CONSIDERING THE UNCERTAINTY OF ESTIMATES IN ENGINEERING ECONOMIC ANALYSIS

by 557

JON HOWARD ESHELMAN

B.S., Kansas State University, 1970

A MASTER'S REPORT

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Civil Engineering

KANSAS STATE UNIVERSITY

Manhattan, Kansas

1971

Approved by:

Monroe L. Funk

Major Professor

TABLE OF CONTENTS

LIST OF TABLES

LIST OF PLATES

LIST OF FIGURES

ACKNOWLEDGEMENT

I would like to acknowledge and publicly thank the Automotive Safety Foundation, now part of the Highway Users' Federation for Safety and Mobility, for their financial support without which I would have been unable to earn my Master's Degree at this time.

INTRODUCTION

Engineering economic analysis is an analysis technique which basically enables engineers to determine the financial implications of proposed capital investments. In order to perform the analysis, the engineer must make estimates of such things as project life, interest rate, initial cost, salvage value, operating cost, and maintenance cost. The traditional method of economic analysis, and the one still being used today, requires single-valued estimates of these variables -- initial cost = $654,510 , project life = 40 years, operating cost = $47,500 per year, and so forth. It is obvious that such things as initial cost can be estimated relatively accurately, whereas project life can only be guessed at in many cases. Under further scrutiny, however, a shadow of uncertainty can be cast over even the supposedly accurate estimates. Such unforeseen happenings as strikes, price of material increases, or construction difficulties could make even these estimates very inaccurate.

It seems that the variables involved in engineering economic analysis could be more realistically described by some form of probability distribution rather than by single-valued estimates. (Admittedly the distribution for initial cost would generally be much "narrower" than that for project life or salvage value.) If it is true that this approach is more realistic, then are we justified in continuing to ignore the uncertainty inherent in our estimates? There are two possible justifications for continuing to ignore uncertainty: (1) if the analysis were very cumbersome mathematically, or (2) if the change noted in the final result were insignificant. The speed and availability of computers answers the first attempted justification concerning mathematics. The second justification remains to be explored.

The primary purpose of this paper is to suggest an analysis technique which would allow engineers to quantitatively consider uncertainty in the variables which must be estimated when performing an engineering economic analysis. The technique which will be reviewed and suggested is general enough so that any probability distribution can be used and uncertainty in any number of variables can be handled. The difficulty involved in determining an appropriate probability distribution for the various variables is recognized but will not be treated in this paper. Admittedly the technique is somewhat academic until this can be done, but for development purposes these distributions are assumed to be given.

The secondary purpose of this paper is to demonstrate the analysis technique by means of an example analysis in which project life is taken to be the only variable with uncertainty in its estimate. Project life was chosen because it probably contains more uncertainty of estimate than any other variable, especially in the case of long-term capital investments which civil engineers and planners continually encounter. The conclusions from this example analysis will also answer, at least for this specific case, the second attempted justification discussed above, that of insignificant change in the final result.

DEVELOPMENT OF THE ANALYSIS TECHNIQUE

This technique(1) focuses on the calculation of the annual cost of capital recovery because annual cost is a popular and widely understood basis with which to analyze and compare proposed investments. The variables involved in the annual cost calculation which might be subject to various degrees of uncertainty in their estimate include such things as project life, interest rate, initial cost, salvage value, operating cost, maintenance cost, and possibly others.

In order to simplify the discussion which will follow, symbols are shown below in Table 1 along with the frequently used terms which they represent.

Explanation of Symbols

(Table 1.)

Term	Symbol
Annual Cost of Capital Recovery ($/yr.)	R
Project Life (yrs.)	n
Interest Rate (%/yr.)	i
Initial Cost ($)	P
Capital Recovery Factor	CRF
Salvage Value ($)	SV
Operating Cost ($/yr.)	OC
Maintenance Cost ($/yr.)	MC
Sinking Fund Factor	SFF

The capital recovery factor which is referred to is defined according to traditional economic analysis. It is a factor which, when multiplied by a present cost, converts that present cost into a uniform series of end-of-period payments. The formula is

$$(CRF\text{-}i\text{-}n) = \frac{i(1+i)^n}{(1+i)^n - 1} \quad .$$

The symbol (CRF-i-n) is a shorthand way of referring to the capital recovery factor for interest rate i and project life n.

The sinking fund factor which is referred to is also defined according to traditional economic analysis. It is a factor which, when multiplied by a future sum of money, converts that future sum into a uniform series of end-of-period payments. Its formula is

$$(SFF\text{-}i\text{-}n) = \frac{i}{(1+i)^n - 1} \quad .$$

These two formulas for the CRF and the SFF are derived in Appendix B.

