Battelle

Pacific Northwest Laboratories
Richland, Washington 99352

AEC Research and Development Report

OPS: RISK ANALYSIS IN CAPITAL
INVESTMENT AND DEVELOPMENT
OPPORTUNITIES

MAY 1972

NOTICE

This report was prepared as an account of work sponsored by the United States Government. Neither the United States nor the United States Atomic Energy Commission, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately-owned rights.

PACIFIC NORTHWEST LABORATORY

operated by

BATTELLE

for the

U.S. ATOMIC ENERGY COMMISSION

Under Contract AT(45-1)-1830

Printed in the United States of America

Available from

National Technical Information Service

U.S. Department of Commerce

5285 Port Royal Road

Springfield, Virginia 22151

Price: Printed Copy \$3.00; Microfiche \$0.95

OPS: RISK ANALYSIS IN CAPITAL INVESTMENT AND DEVELOPMENT OPPORTUNITIES

R. W. Legan
R. L. Engel

Systems Engineering Department

May, 1972

BATTELLE
PACIFIC NORTHWEST LABORATORIES
RICHLAND, WASHINGTON 99352

TABLE OF CONTENTS

	<u>Page</u>
SUMMARY	iii
LIST OF FIGURES	iv
LIST OF TABLES	v
INTRODUCTION	1
DISCUSSION	4
Alternative Analytical Methods.	4
OPS Program	7
Probability Distribution	8
Monte Carlo	10
Program Details.	11
Sample Runs of the OPS Program.	17
ACKNOWLEDGEMENTS	26
REFERENCES	26
APPENDIX A: RANDOM VARIABLES AND DISTRIBUTION	A.1
Beta Distribution	A.8
APPENDIX B: NEWTON-RAPHSON ALGORITHM	B.1
APPENDIX C: INPUT FORMATS.	C.1
DISTRIBUTION	Distr-1

OPS: RISK ANALYSIS IN CAPITAL INVESTMENT AND DEVELOPMENT OPPORTUNITIES

R. W. Legan

R. L. Engel

SUMMARY

Decision makers are constantly faced with the problem of analyzing the potential profit of a proposed investment or with the choice between alternative courses of action. The more commonly used methods for such analysis are discussed as well as their individual advantages and disadvantages. A common shortcoming of all past methods has been their failure to provide an indication of the risk inherent in a given investment; while a project may indicate a certain rate of return, no analysis is provided to show the probability of actually realizing that return.

A computer program, OPS, was developed by Battelle-Northwest to perform such a risk analysis. This program uses the Monte Carlo method to determine probability distributions for the calculated rate of return, present worth, and benefit-cost ratios for investment opportunities. The logic flow of the program is described as well as its input/output characteristics. A sample problem is also included.

The theory of random variables and their application to probability distributions is discussed. The Newton-Raphson algorithm for the solution of the rate-of-return polynomial equation is explained. Detailed user instructions are also included.

LIST OF FIGURES

	<u>Page</u>
1 Probability Distribution Functions	9
2 Flow Chart of Monte Carlo Simulation as Applied in OPS.	12
3 Typical Frequency Histogram for Present Worth Calculation.	19

LIST OF TABLES

	<u>Page</u>
I Listing of OPS Input Parameters	14
II Typical Results of Present Worth Calculation.	16
III Typical OPS Statistical Data for Present Worth	18
IV Project A Cash Flow	20
V Project B Cash Flow	20
VI OPS Analysis of Project A	22
VII OPS Analysis of Project B	23
VIII OPS Analysis of Project A with 3% Inflation	24
IX OPS Analysis of Project B with 3% Inflation	25

LIST OF TABLES

Page

1	Listing of QPS Input Parameters	14
II	Typical Results of Present Worth Calculations	16
III	Typical QPS Statistical Data for Present Worth	18
IV	Project A Cash Flow	20
V	Project B Cash Flow	20
VI	QPS Analysis of Project A	22
VII	QPS Analysis of Project B	23
VIII	QPS Analysis of Project A with 3% Inflation	24
IX	QPS Analysis of Project B with 3% Inflation	25

OPS: RISK ANALYSIS IN CAPITAL INVESTMENT AND DEVELOPMENT OPPORTUNITIES

R. W. Legan

R. L. Engel

INTRODUCTION

Decision makers always have been faced with the problem of choosing between alternative investment possibilities. To quote from a basic text on the subject: "The conduct of a business enterprise requires a successive series of business decisions--decisions between possible alternatives with reference to the future."¹ The success of management can be judged by its ability to make the optimum selection of alternatives through the accurate, subjective analysis of all pertinent factors. There are numerous published methods for performing such analyses. Some of the more common ones are the rate of return, present worth, and benefit cost techniques.

Recently, including discount analysis in these various forms of economic analysis has become popular. It also appears that use of these methods may increase because of the desire to articulate difficult alternatives for public review. To simplify such review, there is a strong desire to apply a system that develops a single ratio or small group of numbers as a basis for meaningful decisions. Discount analysis, as practiced in the past, appears to offer a reasonable means of assisting in decision making with respect to relatively well developed alternatives for accomplishing fairly firm objectives, such as building a large standard design process plant in the industrial sector or a waterworks in the public sector. In these instances, the uncertainties involved generally are relatively small and can be accounted for through concepts of economic risk associated with classical private and governmental financial resources and investment.

This document is concerned with the inadequacy of applying these classical discount analysis methods to situations involving larger uncertainties, such as those related to introduction of new products into the market place, construction of new untested facilities, or

prediction of events far in the future. In particular, two such areas come to mind: 1) evaluations of projects based on results from recent or future research and development and 2) evaluations of the worth of preserving facets of our natural environment. Often, evaluation of these areas has been considered to include intangibles beyond numerical quantification. The careful scientific analyst, of course, has rationalized applying past methods of economic analysis in areas of uncertainty by adding suitable disclaimers to the credibility of the results in each instance. However, since decisions involving these areas are evermore in the public eye, the danger of total misinterpretation of results is large. Often, during reporting to the public, disclaimers are withheld or omitted, and in the event disclaimers are not omitted, there is a distinct danger of error in both directions by the public either under- or overcompensating for the disclaimers. Perhaps, then, it is best to sum up that the disclaimers may be either lost, ignored, or misinterpreted.

The degree of difficulty in making economic decisions can vary widely depending on the number of factors to be considered, the size of the investment (and its attendant financing), and the economic lifetime of the project. It is axiomatic that a planner can predict with more accuracy and confidence near-term economic trends as opposed to those of the long term. The methods used by the decision makers will vary, then, with both the complexity and time span of the project. It is the intent of this document to describe some of the more commonly used analytic methods, to discuss their strengths and weaknesses, and, finally, to describe a computer program which uses the Monte Carlo technique developed for neutron physics analysis to solve the complicated problem of determining the probability of realizing the calculated returns. Such a solution assists the decision maker when evaluating competitive investments and by providing a description of the probabilities of success inherent in each of the alternatives.

The various methods of discounted benefit cost analysis are described in the literature. While it is tempting to provide a thorough treatment of the subject in this report, we shall only briefly describe the salient features to acquaint the newcomer with discount analysis and, at the same

time, provide a brief review for the more experienced reader. In essence, discount analysis is applied when numerous expenditures and receipts must be accounted for when making a decision and when the length of time involved is greater than a single accounting period, possibly as long as decades. Thus, discount analysis is involved with what are known as impressment expenditures; the crucial dimension that distinguishes them from other types of expenditures is the elapse of time; investments are made now in anticipation of future returns. Obviously, a sum received today is worth more than the same sum to be received tomorrow. For the sum received and invested today will grow in value by tomorrow; or, the sum received and spent today for a pleasure is spent on the basis that the pleasure of today is worth more than a larger pleasure tomorrow. Thus, in evaluating investment projects, it is fundamental to consider the timing of returns on the investment and the amount of future return that imparts adequate future satisfaction so that the tendency to spend today for returns today is overridden by the satisfaction of anticipated returns.

