (Kahn and Marshall (1953)) produces an unbiased estimator of θ_2, the expected value of $g(Y_i)$ when the $\underset{\sim}{X}_i$ are from f_2. The other method is a rejection method (see Kennedy and Gentle (1980)) which leaves the output variable Y with the density induced by letting X have density f_2.

In the first method (I), for each vector $\underset{\sim}{X}_i$, let $w_i = f_2(\underset{\sim}{X}_i) / f_1(\underset{\sim}{X}_i)$. It is shown in section 2 that the estimator $g_w = \Sigma w_i g(y_i)/n$ is an unbiased estimator of θ_2.

For the second method (II) a uniform upper bound, M, on the ratio $f_2(\underset{\sim}{X})/f_1(\underset{\sim}{X})$ is assumed to exist. Let the random variable V given the vector $\underset{\sim}{X}_i = x$ have a uniform distribution between 0 and $Mf_1(x)$. For each i, retain the pair $(\underset{\sim}{X}_i, Y_i)$ in the sample if the realization v of V is less than $f_2(X_i)$. In section 3 this rejection scheme is shown to leave the selected vectors $\underset{\sim}{X}$ with density $f_2(\underset{\sim}{X})$, and hence the selected values of Y with the desired density.

The efficiencies of methods I and II are examined in sections 2 and 3, respectively. The results given in these two sections demonstrate the necessity of understanding the methods before their application. The efficiency of each method is shown to decrease rapidly as large differences occur between the densities $f_1(\underset{\sim}{X})$ and $f_2(\underset{\sim}{X})$. However, the efficiency of method I may be larger than 1.0 for small differences between the two densities. An example of the simulation of system unavailabilities derived from fault trees is given in section 4.

2. Method I: The Weighting Method

Letting $g_w(\underset{\sim}{Y}) = \Sigma w_i g(Y_i)/n$ where $w_i = f_2(\underset{\sim}{X}_i)/f_1(\underset{\sim}{X}_i)$, and $\theta_2 = E_{f_2}\{\Sigma g(Y_i)/n\}$, we have $E_{f_1}\{g_w(\underset{\sim}{Y})\}$

$$= \Sigma\ E_{f_1}\{f_2(X_i)/f_1(X_i)g(Y_i)\}/n$$

$$\vdots$$

$$= \Sigma \int g(Y_i)\ f_2(\underset{\sim}{X}_i)d\underset{\sim}{X}_i/n = \theta_2.$$

Therefore, the estimator $g_w(\underset{\sim}{Y})$ is an unbiased estimator of θ_2. Although this method closely resembles importance sampling, its intent is different. In importance sampling the goal is to obtain an unbaised estimator of θ_2 with a smaller variance than would be attained by sampling from f_2. In the weighting method presented here, the goal is only to obtain an unbiased estimator of θ_2. Since f_1 and f_2 are fixed, smaller variances can not be the aim.

The efficiency of $g_w(\underset{\sim}{Y})$ is measured by its variance relative to the variance of $\Sigma g(Y_i)/n$ when the corresponding values of X come from f_2. To assist in the study of the efficiency of this method, we let $H(\underset{\sim}{x}) = g(h(\underset{\sim}{x}))$ and assume that the expected values of H and H^2 with respect to densities indexed by a parameter θ may be expressed as $E_\theta\{H^2(x)\} = \phi(\theta)$ and $E_\theta\{H^2(x)\} = \psi(\theta)$. Then with $w = f(\underset{\sim}{x}_i;\theta_2)/f(\underset{\sim}{x}_i;\theta_1)$ the variance of the weighted estimator $wH(\underset{\sim}{x})$ is $E_{\theta_2}\{wH^2(x)\} - \phi^2(\theta_2)$.

For many of the common densities, $E_{\theta_2}\{wH^2(x)\}$ is expressible as a function of $\psi(\theta^*)$ for some θ^* in the parameter space of θ. Table 1 contains the expected value under f_2 of $wH^2(x)$. These values may be used to obtain the efficiency of the procedure. For example, suppose f_1 is normal with mean 0 and variance 1, and a change to a normal variate with mean 0 and variance 1/2 is desired. Suppose also that $H(x) = x$ so that $\phi(\mu,\sigma^2) = \mu$ and $\psi(\mu,\sigma^2) = \sigma^2 + \mu^2$. Then $E_{\theta_2}\{wH^2(x)\} = .385$, and the efficiency is 77%. While the func ion H(x) would rarely be x, the above exercise does give some indication of the loss of efficiency using this method.

Figures 1 and 2 show the $\log_{10}$ efficiency of method I for normal samples with changes in the means and variances. Once again H(x) is assumed to be x, and the efficiency is measured as the ratio of the variance of the estin ıtor using this method to that of random sampling.

In figure 1 f_1 is taken as a normal with mean 2 and variance 1, while f_2 is normal with mean ranging between .5 and 3 and variance 1. From this figure there is a range of values of the second mean from approximately 1.1 to 2.0, where the efficiency of method I is greater than 1. This is a common occurrence with this method and is not great surprise given the method's close resemblance to importance sampling. The same phenomenon is demonstrated in figure 2 where the means of the densities are both 2.0, the variance of f_1 is 1, and the second variance ranges between .5 and 1.5. In this case the efficiency is greater than 1 when the second variance is between .8 and 1.0. It should be noted that in both cases outside of these ranges the efficiency falls off rapidly.

3. Method II: The Rejection Method

Assume that there exists a uniform bound M such that $f_2(\underset{\sim}{x})/f_1(\underset{\sim}{x}) \leq M$ for all $\underset{\sim}{x}$. Let the random variable V given $\underset{\sim}{X} = \underset{\sim}{x}$ be uniform between 0 and

$Mf_1(\underset{\sim}{x})$. The value $\underset{\sim}{X} = \underset{\sim}{x}$ is accepted as a sample from density $f_2(\underset{\sim}{x})$ if the realization v of V is less than $f_2(x)$. It follows that an arbitrary $\underset{\sim}{x}$ which is selected to remain in the sample has density f_2 since

$\Pr\{\underset{\sim}{X}<\underset{\sim}{x} \mid \underset{\sim}{x} \text{ remains in the sample}\}$

$$= \Pr\{\underset{\sim}{X}<\underset{\sim}{x} \text{ and } \underset{\sim}{x} \text{ remains}\} \,/\, \Pr\{\underset{\sim}{X} \text{ remains}\}$$

$$= \int_{-\infty}^{\underset{\sim}{x}} \Pr\{V<f_2(\underset{\sim}{u}) \mid \underset{\sim}{u}\} d\underset{\sim}{u} \,/\, \int_{-\infty}^{\infty} \Pr\{V<f_2(\underset{\sim}{u}) \mid \underset{\sim}{u}\} f_1(\underset{\sim}{u}) d\underset{\sim}{u}$$

$$= M \int_{-\infty}^{\underset{\sim}{x}} \frac{f_2(\underset{\sim}{u}) f_1(\underset{\sim}{u})}{M\, f_1(\underset{\sim}{u})}\, du = F_2(\underset{\sim}{x}).$$

In random variate generation with the rejection method the analyst chooses f_1 to efficiently generate samples from f_2. In the cases presented here the analyst is not free to pick either f_1 or f_2 as they are assumed known and fixed.

The efficiency of method II can be measured by the probability that a random $\underset{\sim}{x}$ from f_1 is accepted for f_2. This probability is given by the reciprocal of the bound M, since

$\Pr\{\text{ a random } \underset{\sim}{X} \text{ is selected }\} =$

$$= \int_{-\infty}^{\infty} \Pr\{V<f_2(\underset{\sim}{u}) \mid \underset{\sim}{u}\} f_1(\underset{\sim}{u}) du$$

$$= \int_{-\infty}^{\infty} \frac{f_2(\underset{\sim}{u}) f_1(\underset{\sim}{u}) d\underset{\sim}{u}}{M\, f_1(\underset{\sim}{u})} = 1/M \quad .$$

Table 2 contains the bounds M for most of the densities commonly used in simulation studies. It is interesting that M does not exist in some cases. For example if $\sigma_1^2 = \sigma_2^2$ for normal densities, then the bound M exists only for the trivial case $\mu_1 = \mu_2$, while for $\sigma_2^2 = \sigma_1^2 - \varepsilon$ the bound does exist for any values of μ_1, μ_2 and $\varepsilon > 0$.