The starting point of this analysis is the traditional engineering economic analysis approach in which exact estimates are made for all of the variables. The general formula is

$$R = P\ (CRF\text{-}i\text{-}n) - SV\ (SFF\text{-}i\text{-}n) + OC + MC \quad .$$

This approach assumes that there is one and only one combination of values, which is true if the estimates are truly exact. But if some of the variables can assume different values, there may be ten possible combinations of values, or a hundred, or a thousand. The basic idea behind this analysis was to repeat the above calculation for annual cost for each possible combination of values, weighting the result according to its importance (according to the probability of that combination of values occurring), and averaging the weighted results. The final result was to be an "expected value of the annual cost of capital recovery".

Mathematical expectation is a common statistical calculation. In laymen's terms, the expected value of a random variable (X) is the sum of all possible values assumed by X, each multiplied by the probability of X assuming that value. Or in correct mathematical language[2]: "If $(X_1, X_2, X_3, \ldots)$ is the range of a random variable which assumes the value X_i with the probability $p(X_i)$, then the mathematical expectation of X, or the expected value of X, is

$$E(X) = \sum X_i \; p(X_i) \qquad ."$$

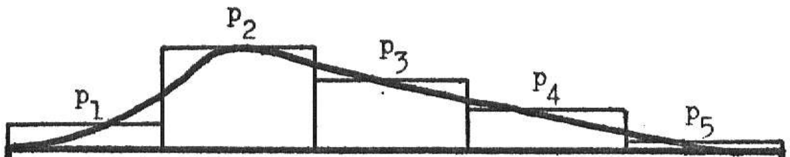
Uncertainty in the Estimate of One Variable

The probability distribution which describes the variability of the true value of the variable would actually be a continuous distribution of some sort. But for this analysis that continuous distribution is to be approximated by a step function. Theoretically the step function can be made to approximate the distribution to any degree of accuracy required by increasing the number of intervals. For the sake of simplicity, a five-step function was used in the development which follows.

The probability of occurrence matrix shows how many combinations of variable values exist and the probability of each possible occurrence. The distribution of the uncertain variable is shown along the top of the matrix (Table 2.), (Remember that this distribution is assumed to be given.) along with the five-step function which is intended to approximate it. Thus X_1, X_2, X_3, X_4, and X_5 are the five possible values, or more correctly, five representative values out of a multitude of possible values which the uncertain variable can assume. And p_1, p_2, p_3, p_4, p_5

Probability of Occurrence Matrix (one-variable)

(Table 2.)



Uncertain Variable				
X_1	X_2	X_3	X_4	X_5
p_1	p_2	p_3	p_4	p_5
$\sum = 1.00$				

are the probabilities associated with each possible value X_i of the variable. This matrix accounts for every possible outcome, therefore the sum of the probabilities equals 1.00 . Since all other variables are single-valued, there are five and only five possible combinations of circumstances which result in five different computations for annual cost.

Annual Cost Matrix (one-variable)

(Table 3.)

Uncertain Variable				
X_1	X_2	X_3	X_4	X_5
R_1	R_2	R_3	R_4	R_5

The annual cost matrix shows the value for the annual cost of capital recovery (R) for each of the five combinations of variables discussed immediately above. X_i is defined exactly the same as it was for the probability of occurrence matrix. To repeat the formula for annual cost, it is

$$R = P\ (CRF\text{-}i\text{-}n) - SV\ (SFF\text{-}i\text{-}n) + OC + MC \quad .$$

Expected Value of Annual Cost Matrix (one-variable)

(Table 4.)

Uncertain Variable				
X_1	X_2	X_3	X_4	X_5
p_1R_1	p_2R_2	p_3R_3	p_4R_4	p_5R_5
$\sum = E(R)$				

The expected value of annual cost matrix shows just what its name implies. The entries in the matrix were computed by multiplying the respective entries in the probability of occurrence matrix and the annual cost matrix. For clarity the definitional formula of expected value, using the symbols which have been employed in this discussion, is shown below.

$$E(R) = \sum R_i\ p_i$$

Variance of Annual Cost Matrix (one-variable)

(Table 5.)