In extending the use of discount analysis, time is important in another sense because of the changing world. The conditions of the future may be different from those of today. This creates much uncertainty in both the future rates of return and future costs. The probability of realizing a desired return on an investment is quite dependent upon the probabilities of future returns and costs actually occurring.

From an investment viewpoint alone, time has another importance in the sophisticated world of today because of inflation. A dollar invested today will be repaid by the inflated dollars of tomorrow. It is desirable to know the rate of return in terms of both today's and tomorrow's dollar so that a measure is obtained of the ability to realize a reasonable rate of return and also be able to replace the proposed facilities when they are obsolete.

The computer program, OPS*, has been designed to assist evaluation of these uncertainties related to future investment possibilities. Each of the cost flow items (installation costs, operating costs, incomes, etc.) used in the evaluation is expressed three times; as a most probable value, the least likely value, and the highest likely value. The code then calculates the probability distribution for the present worth, the rate of return, and the benefit cost ratio. This permits the evaluator to determine the probability that any goal value for these economic parameters will be realized. Much better evaluation of investment possibilities should be possible in comparison to use of classical economic evaluations based on only use of the most probable values for the cash flow items.

DISCUSSION

Alternative Analytical Methods

Probably the most easily understood method of investment analysis is the one using simple interest rate of return. If a \$100 investment offers a net annual profit of \$15, it can be said to be yielding a 15% return on investment. If the lifetime of the investment is short and if its size is small compared to the liquid assets of the company, such an analysis, while not elegant, may be adequate. Assuming that other economic alternatives possess the same time and size characteristics, a sound choice would be that option which yielded the greatest return on investment.

This method of analysis has the obvious advantage of simplicity. Its disadvantages are great, though, when the investment size of one or more of the options becomes large, when the economic lifetimes of the alternatives become long, and when such items as cash flow and salvage value must be included. At this point the time cost of money becomes very important and to neglect it can cause gross errors.

By using a sinking fund calculation, the depreciation of the original investment can be reduced to an annual amount. If this is the only investment and if the revenues and expenses are uniform, a more realistic rate

*OPS is the Roman goddess of the earth as the source of fertility and also the goddess of abundance and wealth in general. The name is appropriate for this code because of the abundance of data provided by the code and because use of the code should assist attainment of greater wealth.

of return may be calculated for it includes an allowance for the reinvested value of the depreciation account. This technique becomes impractical to use, though, when the annual cash flow items are non-uniform or when several capital outlays are required. It has the further disadvantage that the revenues accruing from the depreciation account are assumed to be reinvested in the business at the same rate of return as that used for the depreciation schedule. In order for this approach to be accurate, the depreciation rate used must coincide with the average before tax (revenues from depreciation accounts are tax exempt) return of all capital invested in the business.

A more versatile rate of return calculation is called the internal-rate-of-return method. This method discounts all expenses and revenues and thereby fully accounts for the time value of money. The basic equation for such an analysis is:

$$C = \frac{B_1}{(1+r)} + \frac{B_2}{(1+r)^2} + \frac{B_3}{(1+r)^3} + \dots + \frac{B_n}{(1+r)^n}$$

where C = the initial capital investment

B_1 = net revenue in the first year

B_2 = net revenue in the second year, etc.

When this equation is solved for "r" the resulting rate is called the internal-rate-of-return. Subsequent capital investments can be readily handled by this method by discounting them in the same manner as the revenues. This can be accommodated by combining net revenues and investments in a given year. If, for example, B_2 is redefined as the net revenue of the second year minus the capital expenditures of that year, the same basic equation may be used.

The advantage of using this analytical method is that the results are independent of the size and timing of the required investment. If the

calculated rate of return is materially higher than the cost of borrowing capital, the investment is favorable; if alternatives are being analyzed, the one with the highest rate of return is superior. No decisions need to be made on the appropriate interest rate to be used.

The chief disadvantage of this method is that it is essential that the revenues be reinvested at the same rate of return as calculated, if the results are to be accurate. If the indicated rate is near the average rate of return for all capital investment in the business, the probable error in selection on this basis is small. A second drawback to this technique is the fact that the solution of the basic equation for "r" involves an iterative technique; with a computer available, this is a trivial criticism. Lastly, the solution of a polynomial equation can give more than one answer. Usually extraneous roots may be eliminated by inspection.

A modification of the same basic equation as discussed above may be used to find the net present worth of a project. In this case, though, "r", the present worth factor, is assumed--usually the current cost of capital. If B is defined as net revenue in a year minus any capital expense in that year, and if all terms are gathered on the right hand side, the sign of the sum indicates the desirability of the project. If the sum is positive, the present worth of the net revenues exceeds the present worth of the capital expenditures and the venture is profitable. It is said to have a positive net cash flow. Of two alternatives, the one with the greater net present worth is deemed superior.

One advantage of this method over the internal-rate-of-return is the small inherent error induced by the interest rate used. For the present worth analysis to be accurate, it is only necessary that the annual revenues be reinvested at the assumed interest rate. If this is the current marketplace cost of money, the only error induced will come from time fluctuations in this rate. If this method is used for selecting the better of two alternatives of the same economic lifetime, the errors will probably not be such as to change the validity of the decision.

Even though the use of a rate equal to the cost of borrowing money has technical merit, there are many arguments for using a rate equal to the average return on capital invested in the business. This, in essence, dictates that a new project must show profitability equal to the average of other company projects to be considered. There are further arguments, which are too abstruse to discuss at length here, which claim that all projects should not be analyzed with the same interest rate--the higher the risk or the larger the investment, the higher the interest rate to be used.

There are many variations of the methods discussed briefly here. These include a commonly quoted one which is the benefit-to-cost ratio. This is usually the discounted revenues divided by the discounted costs.

While all of the methods discussed here have unique advantages, they all share a single disadvantage. None of them gives an indication of the probability of a calculated return (or a net present worth) actually being realized from a given investment. Even though it is often necessary to guess at economic lifetimes and cash flows for many years in the future, the results from none of these analyses indicate the uncertainties inherent in the investment being considered. If the rate-of-return as well as the probability of achieving it could be calculated, the decision maker would have a much more satisfactory basis for selection. For example, if Project A indicated a 20% internal-rate-of-return, but only a 30% chance of achieving this return, while Project B showed a 70% chance of returning 15%, Project B might be selected--converse to the choice which might be expected if the probability data were not available.

The following sections describe a computer program which was written to fulfill this need for analysis of the common economic indices, rate-of-return, net present worth, and benefit-cost ratio--all on a probabilistic basis.

OPS Program

A computer program, OPS, has been developed to present a comparative analysis of investment projects, as represented by a cash flow sequence. It estimates the probability distributions of calculated present worths,

present worth benefit-cost ratios, and rate of return. The investigator can use these distributions of the project value indicators to determine the probability that a project will yield the desired economic results. The technique used by OPS to arrive at these results is the Monte Carlo simulation technique.

In the following section the Monte Carlo technique will be described as well as probability functions. The program details of OPS will be discussed. Detailed instructions will be given for use of the program; and, lastly, sample calculations will be illustrated.

Probability Distribution

Suppose that an event is going to occur and that this occurrence can be measured somehow and the measurement will then be expressed by a number. Usually, the exact value of the occurrence will not be known; but often the set of all possible values can be described and each value in the set will, then, have some probability of being the one that occurs. This probability is represented by a number, which must be between zero and one, and the sum of the probabilities of all possible values equals one. A function that expresses this probability relationship is called a probability density function, and the set of all possible values is said to be distributed according to that function.