Figures 3, 4 and 5 show the $\log_{10}$ efficiencies of method II for normal samples and h (x) = x. In figure 3, $\mu_1 = 2$, $\sigma_1^2 = 1$, $\sigma_1^2 = .9$ and μ_2 varies from .5 to 3. Over this range the efficiency covers five orders of magnitude. This points out the impracticality of changing the mean of a normal more than a few tenths of its standard deviation if only small changes in the variance are desired. On the other hand, if larger changes in the variance are made the efficiency does not drop off as rapidly. This is demonstrated in figure 4 where σ_2^2 is fixed at .5. For this case the efficiency is lower at $\mu_2 = 2$ but attains an efficiency of a little less than one order of magnitude at $\mu_2 = .5$.

The loss of efficiency for normal samples with changes in the variance is given in figure 5, where $\mu_1 = \mu_2 = 2$, $\sigma^2 = 1$ and σ_2^2 ranges from .1 to 1. The efficiency loss for this example is one and one-half orders of magnitude at the extreme value $\sigma_2^2 = .1$.

The use of method II can lead to low efficiencies. This is particularly true for large changes in the density functions. However, for reasonable shifts in the density (e.g $\mu_1 = \mu_2 = 2$, $\sigma_1^2 = 1$, $\sigma_2^2 = .9$) the efficiency remains high. For these types of changes the lower efficiency of the method is minor compaired to cost of rerunning the computer code.

4. Example

We give as an example of the uses of the generation schemes presented here a two-out-of-three voting system given in figure 6. (See Henley and Kumamoto 1981) In this system three independent monitors shut the system down if any two of the three signal for a shutdown. The system unavailability for this arrangememt of monitors is given by $P_s = P_1P_2 + P_2P_3 + P_1P_3 - 2P_1P_2P_3$, where P_i is the probability that monitor i signals for a system shutdown. We assume that the analyst desires to study the variabliity of the system unavailability using simulation techniques.

While the methods given here are generally for use in longer running simulation codes than the one implemented, the example was chosen to keep the model mathematically tractable and easy to elucidate. Examples of system unavailability dependent on 13 to 259 components may be found in Martz et al. (1983).

If we assume 10, 15 and 20 obser才erations on the three monitors with 2, 1, and 3 failures respectively, then the system unavailability is estimated by $\hat{P}_s = .049$. Uncertainty in this estimate comes from two sources. First, there is the sampling distribution of the number of monitor failures which is assumed to be binomial. This source of variation was estimated by a simulation study of 10000 observations of $\tilde{P}_s = \tilde{P}_1\tilde{P}_2 + \tilde{P}_2\tilde{P}_3 + \tilde{P}_3\tilde{P}_1 - 2\tilde{P}_1\tilde{P}_2\tilde{P}_3$, and $\tilde{P}_i = X_i/N_i$, and X_i was generated from a binomial $(\hat{P}_i, N_i)$, where $\hat{P}_i$ = 2/10, 1/15, 3/20 for i = 1, 2, 3. A second source of variation in these types of studies is the uncertainty in the values of $\hat{P}_i$, the observed failure rates of the monitors, which are used to generate the binomial samples X_i. As the value of $\hat{P}_i$ changes, the sampling distribution of $\hat{P}_s$ changes, and it is this second source of variability on which the techniques given here attempt to treat.

The estimated mean unavailability is given in Figure 7 for $\hat{P}_1$ not only equal to .20 but also for $\hat{P}_1$ in the range .10 to .30 in steps of .001. These values were generated using the techniques of method I. While it is obvious that the expected unavailability is linear in P_1 for this case, it is interesting to note the observed linear relationship for the estimate $\tilde{P}_s$ as $\hat{P}_1$ moves away from .2. This can be explained by figure 8 which gives the efficiency of the technique for this example as a function of $\hat{P}_1$. For all values of $\hat{P}_1$ between .1 and .2 the efficiency is greater than one. For $\hat{P}_1$ from .2 to .3 the efficiency is less than one but exceeds .12. This leaves the estimate with an effective sample size of more than 1200 (effective sample size = efficiency $\cdot$ sample size) for all parameter values studied here.

Using the same set of 10000 generated values of $\tilde{P}_s$, $\tilde{P}_1$, $\tilde{P}_2$, $\tilde{P}_3$, the second method was used to estimate the 10^{th}, 25^{th}, 50^{th}, 75^{th}, and 90^{th} percentiles of the sampling distribution of $\tilde{P}_s$ for values of $\hat{P}_1$ ranging from .10 to .30. These are given in figure 9, where the estimated percentiles become ragged for $\hat{P}_1$ greater than .28. This is caused by a lack of efficiency which is evident from figure 10. Since the efficiency drops off at a slower rate as the value of $\hat{P}_1$ decreases, one would probably generate the largest value of $\hat{P}_1$ of interest, and using method II estimate percentiles of $\tilde{P}_s$ for smaller values of $\hat{P}_1$.

The computer code used to generate the 10000 values of $\tilde{P}_s$ was by no means long-running. The 10000 observations were generated in 24 seconds on a VAX 11/780. While the techniques presented here are designed to be effective in codes where the majority of the computer time is spent evaluating the function (P_s in our example), they did show improved efficiency in producing the data points of figures 7 and 9. Compared to the 24 seconds used to obtain the data point for $P_1 = .2$, the 200 data points of figure 7 were generated in 31 seconds, or .16 seconds per point while the 200 data points of figure 9 were obtained in 123 seconds or .62 seconds per point.

For computer codes in which a large fraction of time is used to calculate the function, these savings would be even more dramatic.

5. Conclusions

Two methods were given to reduce the computer time necessary to investigate changes in the distributions of the random inputs to large simulation computer codes. The first method produced unbiased estimates of functions, g(y) of the output variables Y. In the second method a subset of the random outcomes $(\underset{\sim}{X}_i, Y_i)$ were selected so that the $\underset{\sim}{X}_i$ have the desired distribution. Efficiencies of these methods were investigated, and an example showed the potential of these techniques to save large amounts of computer time.

References

Dillon, R. T., Lantz, R. B. and Pahwa, S. B. (1978). Risk methodology for geologic disposal of radioactive waste: The Sandia Waste Isolation Flow and Transport (SWIFT) Model. SAND 78-1267, Sandia Laboratories, Albuquerque, NM.

Goodman, J. and Koch, J. E. (1982). The probability of a tornado missile hitting a target. Nuc. Eng. and Design, 75, 125-155.

Hirt, C. W. and Romero, N. C. (1975). Application of a drift-flux model to flashing in straight pipes. Los Alamos National Laboratory Report LA-6005-MS, Los Alamos, NM.

Iman R. L. and Conover W. J. (1980), Small sample sensitivity analysis techniques for computer models, with an application to risk assessment. Communication in Statistics, A9, 1749-1874.

Kahn, H. and Marshall, A. W. (1953)." Methods of reducing sample size in Monte Carlo computations," J. Oper. Res. Soc. Amer., 1, 263-271.

Kennedy, W. J. Jr. and Gentle, J. E. (1980). Statistical Computing, Marcel Dekker Inc., New York.

Levinson, S. H. and Yeater (1983). "Methodology to evaluate the effectiveness of fine protection systems on nuclear power plants", Nuc. Eng. and Design, 76, 161-182.