Uncertain Variable				
X_1	X_2	X_3	X_4	X_5
$p_1(R_1)^2$	$p_2(R_2)^2$	$p_3(R_3)^2$	$p_4(R_4)^2$	$p_5(R_5)^2$
$\sum = E(R^2)$				

$$V_R = E(R^2) - (E(R))^2$$

The variance of annual cost matrix just shows the expected value of annual cost squared (R^2). The variance (V_R) is then computed from the formula shown just below the matrix. For the derivation of this formula, see Appendix C. The expected value of R^2 is analogous to the expected value of R, as shown here.

$$E(R^2) = \sum (R_i)^2 p_i$$

The standard deviation, which is the statistic that will actually be used most of the time, is defined as the positive square root of the variance. (Standard deviation = SD)

Uncertainty in the Estimate of Two Variables

Generalizing somewhat, the next case considered is that in which there is uncertainty in the estimates for two variables, each distribution once again being approximated by a five-step function.

The probability of occurrence matrix was discussed thoroughly in the previous analysis, but a few things do need to be pointed out in Table 6 for the case of uncertainty in the estimate of two variables. The p_1 associated with X_1 and the p_1 associated with Y_1 are not necessarily equal probabilities, which the notation might lead one to believe. This notation was used to simplify the explanation later in the paper. The probabilities associated with X sum to 1.00, as do the probabilities associated with Y. The sum of all entries in the matrix also sum to 1.00 because the 25 matrix elements account for every possible occurrence when two variables are allowed to assume any one of five values, and the sum of probabilities after every possible outcome has been accounted for must be 1.00. The entries in this matrix are computed by multiplying the probability associated with a particular row and the probability associated with a particular column. ($p_{23} = p_2(X) \cdot p_3(Y)$) This is consistent with the laws of probability for the simultaneous occurrence of two independent events,[3] which state that

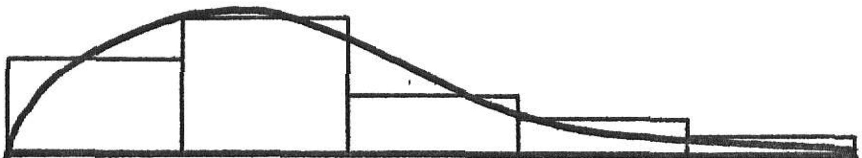
$$P(A \cap B) = P(A) \times P(B)$$

or

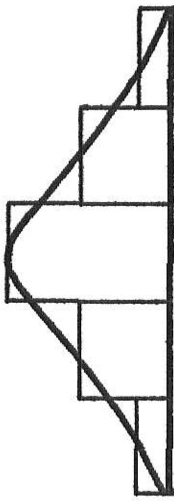
$$p(X_2 \cap Y_3) = p_2(X) \times p_3(Y) \quad .$$

Probability of Occurrence Matrix (two-variables)

(Table 6.)



			Uncertain Variable				
			Y_1	Y_2	Y_3	Y_4	Y_5
			p_1	p_2	p_3	p_4	p_5
Uncertain Variable	X_1	p_1'	p_{11}	p_{12}	p_{13}	p_{14}	p_{15}
	X_2	p_2'	p_{21}	p_{22}	p_{23}	p_{24}	p_{25}
	X_3	p_3'	p_{31}	p_{32}	p_{33}	p_{34}	p_{35}
	X_4	p_4'	p_{41}	p_{42}	p_{43}	p_{44}	p_{45}
	X_5	p_5'	p_{51}	p_{52}	p_{53}	p_{54}	p_{55}
			$\Sigma = 1.00$				



Annual Cost Matrix (two-variables)

(Table 7.)

		Uncertain Variable				
		Y_1	Y_2	Y_3	Y_4	Y_5
Uncertain Variable	X_1	R_{11}	R_{12}	R_{13}	R_{14}	R_{15}
	X_2	R_{21}	R_{22}	R_{23}	R_{24}	R_{25}
	X_3	R_{31}	R_{32}	R_{33}	R_{34}	R_{35}
	X_4	R_{41}	R_{42}	R_{43}	R_{44}	R_{45}
	X_5	R_{51}	R_{52}	R_{53}	R_{54}	R_{55}

The annual cost matrix represents the results of 25 computations of annual cost using the 25 possible combinations of variable values.

Expected Value of Annual Cost Matrix (two-variables)

(Table 8.)