Figure 1 shows three simple probability density functions. Figure 1a is called a uniform distribution, each value between 0 and 1 has equal probability of occurring. If the event represented by this distribution were repeated a large number of times, each value should occur about the same number of times. Figure 1b is a distribution in which the value $1/2$ is twice as likely to occur as either value $1/4$ or $3/4$. Probability distributions of various shapes can be used; the most common, known as the normal (or Gaussian) and which is always symmetric, is shown in Figure 1c. The distribution used in the OPS Program is known as the beta distribution. It was selected because it is very versatile, can take on many shapes, and does not need to be symmetric. A more complete discussion of probability distributions, and in particular the beta distribution, is contained in Appendix A.

PROBABILITY

1a

PROBABILITY

1b

PROBABILITY

1c

FIGURE 1
Probability Distribution Functions

The events of interest in OPS are the items in a cash flow sequence. Future expenditures or incomes are not known exactly, not even the probabilities of possible values are known exactly. However, the most probable value and a likely range of possible values usually can be estimated. This range should be estimated so that the investigator is reasonably sure that the value will fall in this range. If a range were all inclusive, it probably would be too broad for useful analysis. Using the three input parameters, most probable value, lower likely value, and upper likely value, the OPS Program estimates a beta distribution which represents the probability associated with each possible value in the selected range. The beta distribution parameters are determined from the three input parameters by a method known as matching moments, which is discussed in Appendix A.

Monte Carlo

Monte Carlo is a technique which has been popularized since the introduction of high-speed computers. It is a simulation technique that has been used extensively in neutron physics because the transmission of beams of neutrons through scattering media is a random process so complicated that analytical solution is practically impossible. The technique is useful for simulating performance of other complex systems when the distributions of system input components are known or estimated. The technique estimates a distribution of system performance values. It has also been referred to as synthetic sampling or empirical sampling.

The technique is based upon selecting a random value from a distribution of problem input values. To use this technique, each system input component is represented by a probability distribution. Then one value is selected randomly from each input component distribution. Treating these selected values as system components known with certainty, a system performance value is calculated. These sampling-calculating steps are repeated many times, always selecting new random values from each input distribution. The results from these many calculations are then summarized and plotted, resulting in an approximation of the probability distribution of the system performance. In the OPS Program, the system

input components are the items in a cash flow sequence and the system performance values are net present worth, rate-of-return, or benefit-cost ratio. Figure 2 is a flow chart of the monte Carlo technique as applied to a cash flow sequence by the OPS Program.

There are tables of random numbers available for use in selecting values from the input component distributions, but with the calculational speed of modern computing equipment, it is more economical to generate random values. In practice this is a pseudo-random number set since it is generated by a systematic mathematical process, where randomness is defined with respect to the lack of any short-range correlation of digits or sequences of digits. The basic random number generator produces a non-repeating series of integer values from a large interval of values. These random integers are translated into real numbers whose frequency of occurrence will resemble the input parameter beta distribution of interest. The following four steps are required to make the desired transformation:

1. Divide the random integer by the length of the range of integers, resulting in a uniformly distributed set of random real numbers.
2. Transform the uniformly distributed values into a normally distributed set of values symmetric about 0 with variance 1.⁽²⁾
3. Several values from the normal distribution are translated into a single value from a beta distribution on the interval between 0 and 1.⁽³⁾
4. Translate to the proper interval.

A single sample from this beta distribution might be any allowable value, but a frequency plot of values generated will resemble the proper beta distribution function when the number of samples generated is large enough.

Program Details

The OPS Program gives the user considerable flexibility as to the projects that can be analyzed. First of all, a single project can be analyzed or two projects compared. When two projects are to be compared, they can

FIGURE 2

Flow Chart of Monte Carlo Simulation as Applied in OPS

be analyzed independently and then the cash flow difference between the two projects can be treated as another cash flow and analyzed. Up to five present worth factors can be specified. Up to five end-of-project times can be specified. An inflation percentage can be specified. Finally, either discrete or continuous interest can be used. Usually interest is compounded a small number of times annually, such as quarterly. This corresponds to the discrete interest convention. If the compounding period gets smaller, such as daily, it is very nearly equal to continuous compounding, so the user has the option of using continuous interest.

Table I is a sample of cash flow input data for two alternate projects (A and B). A general listing of the input data for the code is presented in Appendix C.

The OPS Program operates in the following manner: first the input parameters, consisting of the above option specifications and the cash flow values are read into the computer. Detailed input instructions are given in Appendix C. Three values are specified for each cash flow item, the most likely value, and the highest and lowest values of the range in which the user is reasonably sure the cost will fall. The most likely value is the value where the probability distribution function peaks. Each cash flow item is represented by a beta distribution in OPS, so the program first calculates the beta parameters (γ, η) from the three input values, as described in Appendix A.

Next, the computer selects a single value randomly from each distribution. Then five (or three) economic attributes are calculated by use of these selected values:

1. Net present worth (present worth benefits minus present worth costs).
2. Present worth benefit-cost ratio.
3. Rate of return (determined by the Newton-Raphson iterative technique discussed in Appendix B).
4. Inflated benefit-cost (optional--whenever an inflation rate is specified by the user).
5. Inflated benefit-cost ratio (also optional).

TABLE I

Listing of OPS Input Parameters

The Following Data is Input to This Program

No. of Cash Flow Cards 10

	Year of Cost or Benefit	Project A Cost or Benefit			Project B Cost or Benefit		
		Probable Value	Minimum Value	Maximum Value	Probable Value	Minimum Value	Maximum Value
Cash Flow 1	1.00	-1000.00	-980.00	-1030.00	-1000.00	-980.00	-1030.00
Cash Flow 2	2.00	100.00	60.00	160.00	.00	-40.00	-60.00
Cash Flow 3	3.00	400.00	340.00	490.00	100.00	40.00	190.00
Cash Flow 4	4.00	300.00	220.00	420.00	200.00	120.00	320.00
Cash Flow 5	5.00	200.00	100.00	350.00	200.00	100.00	350.00
Cash Flow 6	6.00	100.00	-20.00	280.00	200.00	80.00	380.00
Cash Flow 7	7.00	100.00	-40.00	310.00	200.00	60.00	410.00
Cash Flow 8	8.00	100.00	-60.00	340.00	200.00	40.00	440.00
Cash Flow 9	9.00	100.00	-80.00	370.00	200.00	20.00	470.00
Cash Flow 10	10.00	100.00	-100.00	400.00	200.00	.00	500.00
Future Cash Flow		100.00	-120.00	430.00	200.00	-20.00	530.00
Present Worth Factors			5.00	7.50	10.00	12.50	
Time Periods	10						
Rate of Inflation	3.00						

IN THE FOLLOWING CALCULATIONS END OF YEAR (DISCRETE) INTEREST IS USED.

The selected values read into the computer are assumed to be not inflated. Then for the calculations of inflated values, the inflation factor is applied to the random samples after which the present worth factors are applied. The three (or five) calculated values are stored and later used to describe the probability distributions for the calculated project attributes. This random selecting--calculational procedure continues until conditions satisfying convergence control are satisfied. Convergence is tested every 100 simulations. Three convergence parameters are specified: a minimum number of simulations, a maximum number of simulations, and a ratio applied to the average present worth benefit-cost. That is, if μ_i is the average present worth benefit-cost, then convergence is satisfied when:

$$1. - \frac{\mu_i}{\mu_{i-1}} \leq \text{ratio}$$

Usually a minimum of from 300 to 500 iterations and a maximum of 600 to 800 iterations will yield adequate convergence for most problems. Within these limits, a convergence ratio of 0.005 should be specified. A minimum of 100 and a maximum of 2000 iterations are assumed if not specified.