Martz, H. F. Beckman, R. J., Campbell, K., Whiteman, D. E. and Booker, J. M. (1983). "A comparison of methods for uncertainty analysis of nuclear power plant safety system fault tree models." Los Alamos National Laboratory Report, LA-9729-MS. Los Alamos, N.M.

McKay, M. D., Conover, W. J. and Beckman, R. J. (1979), "A Comparison of three methods for selecting values of input variables in the analysis of output from a computer code," Technometrics, 21, 239-245.

Henley, E. and Kumamota, H. (1981). Reliability Engineering and Risk Assessment, Prentice-Hall, New Jersey.

TABLE 1

Family	$E_{\theta_2}(\psi H^2(\underline{x}))$	Restrictions and Comments
Normal $N(\mu, \sigma^2)$	$\exp((\mu_2 - \mu_1)^2/(2\sigma_1^2 - \sigma_2^2))\ \sigma_1^2/[\sigma_2^2(2\sigma_1^2 - \sigma_2^2)]^{1/2}$ X $\psi((2\sigma_1^2\mu_2 - \sigma_2^2\mu_1)/(2\sigma_1^2 - \sigma_2^2),\ \sigma_1^2\sigma_2^2/(2\sigma_1^2 - \sigma_2^2))$	$E(H^2) = \psi(\mu, \sigma^2)$. The expectation does not exist in this form for $\sigma_2^2 > 2\sigma_1^2$
Beta (α, β)	$[B(\alpha_1, \beta_1)\ B(2\alpha_2 - \alpha_1, 2\beta_2 - \beta_1)/B^2(\alpha_2, \beta_2)]$ X $\psi(2\alpha_2 - \alpha_1, 2\beta_2 - \beta_1)$	$E(H^2) = \psi(\alpha, \beta)$. The expectation does not exist in this form for $\alpha_1 > 2\alpha_2$ or $\beta_1 > 2\beta_2$.
Gamma [illegible]	$\Gamma(\alpha_1)\ \beta_1^{\alpha_1}\Gamma(2\alpha_2 - \alpha_1)\ [\beta_2\beta_1/(2\beta_1 - \beta_2)]^{2\alpha_2 - \alpha_1}\ (\Gamma^2(\alpha_2)\beta_2^{2\alpha_2})$ X $\psi(2\alpha_2 - \alpha_1, \beta_1\beta_2/(2\beta_1 - \beta_2))$	$E(H^2) = \psi(\alpha, \beta)$. The expectation does not exist in this form for $\alpha_1 > 2\alpha_2$ or $\beta_2 > 2\beta_1$.
Lognormal [illegible] $N(\mu, \sigma^2)$	Same as that of a normal	$E(H^2) = \psi(\mu, \sigma^2)$. While the functional form of $E_{\theta_2}(\psi H^2)$ is the same as that a normal, the functions are different.
Poisson [illegible]	$\exp((\lambda_2 - \lambda_1)^2/\lambda_1)\psi(\lambda_2^2/\lambda_1)$	$E(H^2) = \psi(\lambda)$
Binomial $B(N, p)$	$((p_2^2 - 2p_1p_2 + p_1)/((1 - p_1)p_1))^N\ \psi(p_2^2(1 - p_1)/(p_2^2 - 2p_1p_2 + p_1))$	$E_1(H^2(x)) = \psi(p)$, $N_1 = N_2 = N$
Exponential Family $c(\theta)B(x)\exp(\Sigma\theta_j T_j(\underline{x}))$	$C^2(\theta_2)/[C(\theta_1)C(2\theta_2 - \theta_1)] \cdot \psi(2\theta_2 - \theta_1)$	$E_{\underline{\theta}}(H^2(x)) = \psi(\theta)$. This form exists only when $C(2\theta_2 - \theta_1)$ is defined.

TABLE 2

Family	M	Restrictions and Comments
Normal or Lognormal X or $\ln X \sim N(\mu, \sigma^2)$	$(\sigma_1/\sigma_2)\exp(.5(\mu_2 - \mu_1)^2/(\sigma_1^2 - \sigma_2^2))$	The bound only exists for $\sigma_1^2 > \sigma_2^2$
Beta $B(\alpha, \beta)$	$CX_0^{(\alpha_2-\alpha_1)}(1 - X_0)^{\beta_2-\beta_1}$ where $C = \frac{\Gamma(\alpha_2 + \beta_2)\Gamma(\alpha_1)\Gamma(\beta_1)}{\Gamma(\alpha_1 + \beta_1)\Gamma(\alpha_2)\Gamma(\beta_2)}$, $X_0 = (\alpha_2 - \alpha_1)/(\alpha_2 - \alpha_1 + \beta_2 - \beta_1)$	This bound holds for a) $\alpha_2 > \alpha_1$, $\beta_2 > \beta_1$ b) $\alpha_2 < \alpha_1$, $\beta_2 > \beta_1$ and $(\alpha_1 - \alpha_2) > (\beta_2 - \beta_1)$
	C	a) $\alpha_1 = \alpha_2$, $\beta_2 > \beta_1$ b) $\alpha_2 > \alpha_1$, $\beta_2 = \beta_1$
Gamma (α, β) $\beta^\alpha x^{\alpha-1}\exp(-x\beta)/\Gamma(\alpha)$	$CX_0^{\alpha_2-\alpha_1}\exp((\beta_1 - \beta_2)X_0)$, where $C = \Gamma(\alpha_2)\beta_2^{\alpha_2}/\Gamma(\alpha_1)\beta_1^{\alpha_1}$ and $X_0 = (\alpha_2 - \alpha_1)/(\beta_2 - \beta_1)$	for $\alpha_2 > \alpha_1$ and $\beta_2 > \beta_1$
	$(\beta_2/\beta_1)^\alpha$	for $\alpha = \alpha_1 = \alpha_2$, $\beta_2 > \beta_1$
Poisson	$\exp(\lambda_1 - \lambda_2)$	for $\lambda_2 < \lambda_1$
Binomial (N, p)	$(1 - p)^{N_2-N_1}$	for $p_1 = p_2 = p$, $N_2 < N_1$
	$(1 - p_2)^{N_2}/(1 - p_1)^{N_1}$	for $p_1 > p_2$, $N_2 < N_1$
	No closed form solution	for $p_2 > p_1$, $N_2 < N_1$