		Uncertain Variable				
		Y_1	Y_2	Y_3	Y_4	Y_5
Uncertain Variable	X_1	$p_{11}R_{11}$	$p_{12}R_{12}$	$p_{13}R_{13}$	$p_{14}R_{14}$	$p_{15}R_{15}$
	X_2	$p_{21}R_{21}$	$p_{22}R_{22}$	$p_{23}R_{23}$	$p_{24}R_{24}$	$p_{25}R_{25}$
	X_3	$p_{31}R_{31}$	$p_{32}R_{32}$	$p_{33}R_{33}$	$p_{34}R_{34}$	$p_{35}R_{35}$
	X_4	$p_{41}R_{41}$	$p_{42}R_{42}$	$p_{43}R_{43}$	$p_{44}R_{44}$	$p_{45}R_{45}$
	X_5	$p_{51}R_{51}$	$p_{52}R_{52}$	$p_{53}R_{53}$	$p_{54}R_{54}$	$p_{55}R_{55}$
		$\Sigma = E(R)$				

The elements in the expected value of annual cost matrix are the product of the respective elements in the probability of occurrence matrix and the annual cost matrix.

Variance of Annual Cost Matrix (two-variables)

(Table 9.)

		Uncertain Variable				
		Y_1	Y_2	Y_3	Y_4	Y_5
Uncertain Variable	X_1	$p_{11}(R_{11})^2$	$p_{12}(R_{12})^2$	$p_{13}(R_{13})^2$	$p_{14}(R_{14})^2$	$p_{15}(R_{15})^2$
	X_2	$p_{21}(R_{21})^2$	$p_{22}(R_{22})^2$	$p_{23}(R_{23})^2$	$p_{24}(R_{24})^2$	$p_{25}(R_{25})^2$
	X_3	$p_{31}(R_{31})^2$	$p_{32}(R_{32})^2$	$p_{33}(R_{33})^2$	$p_{34}(R_{34})^2$	$p_{35}(R_{35})^2$
	X_4	$p_{41}(R_{41})^2$	$p_{42}(R_{42})^2$	$p_{43}(R_{43})^2$	$p_{44}(R_{44})^2$	$p_{45}(R_{45})^2$
	X_5	$p_{51}(R_{51})^2$	$p_{52}(R_{52})^2$	$p_{53}(R_{53})^2$	$p_{54}(R_{54})^2$	$p_{55}(R_{55})^2$
		$\sum = E(R^2)$				

$$V_R = E(R^2) - (E(R))^2$$

The variance of annual cost matrix follows immediately and logically the expected value matrix. The theory was previously explained.

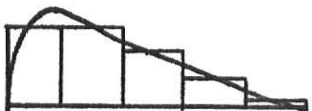
<u>Uncertainty in the Estimate of Three Variables</u>

At this point it begins to become difficult to visualize what is occurring. The two-dimensional matrix is no longer adequate and the analysis changes to a three-dimensional array -- a triple subscript array.

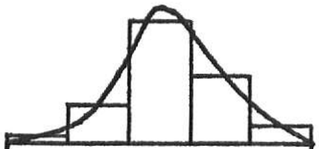
Probability of Occurrence Array (three-variables)

(Table 10.)

Uncertain Variable				
X_1	X_2	X_3	X_4	X_5
p_1	p_2	p_3	p_4	p_5
$\Sigma = 1.00$				



Uncertain Variable				
Y_1	Y_2	Y_3	Y_4	Y_5
p_1'	p_2'	p_3'	p_4'	p_5'
$\Sigma = 1.00$				



Uncertain Variable				
Z_1	Z_2	Z_3	Z_4	Z_5
p_1''	p_2''	p_3''	p_4''	p_5''
$\Sigma = 1.00$				

p_{111}	p_{112}	p_{113}	p_{114}	p_{115}
p_{121}	p_{122}	p_{123}	p_{124}	p_{125}
p_{131}	p_{132}	p_{133}	p_{134}	p_{135}
p_{141}	p_{142}	p_{143}	p_{144}	p_{145}
p_{151}	p_{152}	p_{153}	p_{154}	p_{155}
p_{211}	p_{212}	p_{213}	p_{214}	p_{215}
p_{221}	p_{222}	p_{223}	p_{224}	p_{225}
⋮	⋮	⋮	⋮	⋮
p_{541}	p_{542}	p_{543}	p_{544}	p_{545}
p_{551}	p_{552}	p_{553}	p_{554}	p_{555}
$\Sigma = 1.00$				

Once again, the probabilities associated with X,Y, and Z are not necessarily equal even though the notation may suggest that they are. The theory of probability concerning simultaneous events was discussed during the previous analysis, so it is not repeated here. The elements in this probability of occurrence array are the product of three probabilities -- one associated with X, one with Y, and one with Z. (p_{XYZ}) ($p_{543} = p_5(X) \times p_4(Y) \times p_3(Z)$) The 125 elements in this array account for all possible combinations of X,Y, and Z. Once again the sum of all probabilities equals 1.00 .