This simulation method is performed for each case, where a case is determined by an interest rate, a project end-time, and a project (either project A, B, or A-B). Three types of output are available for each attribute calculated for each use. Table II is the results of an OPS calculation based on the data of Table I and presents the following information for the calculation of the present worth:

1. Number of simulation iterations.
2. Average value of present worth.
3. Standard deviation for present worth.
4. Maximum value observed.
5. Minimum value observed.
6. Skewness estimate.

Skewness is a measure of assymetry.

If this value is less than 0, the distribution is skewed to the left (long left tail). If this value is greater than 0, the distribution is skewed to the right (long right tail).

TABLE II

Typical Results of Present Worth Calculation

Project A

The Distribution of Present Worth (Benefit Minus Costs Discounted) at an Interest Rate of .00 and thru the 10 Cash Flow Period

The Sampling From the Beta Distribution of each Cost or Benefit (The Cash Flow Diagram) is Checked Every 100 Times and if the Percent Change is Less Than 0.5, The Sampling Stops

Number of Entries in Table	300
Average	5.9765278+02
Standard Deviation	.1485803+03
The Maximum Value is	1069.58730
The Minimum Value is	194.47788
Skewness Estimate From Moments	.20902
Skewness Estimate From K-Statistics	.21042

The Distribution is Skewed to the Right That is it has a Right (High) Tail

Kurtosis Estimate From Moments	3.07823
Kurtosis Estimate From K-Statistics	3.09982

The Distribution is More Peaked Than a Normal Distribution.

7. Kurtosis estimate.

Kurtosis is a measure of peakedness.

The normal distribution is used as a peakedness standard.

If this value is greater than 3.0, the distribution is more peaked than a standard normal distribution; if less than 3.0, the distribution is less peaked than a normal distribution. A flat distribution has a value of 1.8.

Table III is a frequency table. This table presents, for several intervals of calculated values:

1. The frequency
2. The relative frequency (percent)
3. Cumulative percentages
4. Cumulative remainder (100 minus item 3)

The cumulative percentage is a powerful analytical guide. It estimates the probability that the calculated attribute will be less than a certain amount. Table III may be suppressed by input parameter control. Figure III, which may also be suppressed by input parameter, is a histogram of the distribution and presents a graph of the probability density function of the calculated attribute.

Sample Runs of the OPS Program

The two fictitious projects described on Table I were selected to demonstrate the OPS analysis. Each project is represented by different cash flows, but both have a total cost of \$1000 and a total benefit of \$1500. The values for the two cash flows are repeated in Tables IV and V. These cash flow streams were input as the most probable cash flows into the OPS Program. Uncertainty associated with each cash flow item was specified as a function of the year (distance in the future) as follows:

Minimum Value = Probable Value - Year Multiplied by \$20

Maximum Value = Probable Value + Year Multiplied by \$30

TABLE III

Typical OPS Statistical Data for Present Worth

UPPER LIMIT	OBSERVED FREQUENCY	PERCENT OF TOTAL	CUMULATIVE PERCENTAGE	CUMULATIVE REMAINDER
210.00	2	.67	.67	99.33
240.00	0	.00	.67	99.33
270.00	0	.00	.67	99.33
300.00	1	.33	1.00	99.00
330.00	3	1.00	2.00	98.00
360.00	6	2.00	4.00	96.00
390.00	14	4.67	8.67	91.33
420.00	14	4.67	13.33	86.67
450.00	11	3.67	17.00	83.00
480.00	16	5.33	22.33	77.67
510.00	19	6.33	28.67	71.33
540.00	22	7.33	36.00	64.00
570.00	30	10.00	46.00	54.00
600.00	15	5.00	51.00	49.00
630.00	20	6.67	57.67	42.33
660.00	23	7.67	65.33	34.67
690.00	26	8.67	74.00	26.00
720.00	21	7.00	81.00	19.00
750.00	16	5.33	86.33	13.67
780.00	13	4.33	90.67	9.33
810.00	3	1.00	91.67	8.33
840.00	10	3.33	95.00	5.00
870.00	5	1.67	96.67	3.33
900.00	2	.67	97.33	2.67
930.00	2	.67	98.00	2.00
960.00	1	.33	98.33	1.67
990.00	1	.33	98.67	1.33
1020.00	3	1.00	99.67	.33
1050.00	0	.00	99.67	.33
1080.00	1	.33	100.00	.00

Project A
The Distribution of Present Worth (Benefit Minus Costs Discounted at an Interest
Rate of .10 and Thru the 10 Cash Flow Period

```

210.000+****
240.000+
270.000+
300.000+***
330.000+*****
360.000+*****
390.000+*****
420.000+*****
450.000+*****
480.000+*****
510.000+*****
540.000+*****
570.000+*****
600.000+AAAAAAAAAAAAAAAAAAAAAAAAAAAA
630.000+*****
660.000+*****
690.000+*****
720.000+*****
750.000+*****
780.000+*****
810.000+*****
840.000+*****
870.000+*****
900.000+*****
930.000+*****
960.000+**
990.000+**
1020.000+*****
1050.000+
1080.000+**

```

FIGURE 3 Typical Frequency Histogram for Present Worth Calculation

TABLE IV
Project A Cash Flow

<u>Year</u>	<u>Probable Value</u>	<u>Minimum Value</u>	<u>Maximum Value</u>	<u>Calculated Average Value</u>	<u>Typical Generated Average Value</u>
1	-1000	-1030	-980	1001.67	-1001.94
2	100	60	160	103.33	103.80
3	400	340	490	405.00	405.15
4	300	220	420	306.67	305.22
5	200	100	350	208.33	212.36
6	100	-20	280	110.00	109.82
7	100	-40	310	111.67	116.71
8	100	-60	340	113.33	113.54
9	100	-80	370	115.00	111.73
10	100	-100	400	116.67	121.27

TABLE V
Project B Cash Flow

<u>Year</u>	<u>Probable Value</u>	<u>Minimum Value</u>	<u>Maximum Value</u>	<u>Calculated Average Value</u>	<u>Typical Generated Average Value</u>
1	-1000	-1030	-980	-1001.67	-1002.05
2	0	-40	60	3.33	2.12
3	100	40	190	105.00	105.45
4	200	120	320	206.67	207.22
5	200	100	350	208.33	209.45
6	200	80	380	210.00	208.34
7	200	60	410	211.67	218.86
8	200	40	440	213.33	218.08
9	200	20	470	215.00	219.46
10	200	0	500	216.67	221.58

The column headed "Calculated Average Value" is the average value of the beta distribution representing the cash flow item. It is calculated from the probable value, minimum value, and maximum value as described in the discussion of matching moments in Appendix A. Note that this average value does not coincide with the most probable value. The column headed "Generated Average Value" is the average of the random values generated by OPS in the analysis of this sample problem.

The OPS analysis of these two sample projects assuming no inflation is summarized in Tables VI and VII, and the 3% inflated analysis in Tables VIII and IX. Project A, which has earlier return, is clearly the superior project.

The greatest advantage of this type of analysis over more conventional ones is the calculation of the probability of achieving a given success from an investment. For example, Project A indicates a 12.0% rate-of-return which might be an acceptable figure for some companies. More important, though, a company might want to know with a high degree of certainty that their investment would yield at least a 10% return. The OPS analysis shows that Project A fills this requirement while Project B does not.