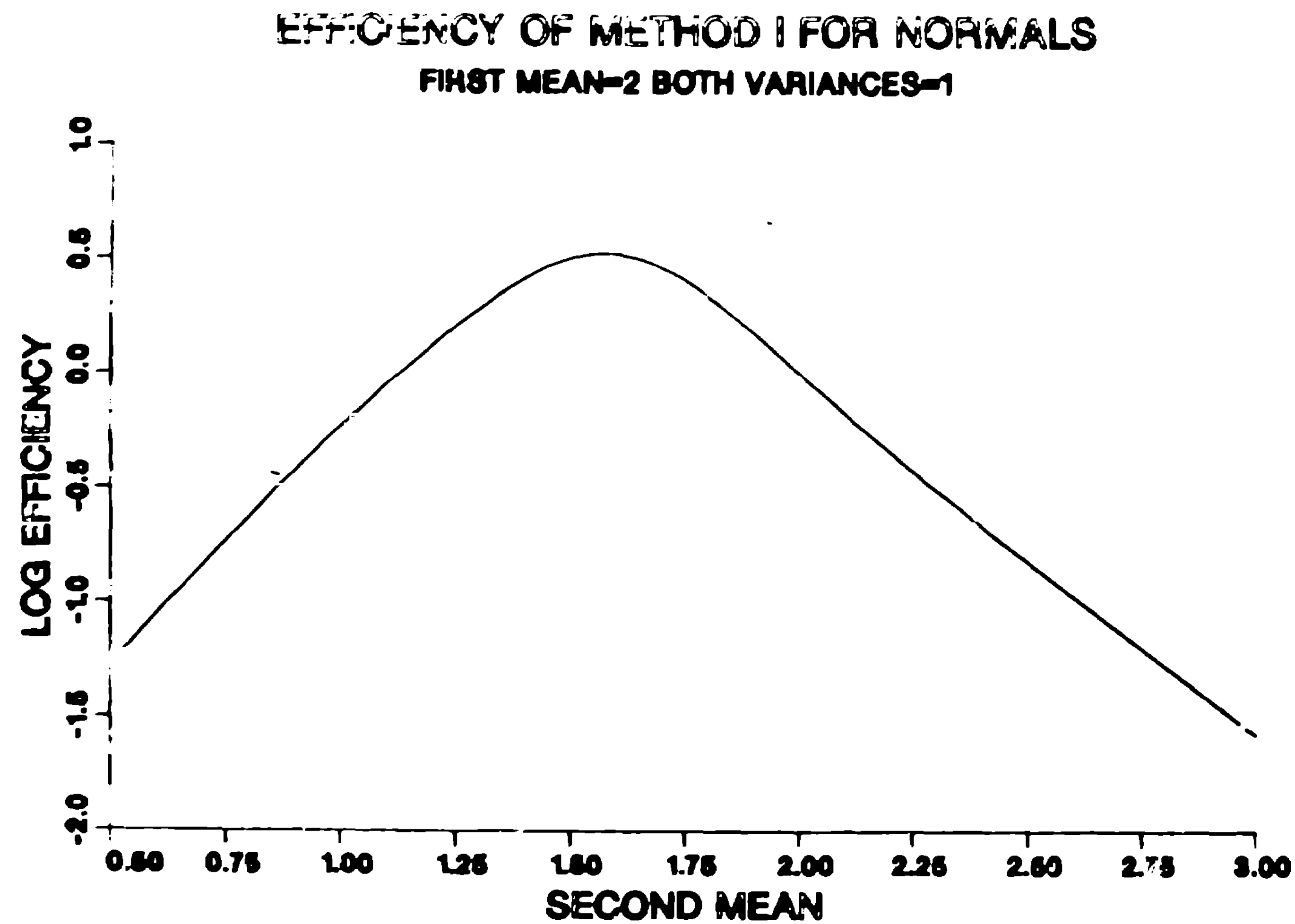

EFFICIENCY OF METHOD I FOR NORMALS
FIRST MEAN=2 BOTH VARIANCES=1
LOG EFFICIENCY
1.0
0.5
0.0
-0.5
-1.0
-1.5
-2.0
0.50
0.75
1.00
1.25
1.50
1.75
2.00
2.25
2.50
2.75
3.00
SECOND MEAN

Figure 1

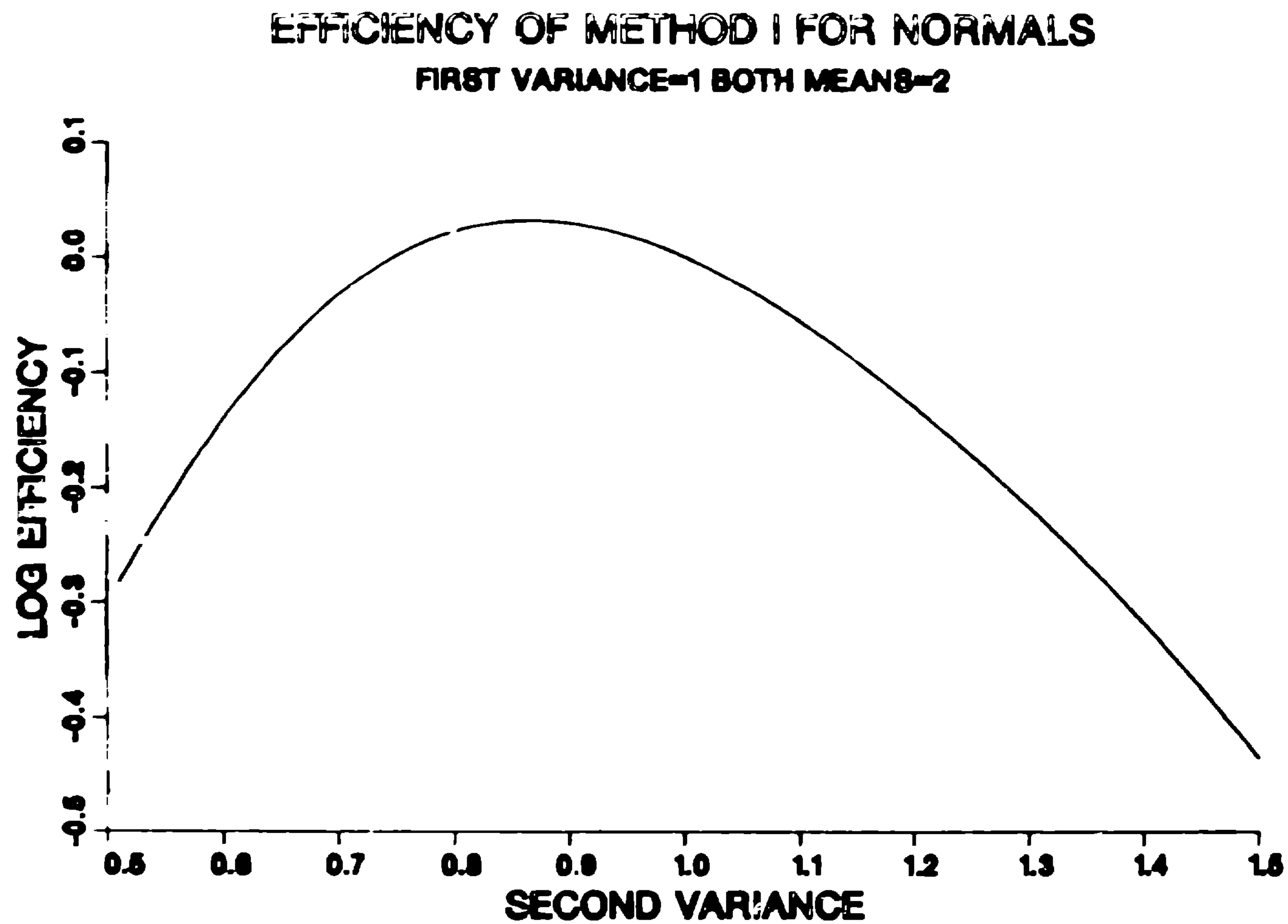

EFFICIENCY OF METHOD I FOR NORMALS
FIRST VARIANCE=1 BOTH MEANS=2
LOG EFFICIENCY
0.1
0.0
-0.1
-0.2
-0.3
-0.4
-0.5
0.5
0.6
0.7
0.8
0.9
1.0
1.1
1.2
1.3
1.4
1.5
SECOND VARIANCE