Annual Cost Array (three-variables)

(Table 11.)

R_{111}	R_{112}	R_{113}	R_{114}	R_{115}
R_{121}	R_{122}	R_{123}	R_{124}	R_{125}
R_{131}	R_{132}	R_{133}	R_{134}	R_{135}
R_{141}	R_{142}	R_{143}	R_{144}	R_{145}
R_{151}	R_{152}	R_{153}	R_{154}	R_{155}
R_{211}	R_{212}	R_{213}	R_{214}	R_{215}
R_{221}	R_{222}	R_{223}	R_{224}	R_{225}
⋮	⋮	⋮	⋮	⋮
R_{541}	R_{542}	R_{543}	R_{544}	R_{545}
R_{551}	R_{552}	R_{553}	R_{554}	R_{555}

The annual cost array represents the results of 125 computations of annual cost using the 125 possible combinations of variable values.

The elements in the expected value of annual cost array (Table 12.) are the product of the respective elements in the probability of occurrence array and the annual cost array.

Expected Value of Annual Cost Array (three-variables)

(Table 12.)

$^{P}111^{R}111$	$^{P}112^{R}112$	$^{P}113^{R}113$	$^{P}114^{R}114$	$^{P}115^{R}115$
$^{P}121^{R}121$	$^{P}122^{R}122$	$^{P}123^{R}123$	$^{P}124\ ^{R}124$	$^{P}125^{R}125$
$^{P}131^{R}131$	$^{P}132^{R}132$	$^{P}133^{R}133$	$^{P}134^{R}134$	$^{P}135^{R}135$
,	,	,	,	,
,	,	,	,	,
,	,	,	,	,
$^{P}541^{R}541$	$^{P}542^{R}542$	$^{P}543^{R}543$	$^{P}544^{R}544$	$^{P}545^{R}545$
$^{P}551^{R}551$	$^{P}552^{R}552$	$^{P}553^{R}553$	$^{P}554^{R}554$	$^{P}555^{R}555$
		$\sum = E(R)$		

The variance of annual cost array follows immediately and logically the expected value array.

Variance of Annual Cost Array (three-variables)

(Table 13.)

$^{P}111^{(R}111^{)2}$	$^{P}112^{(R}112^{)2}$	$^{P}113^{(R}113^{)2}$	$^{P}114^{(R}114^{)2}$	$^{P}115^{(R}115^{)2}$
$^{P}121^{(R}121^{)2}$	$^{P}122^{(R}122^{)2}$	$^{P}123^{(R}123^{)2}$	$^{P}124(^{R}124^{)2}$	$^{P}125^{(R}125^{)2}$
$^{P}131^{(R}131^{)2}$	$^{P}132^{(R}132^{)2}$	$^{P}133^{(R}133^{)2}$	$^{P}134^{(R}134^{)2}$	$^{P}135^{(R}135^{)2}$
,	,	,	,	,
,	,	,	,	,
,	,	,	,	,
$^{P}541^{(R}541^{)2}$	$^{P}542^{(R}542^{)2}$	$^{P}543^{(R}543^{)2}$	$^{P}544^{(R}544^{)2}$	$^{P}545^{(R}545^{)2}$
$^{P}551^{(R}551^{)2}$	$^{P}552^{(R}552^{)2}$	$^{P}553^{(R}553^{)2}$	$^{P}554^{(R}554^{)2}$	$^{P}555^{(R}555^{)2}$
		$\sum = E(R^2)$		

$$V_R = E(R^2) - (E(R))^2$$

Extension to Uncertainty in Any Number of Variables

From this last analysis, allowing uncertainty in three variables, it can easily be seen how to expand the analysis to allow uncertainty in more variables. Every additional variable which contains uncertainty would increase the number of subscripts (the dimension of the array) by one also.

If N equals the number of steps in the step function used to approximate the probability distributions of the uncertain variables (in the above case, 5) and w equals the number of uncertain variables, then

$$N^w = \text{number of elements in the array.}$$

When N = 5, the progression is 5, 25, 125, 625, 3125, 15625 . When N = 10, the progression is 10, 100, 1000, 10000, 100000, 1000000 . The number of computations can get very large very quickly. To answer this problem there has been an approximation technique developed in which a random number generator is used as a tool for generating combinations of variable values which, over the course of many iterations, will approximately reflect the possible outcomes of our initial probability distributions, i.e., will approximately reflect the possible outcomes described by the probability of occurrence array.