TABLE VI

OPS Analysis of Project "A"

	<u>Interest Rates</u>				
	<u>0%</u>	<u>5%</u>	<u>7.5%</u>	<u>10%</u>	<u>12.5%</u>
Average B-C	597.6	297.2	184.0	88.4	14.0
Minimum B-C	194.5	-0.3	-68.8	-151.3	-161.2
Maximum B-C	1069.6	552.0	458.7	296.0	278.1
Benefit/Cost	1.59	1.31	1.20	1.10	1.02

Probability of
Earning at Least

-200	100	100	100	100	100
-100	100	100	100	98.7	96.5
0	100	99.8	98.2	86.7	55.2
100	100	96.8	83.4	43.7	9.5
200	100	79.0	41.8	10.3	0.5
300	99.0	48.8	8.8	0	0
400	89.8	17.0	0.8	0	0
500	73.5	3.0	0	0	0
600	49.0	0	0	0	0
700	23.7	0	0	0	0
800	8.7	0	0	0	0
900	2.7	0	0	0	0
1000	1.0	0	0	0	0
1100	0	0	0	0	0

Rate of ReturnCumulative Probability

18%	1
16%	6
14%	29
12.8%	50
12%	65
10%	87
8%	99
6%	100

TABLE VII
OPS Analysis of Project "B"

	Interest Rates				
	0%	5%	7.5%	10%	12.5%
Average B-C	608.5	203.7	58.0	-49.1	-141.3
Minimum B-C	153.1	-131.4	-191.4	-293.2	-316.2
Maximum B-C	1003.1	479.4	307.3	214.5	25.1
Benefit/Cost	1.60	1.21	1.06	0.95	0.84
Probability of Earning at Least					
-400	100	100	100	100	100
-300	100	100	100	100	99.6
-200	100	100	100	97.6	80.6
-100	100	99.7	95.5	73.0	28.6
0	100	96.3	73.5	27.0	1.2
100	100	84.0	33.0	3.4	0
200	99.7	53.0	9.0	0.2	0
300	98.3	16.3	0.5	0	0
400	89.0	3.3	0	0	0
500	73.7	0	0	0	0
600	53.7	0	0	0	0
700	32.0	0	0	0	0
800	13.3	0	0	0	0
900	3.0	0	0	0	0
1000	0.3	0	0	0	0
1100	0	0	0	0	0
Rate of Return		Cumulative Probability			
14%		0			
12%		3			
10%		30			
8.9%		50			
8%		67			
6%		92			
4%		99			
2%		100			

TABLE VIII

OPS Analysis of Project "A" With 3% Inflation

	Interest Rates				
	0%	5%	7.5%	10%	12.5%
Average B-C	837.8	469.0	329.0	211.1	121.0
Minimum B-C	317.0	106.6	16.4	-82.8	-94.3
Maximum B-C	1428.6	807.0	666.8	470.1	445.4
Benefit/Cost	1.81	1.48	1.34	1.23	1.13
Probability of Earning at Least					
-100	100	100	100	100	100
0	100	100	100	98.7	93.2
100	100	100	98.2	88.3	56.8
200	100	97.5	86.6	53.3	16.5
300	100	89.2	61.0	17.0	2.0
400	99.3	68.2	25.0	4.0	0.5
500	98.1	41.8	6.0	0	0
600	89.3	17.2	0.6	0	0
700	76.2	4.2	0	0	0
800	55.6	0.2	0	0	0
900	39.0	0	0	0	0
1000	17.2	0	0	0	0
1100	8.2	0	0	0	0
1200	2.7	0	0	0	0
1300	1.7	0	0	0	0
1400	0.3	0	0	0	0
1500	0	0	0	0	0

TABLE IX

OPS Analysis of Project "B" With 3% Inflation

	Interest Rates				
	0%	5%	7.5%	10%	12.5%
Average B-C	939.7	424.6	238.0	102.0	-15.5
Minimum B-C	372.8	4.86	-70.5	-195.8	-233.1
Maximum B-C	1429.9	777.1	554.4	412.2	184.2
Benefit/Cost	1.91	1.43	1.25	1.11	0.98

Probability of
Earning at Least

-300	100	100	100	100	100
-200	100	100	100	100	99.2
-100	100	100	100	98.6	83.4
0	100	100	98.2	86.6	46.2
100	100	99.3	87.5	47.6	6.6
200	100	95.3	61.2	17.6	0
300	100	81.7	29.8	2.4	0
400	99.7	58.7	10.2	0.4	0
500	98.8	29.7	1.2	0	0
600	96.0	9.0	0	0	0
700	86.4	2.0	0	0	0
800	73.8	0	0	0	0
900	57.0	0	0	0	0
1000	40.3	0	0	0	0
1100	23.5	0	0	0	0
1200	12.3	0	0	0	0
1300	3.8	0	0	0	0
1400	0.8	0	0	0	0
1500	0	0	0	0	0

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the assistance of R. L. Hooper for his investigations of several random number generators and his authorship of the appendix on probability. We also wish to thank E. A. Eschbach, who provided the initial scope of the project, and J. B. Burnham for his editorial assistance.

REFERENCES

1. E. L. Grant, Principles of Engineering Economy, Revised First Edition, Ronald Press Company, 1941.
2. Abramowitz and Stegun, Handbook of Mathematical Functions, U. S. Department of Commerce, NBS, Applied Math Series 55, 1964.
3. Hahn and Shapiro, Statistical Models in Engineering, John Wiley and Sons, Inc., 1967.
4. F. S. Hillier and G. J. Lieberman, Introduction to Operations Research, Holde-Day, Inc., 1967.

APPENDIX A

RANDOM VARIABLES AND DISTRIBUTIONS

RANDOM VARIABLES AND DISTRIBUTIONS

The discussion of random variables and distributions that follows will be relatively brief and is designed to serve only our purposes here. Our objective is not to teach mathematical techniques although a few mathematical concepts must be introduced. The reader should be aware that we are only touching on the topics mentioned, and he is referred to any good book on statistics or probability for a more complete treatment.

Our particular problem is one of representing the results of a determination in terms of the likelihood that certain results or sets of results will occur. Mathematicians tend to express everything ultimately in terms of set theory. The basic concept in our case becomes then the set of all conceivable outcomes of the experiment, and this set is called the sample space. This space is thought of as consisting of points, each of which represents a certain outcome. For example, if an experiment consists of tossing a penny twice and observing heads or tails, the sample space consists of four points which can be represented by the symbols (H,H) , (H,T) , (T,H) and (T,T) , where H designates heads and T designates tails.

For the moment, let us suppose that the sample space is discrete, that is, it contains only a finite number of points or an infinite number which can be counted (that is, the positive integers $1, 2, 3, \dots$). Associated with each point in a discrete sample space is a number which we call the probability that the outcome represented by that point will occur. There is no reason to define this beyond specifying that these probabilities are non-negative numbers whose sum is one.

Next we need the concept of a random variable, which is a function defined on the sample space to the real numbers. To the many readers who find this kind of abstraction unexciting at best, we hasten to point out that it is sufficient here to think of a random variable as a symbol for a number which is going to be produced by an experiment or determination. Once produced, the number is, of course, no longer random, and is called a realization, or instance of the random variable. The word random applies to the process that produces the number, rather than to the number itself.

To put these ideas together let's go back to the penny tossing experiment and define the random variable R as the number of heads occurring in the two tosses. Clearly the random variable R can take on just three values; 0 for no heads in two tosses, 1 for one head in two tosses and 2 for two heads in two tosses. Further, if the probability of tossing a head is 0.7 and the probability of tossing a tail is 0.3 we can associate a probability with each point in the sample space and consequently with each value of R .