Figure 2

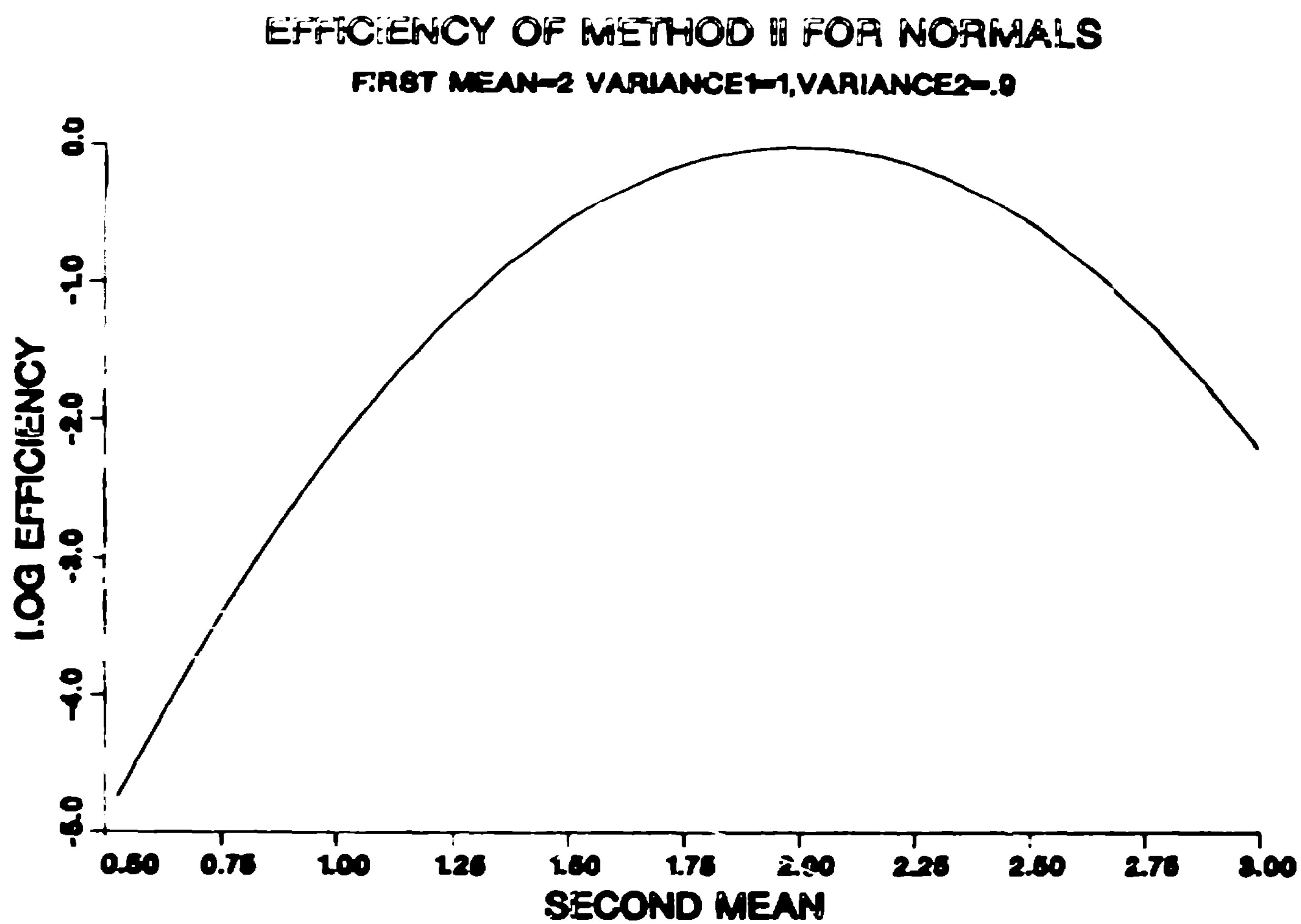

EFFICIENCY OF METHOD II FOR NORMALS
FIRST MEAN=2 VARIANCE1=1,VARIANCE2=.9
LOG EFFICIENCY
0.0
-1.0
-2.0
-3.0
-4.0
-5.0
0.50
0.75
1.00
1.25
1.50
1.75
2.00
2.25
2.50
2.75
3.00
SECOND MEAN

Figure 3

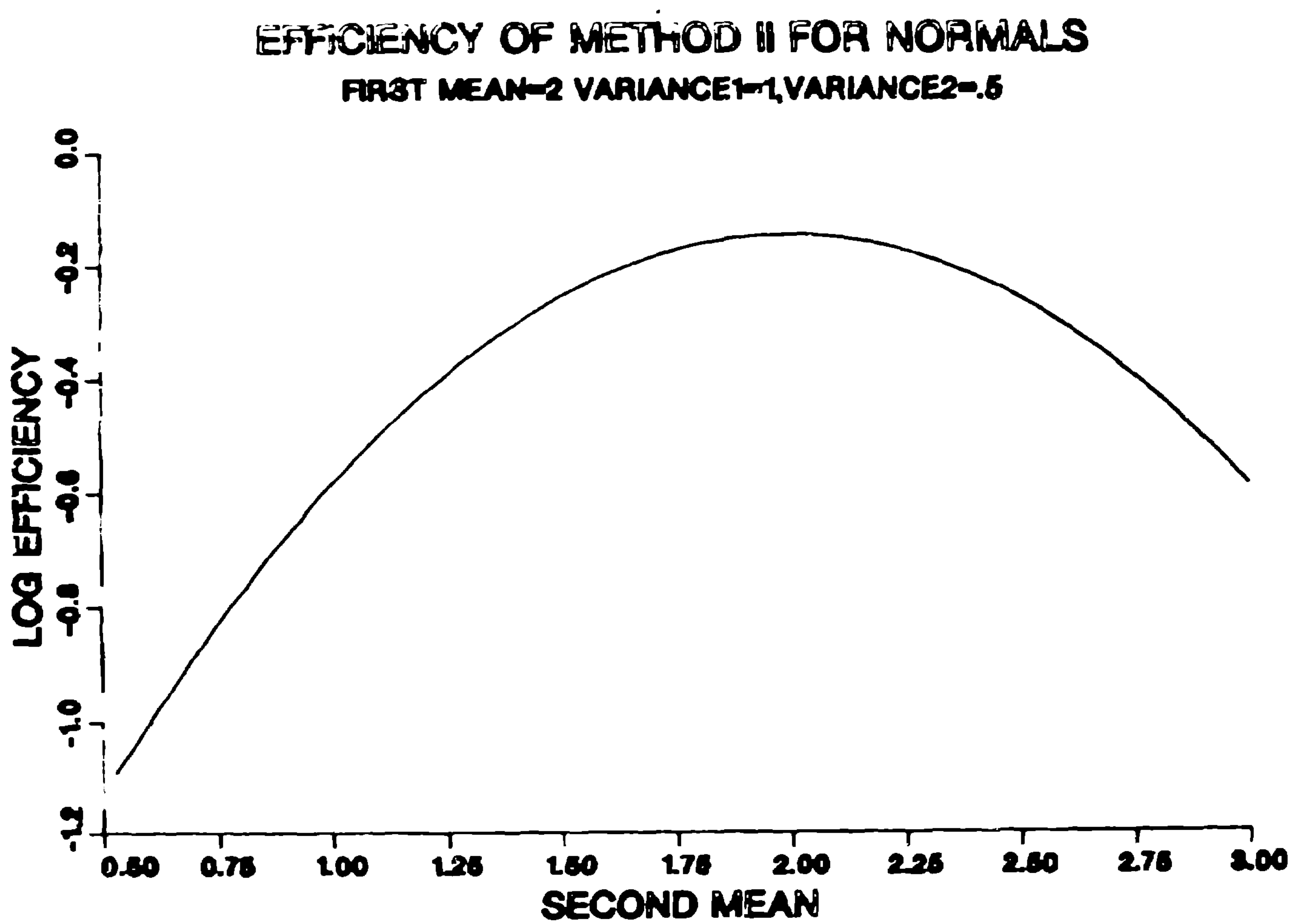

EFFICIENCY OF METHOD II FOR NORMALS
FIRST MEAN=2 VARIANCE1=1,VARIANCE2=.5
LOG EFFICIENCY
0.0
-0.2
-0.4
-0.6
-0.8
-1.0
-1.2
0.50
0.75
1.00
1.25
1.50
1.75
2.00
2.25
2.50
2.75
3.00
SECOND MEAN

Figure 4

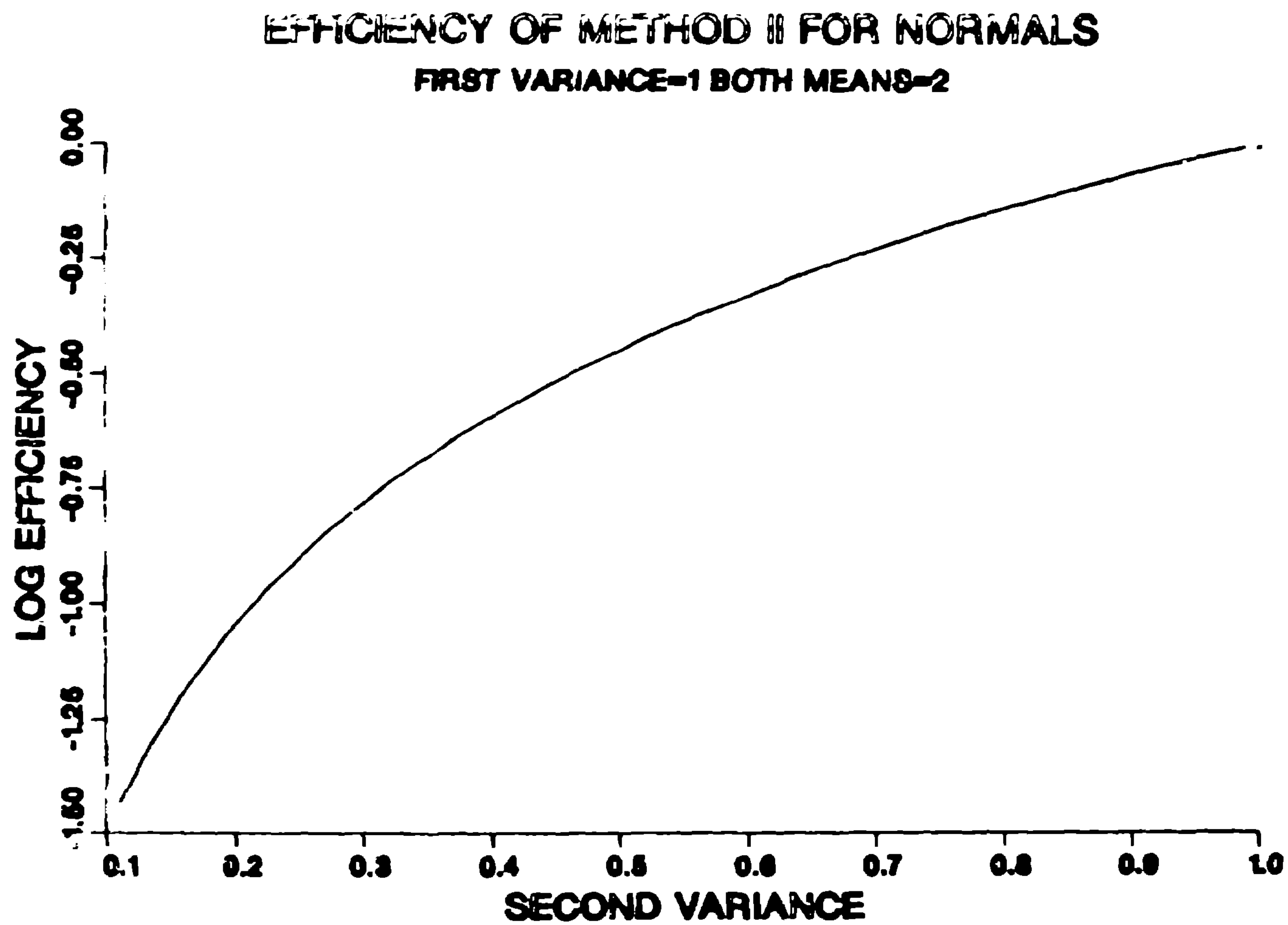
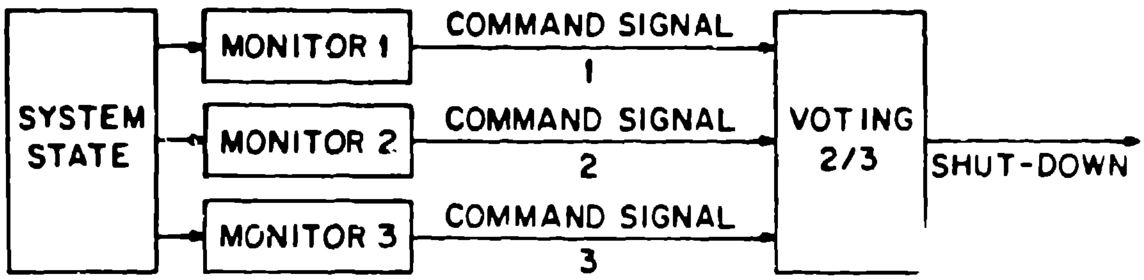
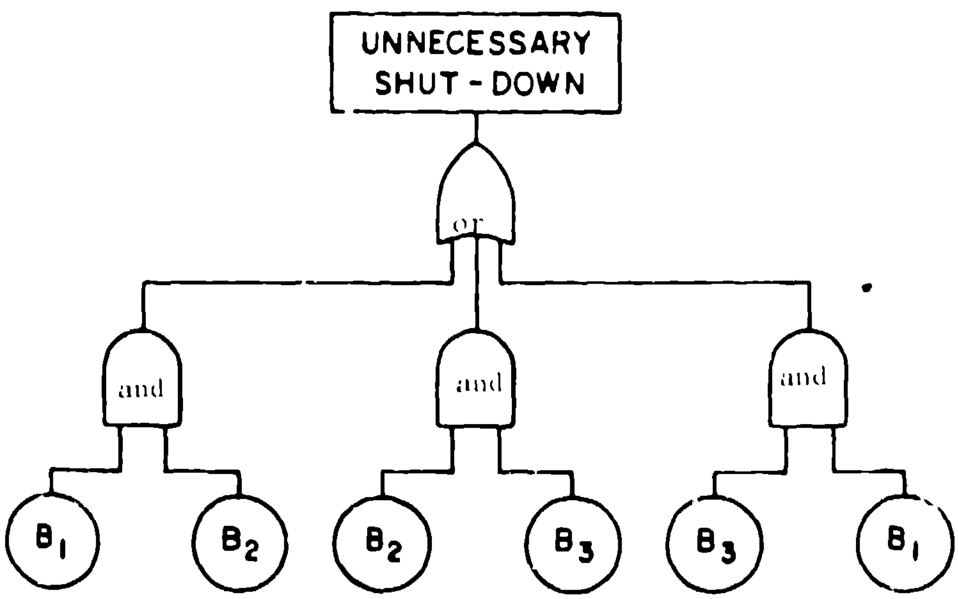

EFFICIENCY OF METHOD II FOR NORMALS
FIRST VARIANCE=1 BOTH MEANS=2
LOG EFFICIENCY
0.00
-0.25
-0.50
-0.75
-1.00
-1.25
-1.50
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
SECOND VARIANCE

Figure 5



Two-out-of-three voting system (from Henley and Kumamoto).



Fault tree for two-out-of-three voting system.