Sample Space:	(H,H)	(H,T)	(T,H)	(T,T)
Probability:	$0.7(0.7)=0.49$	$0.7(0.3)=0.21$	$0.3(0.7)=0.21$	$0.3(0.3)=0.09$
Value R :	2	1	1	0

Therefore the "probability function" which describes the distribution of the random variable R is easily written as:

$$p(X) = \Pr(R=X) \quad \text{for } X = 0,1,2$$

$\therefore$

$$p(0) = \Pr(R=0) = 0.09$$

$$p(1) = \Pr(R=1) = 0.42$$

$$p(2) = \Pr(R=2) = 0.49$$

The function $p(X)$ which describes the distribution of a random variable is called a probability density function. A random variable is also characterized by a related function called its cumulative distribution function (or simply distribution) $P(X)$, defined by

$$P(X) = \Pr(R \leq X)$$

read, "the probability that R is less than or equal to X ." For the coin tossing example, the cumulative distribution function is easily derived as:

$$P(0) = \Pr(R \leq 0) = 0.09 = p(0)$$

$$P(1) = \Pr(R \leq 1) = 0.51 = p(0) + p(1)$$

$$P(2) = \Pr(R \leq 2) = 1.0 = p(0) + p(1) + p(2)$$

The relationship between the probability density function and the cumulative distribution function for a discrete random variable is obvious.

Although it may be possible to express all problems as discrete problems, this would often be highly inefficient. For example, an experiment

may have as its outcome the position of a needle on a meter. The possible outcomes are from a continuum, as opposed to a discrete space. If the scale on the meter runs from 0 to 100 and it can be read at best to a tenth of a unit, then it is true that there are just 1,000 distinguishable positions of the needle, and these make up a discrete sample space that could be used to study the experiment. It is usually easier not to use the discrete model when there are so many points in the space.

When the sample space is not discrete, the concepts of probability, random variable, and distribution are developed in an analogous way, but the procedure for doing so requires the use of measure theory. We shall not go into this, but simply define a random variable as a quantity R characterized by a distribution

$$\Pr(R \leq X) = P(X),$$

where $P(X)$ is any function of a real variable satisfying the conditions

$$(1) P(-\infty) = 0$$

$$(2) P(\infty) = 1$$

$$(3) P(X) \leq P(X+a) \text{ for } a > 0.$$

In the discrete case, R can take on only a set of values $X_1, X_2, X_3, \dots$ and the situation is described by the probability $p(X_i)$ that R is equal to X_i . This probability function tells the whole story. In the nondiscrete case this is usually not possible, but when $P(X)$ is differentiable, the derivative $P'(X)$ plays a role analogous to that of the probability function and so is written $p(X)$ and again called the probability density function. The relation between these two functions may be expressed by

$$P(X) = P'(X) = \frac{d}{dX} P(X)$$

or by

$$P(X) = \int_{-\infty}^X p(t) dt$$

since

$$P(X) = \Pr(R \leq X), \text{ it follows that}$$
$$\Pr(a \leq R \leq b) = P(b) - P(a) = \int_a^b p(t) dt.$$

These ideas may be more apparent by considering Figures A.1 through A.4.

To relate these ideas to the problem at hand we need first to consider the notion of a random number generator. A random number generator is a mechanism by which we can produce a random sample from a completely defined distribution of numbers. A random sample of size n is a set of realizations or instances of the random variable X defined by the distribution $P(X)$ such that every possible sample of size n has an equal chance of being selected. In actual operation the random number generator produces a sample of size n by successive samples of size one.

Suppose we are considering the cost of building project. In estimating the total cost of the project it is logical, if not necessary, to partition the project into units which are fairly independent in a cost sense. A partitioning of this nature could include such things as excavating, electrical wiring, sheet metal work and landscaping. Now if the cost of the i^{th} unit is estimated to be C_i then the estimated cost of the project can be defined as:

$$C_T = \sum_{i=1}^k C_i$$

where k is the number of units.

Here it is important to recognize that while the estimate of the cost of the i^{th} unit is obtained by considering the best information available, there are still a number of factors either overlooked or not anticipated which can cause the true cost when actually incurred to deviate from the estimated cost. This, of course, will in turn cause the true cost of the project to deviate from the estimated total cost. This fact can be used

THIS FIGURE IS SIMPLY A PICTURE OF THE PROBABILITY DENSITY FUNCTION $p(X)$. THE AREA UNDER THE CURVE COMPUTED AS $\int_{-\infty}^{\infty} p(X) dx$ IS ONE IN ORDER FOR $p(X)$ TO QUALIFY AS A PROBABILITY DENSITY FUNCTION.

FIGURE A.1

THE SHADED AREA UNDER THIS CURVE IS $P(a)$ DEFINED AS:

$$P(a) = \Pr(X < a) = \int_{-\infty}^a p(t) dt$$

FIGURE A.2

THE SHADED AREA UNDER THIS CURVE IS $P(b)$
DEFINED AS:

$$P(b) = \Pr(X \leq b) = \int_{-\infty}^b p(t) dt$$

FIGURE A.3

THE SHADED AREA UNDER THIS CURVE IS
 $\Pr(a < X < b)$. BY COMPARING FIGURES 2, 3, AND 4 IT
BECOMES CLEAR THAT:

$$\begin{aligned} \Pr(a < X < b) &= P(b) - P(a) = \int_{-\infty}^b p(t) dt - \int_{-\infty}^a p(t) dt \\ &= \int_a^b p(t) dt \end{aligned}$$

FIGURE A.4

to enhance the cost analysis of the total project if, at the time we are assigning a probable cost (C_i) to the i^{th} unit, we can identify a point C_i^1 which the cost of the i^{th} unit is not likely to go below and a point C_i^2 which the cost is not likely to go above. Obviously

$$C_i^1 \leq C_i \leq C_i^2 .$$

Now if we think of the cost of the i^{th} unit as a random variable then the numbers C_i^1 and C_i^2 described above will define a continuous sample space. Further, if the probability density function $p_i(X)$ describes the distribution of numbers for the sample space so defined, we then have a completely defined distribution of numbers and a random number generator provides us an efficient sampling mechanism.

Before looking at how these ideas are used to perform an analysis of total project cost lets examine two specific probability functions which we have chosen to describe the distribution of unit costs.

Normal Distribution¹

Many random variables in physical situations are distributed in such a way as to have (at least approximately) the probability density function

$$p(X) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(X-\mu)^2}{2\sigma^2}} , (-\infty < X < \infty)$$

-
1. The normal curve was first discovered by A. De Moire (1667-1754), a French mathematician and it was published in 1733 in a privately printed pamphlet. He obtained it while working on certain problems in games of chance which were proposed to him by the gamblers of his day. Actual statistical use of the normal curve began with the work of the famous mathematical astronomers, Laplace (1749-1827) and Gauss (1777-1855), each of whom derived it independently and presumably without knowing of De Moire's treatment. They found that it represented very well the errors of observation in the physical sciences. Since that time, experience has shown that it serves quite well to describe many of the distributions which arise in the fields of biology, education, and sociology. Much of the theory of statistics is built around it.

where μ and σ are real numbers, the latter positive, and X runs over the real numbers. Such random variables are said to have the normal, or Gaussian, distribution. The distribution of a normally distributed random variable depends upon two constants, which are called parameters and they are represented in the above expression by the symbols μ and σ .

The normal curve is bell-shaped and is symmetrical about the line $X = \mu$ as shown in Figure A.5. The distribution of a normally distributed random variable is completely determined by identifying the values of the two parameters μ and σ . μ is the location parameter and its value determines the position of the curve along the X -axis. σ is the scale parameter and its value determines the width of the normal curve. Its value is the distance from μ to the points of inflection (there is a point of inflection on each side of μ and it is the point where the curve breaks over from concave to convex). Figure A.6 shows several normal curves which are located at the same place (same value of μ) but different values of σ . Figure A.7 shows several normal curves with the same value of σ but different values of the location parameter μ .

Note that for smaller values of σ we have a larger proportion of values of the random variable located near the center point μ . This makes the normal curve appear more peaked. The degree of peakedness of a probability distribution curve is referred to as kurtosis. Regardless of the combination of values of μ and σ the normal curve retains its characteristic bell-shaped, symmetrical appearance.

Beta Distribution

The beta distribution is a useful probability model for random variables whose values are limited to a finite interval. The beta probability density function defined over the interval from zero to one is

$$p(X) = \frac{\Gamma(\gamma)\Gamma(\eta)}{\Gamma(\gamma+\eta)} X^{\gamma-1} (1-X)^{\eta-1}, \quad (0 \leq X \leq 1, 0 < \gamma, 0 < \eta)$$

where $\Gamma(\gamma)$ represents the gamma function of any number and is defined by

$$\Gamma(\gamma) = \int_0^{\infty} t^{\gamma-1} e^{-t} dt.$$

FIGURE A.5
A Normal Distribution

FIGURE A.6
Normal Distributions With Constant μ

FIGURE A.7

Normal Distributions with Constant σ

The beta distribution, like the normal distribution, depends upon two parameters which are represented in the above expression by the symbols γ and η . Unlike the normal distribution which has a location parameter and a scale parameter, but no shape parameter, both parameters of the beta distribution are shape parameters. What this means is that with the proper choices of values for the parameters γ and η we can generate a wide variety of distributional shapes. This is illustrated in the following sequence of pictures (Figures A.8 through A.10).

When $\gamma > 1$ and $\eta > 1$ the distribution is single peaked with peak at the point $X = \frac{(\gamma-1)}{(\gamma+\eta-2)}$. For $\eta > \gamma > 1$ the beta distribution is positively skewed meaning that the curve is asymmetric with the tail pointing in the positive X direction. For $\gamma > \eta > 1$ the curve is negatively skewed with the tail pointing in the negative X direction. For $\gamma = \eta > 1$ the beta distribution is symmetric.

When $\gamma < 1$ and $\eta < 1$ the beta distribution curve is U shaped as shown in Figure A.9. The skewness statements made above still apply. The curve is symmetric for $\gamma = \eta$.

FIGURE A.8

Beta Distributions with $\gamma > 1$ and $\eta < 1$

FIGURE A.9

Beta Distributions with $\gamma < 1$ and $\eta < 1$

FIGURE A.10
Beta Distribution with $\gamma < 1$ and $\eta \geq 1$.

For $\gamma < 1$ and $\eta \geq 1$ the beta distribution is reverse J shaped as shown in Figure A.10. This is the extreme example of positive skew. Similarly when $\gamma \geq 1$ and $\eta < 1$ the beta curve is J shaped which implies extreme negative skew.

For our problem we have chosen the beta distribution to represent the distribution of unit costs largely because of the wide variety of shapes available using this distribution. In evaluating the costs of the unit factors we will find that quite often, for a variety of reasons, the cost of a unit factor is more likely to deviate from the estimated value (C_i) in one direction than another. Suppose we have estimated the cost of excavation (C_E) and in evaluating our estimate in terms of assigning lower (C_E^1) and upper (C_E^2) bounds we come to the conclusion that should the true cost deviate from the estimated cost it is more likely to cost less than what was estimated than it is to cost more. If we now think of the cost of excavation as a random variable the distribution which describes the random variable would need to be skewed since a symmetric distribution would imply

that the true cost could deviate from the estimated cost in either direction with equal probability. The beta distribution allowed us to skew the distribution in the required fashion.

The normal distribution is included largely because it is easy to work with. It can also be used to describe the distribution costs where the required distribution is symmetric.

Choosing between normal and beta for symmetric intervals, the upper and lower bounds (which have been identified using superscripts 1 and 2) on the cost of a unit factor as well as its estimated cost ("most probable cost") play a particularly important role since not only do the bounds on the cost define a sample space, but the three numbers together supply enough information to obtain values for the parameters of either the normal distribution or the beta distribution. A brief discussion of how this is done is contained in a footnote.²

2. The following information is available:

- C_i - the estimated "most probable" cost of the i^{th} unit.
- $C_i^{(1)}$ - an estimated cost less than C_i chosen in such a way as be certain that the actual cost of the i^{th} unit is not likely to go below it.
- $C_i^{(2)}$ - an estimated cost greater than C_i chosen in such a way as be certain that the actual cost of the i^{th} unit is not likely to go above it.

If the normal curve is to be used to describe the distribution of a unit's cost then the location of the distribution is estimated simply by setting

$$\mu = C_i.$$

This implies that the sample space defined by the points $C_i^{(1)}$ and $C_i^{(2)}$ is symmetric about the "most probable" point C_i . This means simply that $C_i - C_i^{(1)} = C_i^{(2)} - C_i$. The scale parameter is estimated by assuming that 99% of the distribution is contained in the interval $C_i^{(1)} \leq X \leq C_i^{(2)}$. This, together with the fact that for a normal distribution approximately 99% of the distribution is contained in the interval $\mu - 3\sigma \leq X \leq \mu + 3\sigma$ allows us to compute an estimate of σ by substituting C_i for μ , equating the endpoints of the two intervals and solving for σ (μ could be computed in this way also).

Footnote 2 (Cont.)

$$\sigma = \frac{C_i^1 - C_i^1}{3} = \frac{C_i^2 - C_i^1}{3}.$$

The beta distribution parameters are more difficult to compute. We again assume that 99% of the distribution is contained in the interval $C_i^1 \leq X \leq C_i^2$. The mean and variance of the distribution can be estimated as follows (the mean and variance of a distribution are now terms and they will not be defined beyond pointing out that the mean is the first moment of the distribution about zero and the variance is the second moment about the mean).

$$\text{mean} = \frac{C_i^1 + 4C_i + C_i^2}{6}$$

$$\text{variance} = [1/6 (C_i^2 - C_i^1)]^2$$

These computations are familiar from the PERT analyses (Ref. 4) and justification is available from a number of sources. These calculations yield so-called sample moments and these numbers are then equated to the population moments which are functions of the parameters γ and η for which we are trying to obtain estimates. The equations can be written as:

$$\frac{C_i^1 + 4C_i + C_i^2}{6} = \frac{C_i^2\gamma + C_i^1\eta}{\gamma + \eta}$$

and

$$[1/6(C_i^2 - C_i^1)]^2 = \frac{(C_i^2 - C_i^1)^2 \gamma \eta}{(\gamma + \eta)^2 (\gamma + \eta + 1)}.$$

We have two equations and two unknowns and the solution yields the following:

$$\gamma = A\eta$$

$$\eta = \frac{36A}{(A+1)^3} - \frac{1}{(A+1)}$$

for

$$A = \frac{5C_i^1 - C_i^2 - 4C_i}{C_i^1 - 5C_i^2 + 4C_i}$$

This method of estimation is called "matching moments" and is well described in the statistical literature.⁽³⁾

We now have all that is needed to perform a cost analysis of the project. For each of the unit factors we have an estimate of the "most probable" cost (C_1) along with estimate of cost intervals for each factor. With the appropriate choice of distribution (normal or beta distribution) estimates of the distributional parameters follow as indicated above. This, of course, requires that we think of the unit costs as random variables distributed according to the distributions estimated above across the selected intervals. This information can be summarized as follows:

<u>Unit Factor</u>	<u>Probable Cost</u>	<u>Lower Cost Limit</u>	<u>Upper Cost Limit</u>	<u>Estimated Probability Function</u>
1	C_1	C_1^1	C_1^2	$P_1(X_1)$
2	C_2	C_2^1	C_2^2	$P_2(X_2)$
3	C_3	C_3^1	C_3^2	$P_3(X_3)$
.	.	.	.	
.	.	.	.	
.	.	.	.	
n	C_n	C_n^1	C_n^2	$P_n(X_n)$

For each unit we now have a completely defined distribution of numbers. Under these circumstances an appropriate random number generator can be used to obtain random samples from these distributions and the samples can in turn be evaluated in an attempt to understand the variability in total project cost. The procedure for doing this is straight forward.