Figure 6

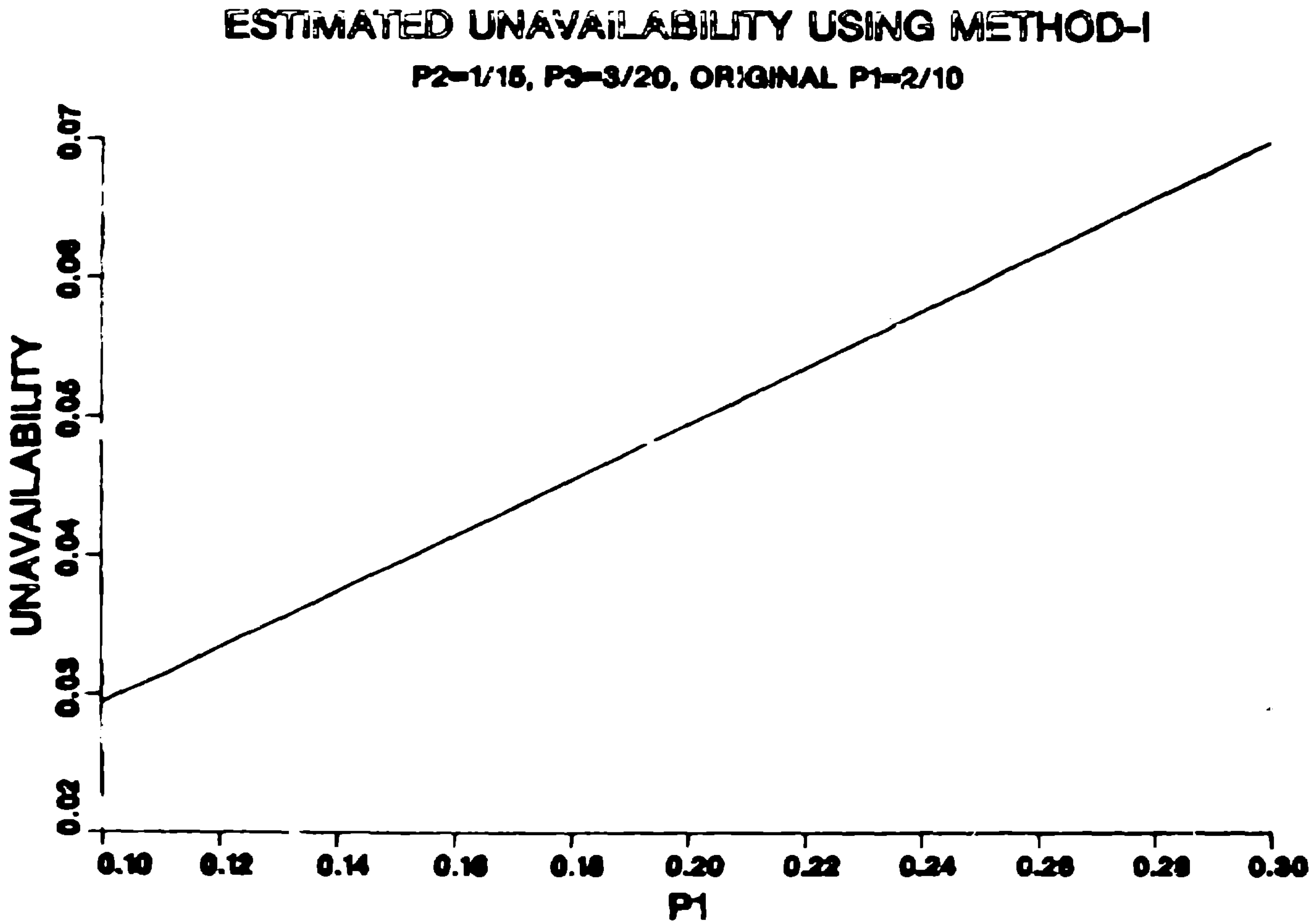

ESTIMATED UNAVAILABILITY USING METHOD-I
P2=1/15, P3=3/20, ORIGINAL P1=2/10
UNAVAILABILITY
0.02
0.03
0.04
0.05
0.06
0.07
0.10
0.12
0.14
0.16
0.18
0.20
0.22
0.24
0.26
0.28
0.30
P1

Figure 7

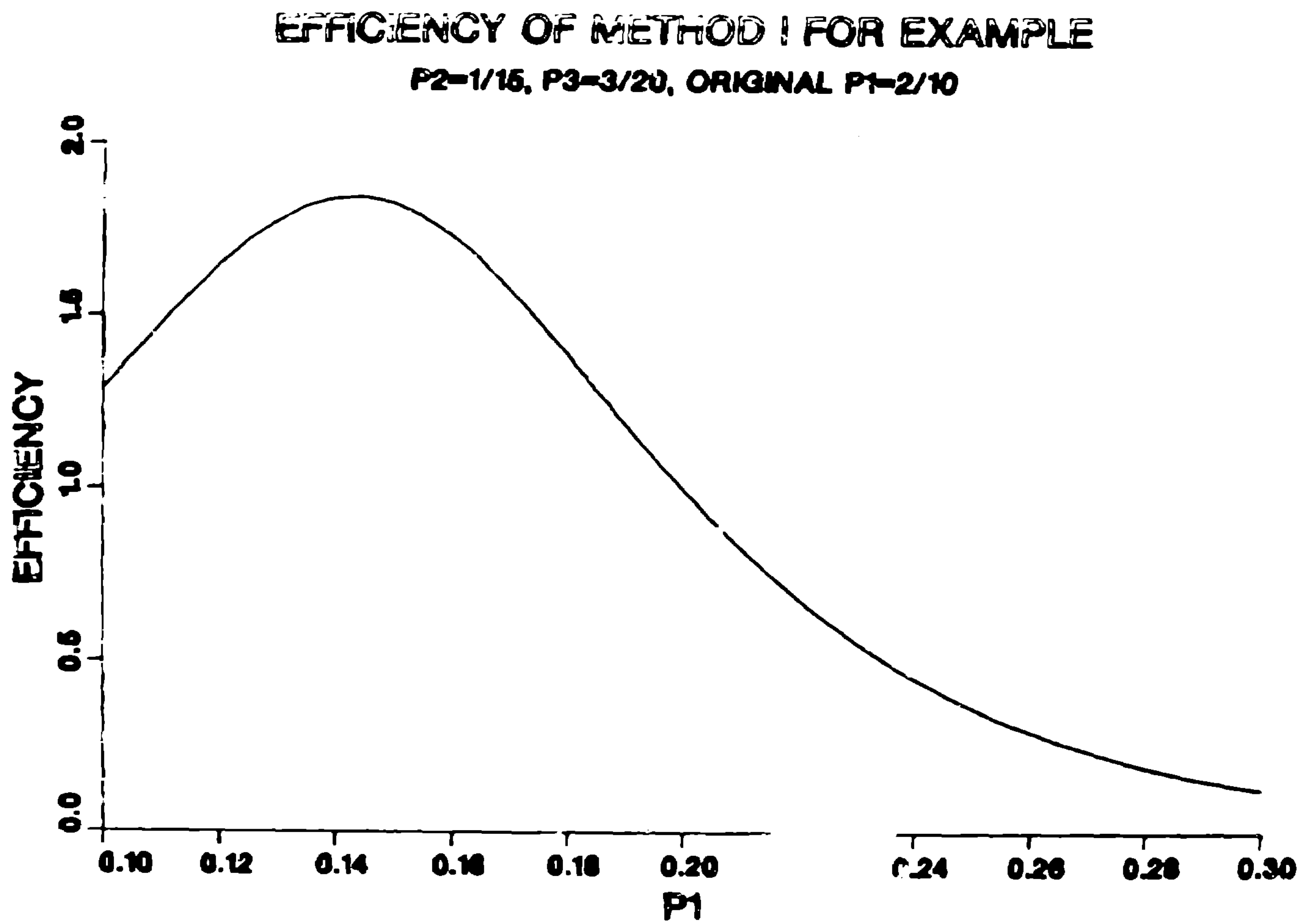

EFFICIENCY OF METHOD I FOR EXAMPLE
P2=1/15, P3=3/20, ORIGINAL P1=2/10
EFFICIENCY
0.0
0.5
1.0
1.5
2.0
0.10
0.12
0.14
0.16
0.18
0.20
0.24
0.26
0.28
0.30
P1