Total project cost was defined as the sum of the costs of a number of unit factors. Now using the random number generator and the above indicated information for each factor, we obtain a random sample of size one from each of the distribution of numbers. That is, for each factor we obtain an instance of the random variable (factor cost) defined by the appropriate distribution function. These numbers are then combined into an estimate of project cost. This procedure is repeated k times with the following result.

Let X_{ij} designate results of sampling from the unit cost distributions. The subscript i is used to identify the factor sampled ($i=1,2,\dots,n$). The subscript j is used to identify the sample ($j=1,2,\dots,k$). By our definition of project costs we have:

$$Y_j = \sum_{i=1}^n X_{ij} \quad \begin{matrix} (i = 1,2,\dots,n) \\ (j = 1,2,\dots,k) \end{matrix}$$

The k estimates of total project costs obtained in this fashion is in fact a sample from the distribution of project costs. At this point we can attempt to fit a distribution to the sample results so obtained, compute probabilities, and do a variety of statistical analyses.

APPENDIX B

NEWTON-RAPHSON ALGORITHM

NEWTON-RAPHSON ALGORITHM

If a cash flow sequence has several terms, the rate of return equation will be a polynomial of higher degree. No easy formula exists to find the roots of higher degree polynomials, so an iterative technique is used. This iterative process is known as the Newton-Raphson method. The geometric significance of the Newton-Raphson iteration (for real roots) can be seen by consideration of Figure B.1.

FIGURE B.1
A Newton-Raphson Iteration

Assume that the present worth as a function of the interest rate is represented by the curve in Figure B.1. The algorithm starts by assuming the root, r_1 . The value of the present worth and the functions slope is calculated and then we move along the straight line representing the slope, represented as line AB, of the curve to obtain the second guess, r_2 . This procedure is repeated until the newest guess is essentially the same as the last guess:

$$r_4 = (1 + \epsilon)r_3$$

where the root is approached as a limit when ϵ approaches zero. On the average, this algorithm took four iterations to get within 0.2 percent of the root. If multiple roots exist, this algorithm could have difficulty in finding a root. If the root finding technique is not converging when the OPS program is used, this fact is printed out and this data point is eliminated for just the rate of return calculation.

FIGURE B.1

A Newton-Raphson Iteration

Assume that the present worth is a function of the interest rate. The function is represented by the curve in Figure B.1. The algorithm starts by assuming the root, r_0 . The value of the present worth and its derivative are calculated with r_0 as the starting interest rate. The next iteration is represented by the line AB of the curve to obtain the second guess, r_1 . The process is repeated until the present worth is essentially zero. The next guess is

APPENDIX C

INPUT FORMATS

INPUT FORMATS

The following input data cards are required by the OPS Program. For all cards except the last two, those columns preceding the first parameter can be used for comments. These comments are printed in the output list of input parameters.

<u>Card</u>	<u>Columns</u>	<u>Mode</u>	<u>Value</u>
1	29-30	Integer	N = the number of cash flow cards that follow this card
2 to N+1	16-20	Real	Year of cost or benefit. This is the time used in the present worthing and inflation expressions. This number does not need to be an integer value.
	21-30	Real	Probable cost or benefit (Project A)
	31-40	Real	Minimum cost or benefit (Project A)
	41-50	Real	Maximum cost or benefit (Project A)
	51-60	Real	Probable cost or benefit (Project B)
	61-70	Real	Minimum cost or benefit (Project B)
	71-80	Real	Maximum cost or benefit (Project B)
N+2	21-80	Real	Six 10 character values as in the previous card. Expected cash flow extended for each year after the N values input above
N+3	20	Integer	Number of interest rates to consider. Maximum is 5.
	30-39	Real	Interest rate in percent
	40-49	Real	Interest rate in percent
	50-59	Real	Interest rate in percent
	60-69	Real	Interest rate in percent
	70-79	Real	Interest rate in percent
N+4	15	Integer	Number of end of projects to be considered. Maximum is 5

Card	Columns	Mode	Value
	16-20 21-25 26-30 31-35 36-40	Integer	Index of the number of cash flow cards that represents the end of project. These indices may be larger than N, the number of cash flow cards, in which case the future cash flow (card N+2) is repeated for each additional year.
N+5	19-26	Real	Rate of inflation (percent)
N+6	1-3	Alpha-betic	Type of interest CON for continuous END for end of year

The last two cards use the free format concept. Each numeric field may be in any columns. Each field is separated by a comma.

Card	Field	Mode	Value
N+7	1	Integer	Which project to start analysis 1 = Project A 2 = Project B 3 = Project A-B
	2	Integer	Which project to end analysis 1 = Project A 2 = Project B 3 = Project A-B
	3	Integer	Frequency table print control 0 = Don't print 1 = Do print
	4	Integer	Histogram print control 0 = Don't print 1 = Do print
N+8	1	Integer	Minimum number of samples in hundreds. That is, if 5 is specified, at least 500 random numbers will be generated
	2	Integer	Maximum number of samples in hundreds
	3	Real	Ratio for convergence test. If ρ_i is the average present worth for i hundred samples, convergence test is satisfied if:

$$\left| 1 - \frac{\rho_i}{\rho_{i-1}} \right| \leq \text{Ratio}$$

DISTRIBUTIONNo of
CopiesNo. of
CopiesOFFSITE

- 17 AEC Headquarters
Washington, D. C.
 F. P. Baranowski
 R. W. Bown
 S. T. Brewer
 P. C. Fine
 E. E. Fowler
 D. K. Geston
 G. C. Gough
 F. D. Haines
 G. W. Johnson
 J. S. Kelly
 R. W. A. LaGassie
 L. Murphy
 F. K. Pittman
 G. F. Quinn
 R. W. Ramsey
 D. Thomas
 M. J. Whitman

- 2 AECOP
Union Carbide Corp.
Oak Ridge, TN 37830
 J. Shacter
 R. A. Watrous

ONSITE

- 6 AEC Richland Operation Office
 O. J. Elgert
 A. G. Fremling
 J. E. Goodwin
 P. G. Holsted
 W. E. Lotz
 F. R. Standerfer

9 Atlantic Richfield Hanford Company

J. B. Fecht
 H. H. Hopkins
 R. D. Jensen
 C. G. Lindberg
 E. T. McIntyre
 J. R. Palmer
 L. M. Richards
 A. E. Smith
 W. G. Smith

6 Douglas United Nuclear

C. L. Abel
 D. H. Bangerter
 T. S. Dodge
 C. D. Harrington
 H. R. Kosmata
 L. W. Lang

2 Westinghouse Hanford

E. A. Evans
 P. L. Hofmann

97 Battelle-Northwest

E. L. Alpen
 J. M. Batch
 W. L. Bathke
 J. B. Burnham
 C. H. Bloomster
 R. S. Carter
 B. M. Cole
 D. L. Condotta
 D. E. Deonigi
 K. Drumheller
 R. L. Engel (50)
 E. A. Eschbach (10)

Battelle-Northwest (cont.)

J. W. Finnegan
J. C. Fox
R. L. Hooper
C. N. Knudsen
C. W. Lindenmeier
R. W. McKee
E. T. Merrill
R. E. Nightingale
D. E. Olesen
R. S. Paul
L. T. Pedersen
L. P. Perrigo
J. E. Rasmussen/C. A. Bennett
W. A. Reardon
W. D. Richmond
A. M. Schneider
A. J. Scott
R. E. Turley
G. L. Wilfert
K. E. Yandon
J. R. Young
Technical Information (5)
Technical Publications