Figure 8

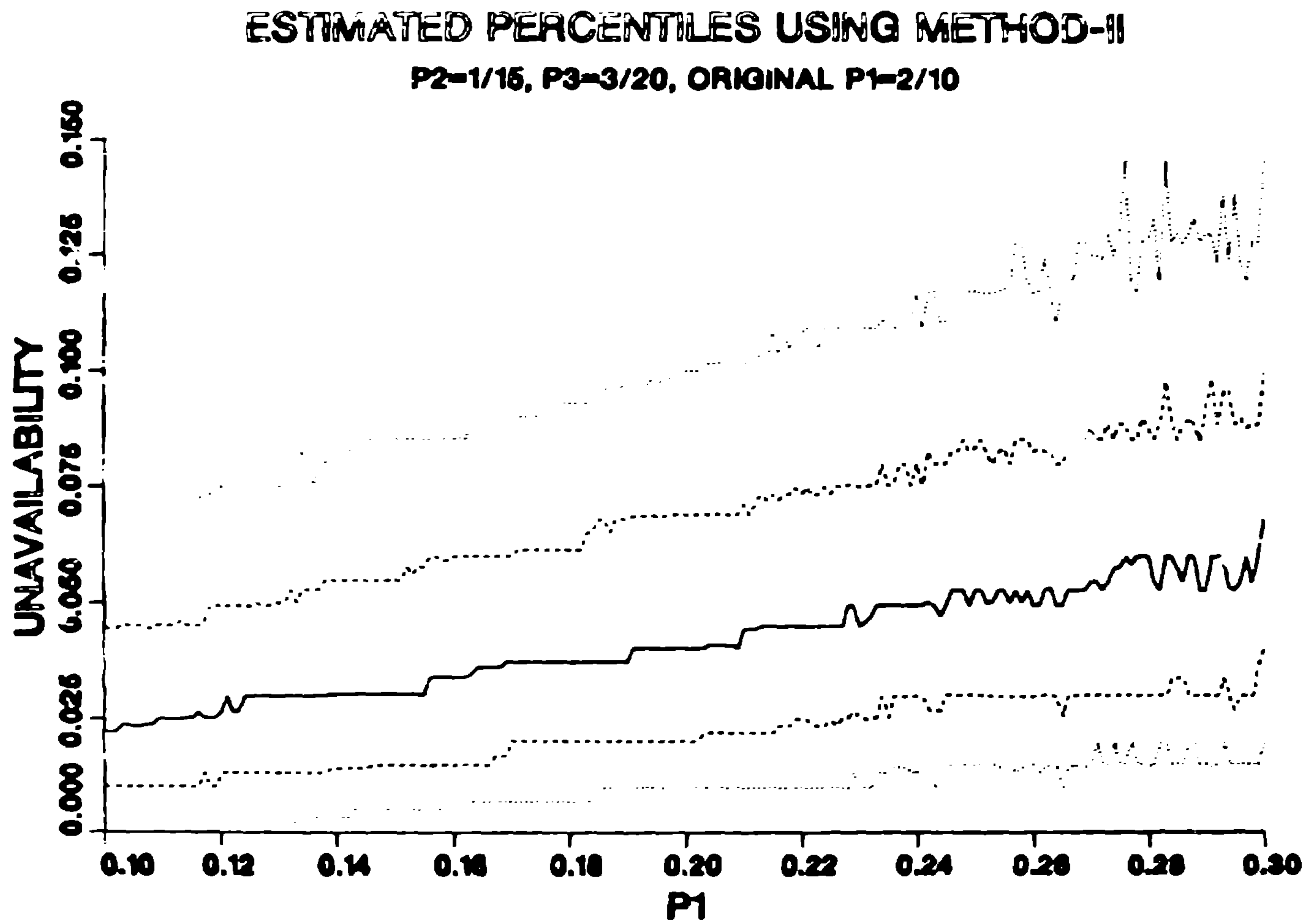

ESTIMATED PERCENTILES USING METHOD-II
P2=1/15, P3=3/20, ORIGINAL P1=2/10
UNAVAILABILITY
0.000
0.025
0.050
0.075
0.100
0.125
0.150
0.10
0.12
0.14
0.16
0.18
0.20
0.22
0.24
0.26
0.28
0.30
P1

Figure 9

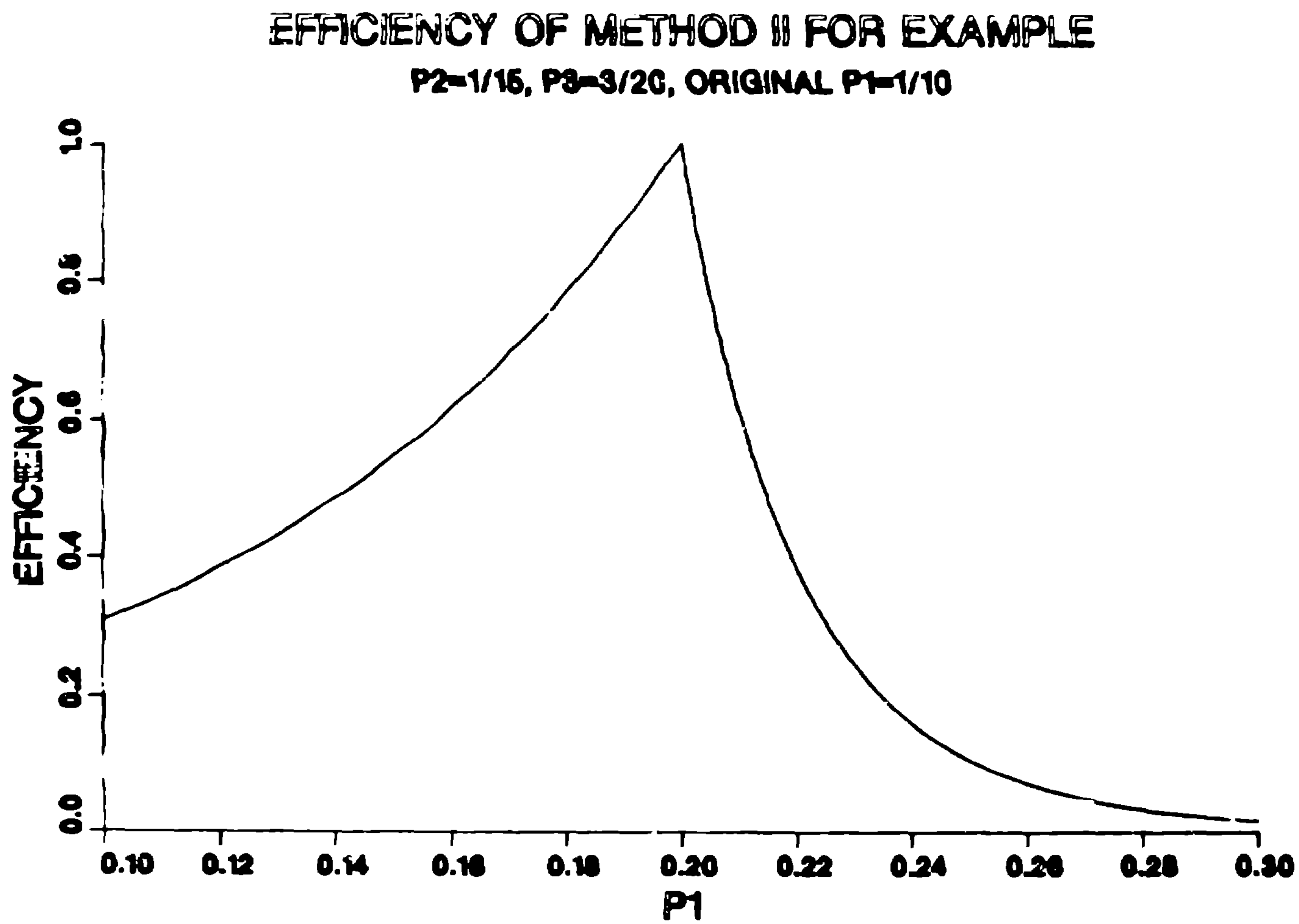

EFFICIENCY OF METHOD II FOR EXAMPLE
P2=1/15, P3=3/20, ORIGINAL P1=1/10
EFFICIENCY
0.0
0.2
0.4
0.6
0.8
1.0
0.10
0.12
0.14
0.16
0.18
0.20
0.22
0.24
0.26
0.28
0.30
P1

Figure 10