REVIEW OF AN APPROXIMATION TECHNIQUE

The beginning point of this approximation technique is the same as for the theoretical expected value technique -- the probability distributions for the variables must be known.[4] The continuous distribution must then be approximated by a step function exactly as before. In this example the distribution is approximated by a four-step function and there are five uncertain variables. The following table shows the possible values for each variable along with the probability of each value occurring.

Demonstration Problem

Example Distributions

(Table 14.)

	Value	Prob.	Value	Prob.	Value	Prob.	Value	Prob.
Variable A	A_1	0.1	A_2	0.2	A_3	0.1	A_4	0.6
Variable B	B_1	0.6	B_2	0.2	B_3	0.1	B_4	0.1
Variable C	C_1	0.2	C_2	0.4	C_3	0.3	C_4	0.1
Variable D	D_1	0.1	D_2	0.1	D_3	0.5	D_4	0.3
Variable E	E_1	0.3	E_2	0.3	E_3	0.2	E_4	0.2

Next a digit value is assigned to each variable value according to its probability of occurrence.

Example Digit Assignments

(Table 15.)

Variable	Digit Value	Value	Digit Value	Value	Digit Value	Value	Digit Value	Value
1st digit (A)	1	A_1	2,3	A_2	4	A_3	5,6,7,8,9,0	A_4
2nd digit (B)	1,2,3,4,5,6	B_1	7,8	B_2	9	B_3	0	B_4
3rd digit (C)	1,2	C_1	3,4,5,6	C_2	7,8,9	C_3	0	C_4
4th digit (D)	1	D_1	2	D_2	3,4,5,6,7	D_3	8,9,0	D_4
5th digit (E)	1,2,3	E_1	4,5,6	E_2	7,8	E_3	9,0	E_4

To understand this table more clearly, look at the bottom row. Note that there are ten digits in this row (as there are in all rows), each representing one "unit of probability", since the probability was measured in tenths. Variable value E_1 has three digits assigned to it because E_1 has a probability of occurrence of 0.3 . E_2 has three digits assigned to it because E_2 also has a probability of occurrence of 0.3 . E_3 has two digits assigned to it because E_3 has a probability of occurrence of 0.2 . Likewise for E_4. And likewise for every other row. This table is actually just a "coded" form of the first table. The left column does demand some explanation, however. It was decided that the first random number generated would be used to determine which value of A would be used, according to which digit value it corresponded with; the second random number generated would be used to determine which value of B would be used, according to which digit value it corresponded with; and so forth until the fifth random number generated determined which value of E was to be used. Thus, the headings "1st digit", "2nd digit", and so on, refer to the order in which the random numbers are generated.

To begin the analysis five random numbers are generated, which in turn determine the values of A,B,C,D, and E to be used, which in turn allows us to compute the annual cost. For example, if the five random numbers generated were

$$2 \quad 1 \quad 5 \quad 5 \quad 9$$

then the variable values to be used would be

$$A_2 \quad B_1 \quad C_2 \quad D_3 \quad E_4$$

and these values could be used to compute the annual cost (R). If this iterative process were repeated many times, one can see that the values with high probability will be chosen more frequently and eventually a good approximation will result for the theoretical probabilities of simultaneous occurrence. This process, in effect, attempts to simulate experience. The more trials made, the closer the approximation of the theoretical value.

For further assurance in the value selection process, below are shown several sets of random numbers along with the variable values they describe.

2 1 0 8 1	A_2 B_1 C_4 D_4 E_1
1 0 4 6 8	A_1 B_4 C_2 D_3 E_3
0 8 3 0 2	A_4 B_2 C_2 D_4 E_1

After a hundred iterations, for instance, there would be a hundred values of annual cost. The arithmetic average of these annual costs would be the approximation for the expected value of the annual cost. Smith suggested that perhaps 1000 iterations would be enough to make a very good approximation.[4]

EXAMPLE ANALYSIS

In order to demonstrate the analysis technique which was developed in this paper, an example analysis follows which shows the effect that uncertainty in the estimate of project life has on the computation of annual cost.

Assumptions

A normal distribution was chosen with which to describe the possible outcomes of project life for several reasons: (1) It seems reasonable that the true life could be either above or below the average life with equal probability. (2) It seems reasonable that the true life would have a high probability of being relatively near the average life. (3) The normal distribution satisfies the above conditions and is at the same time easy to handle mathematically.

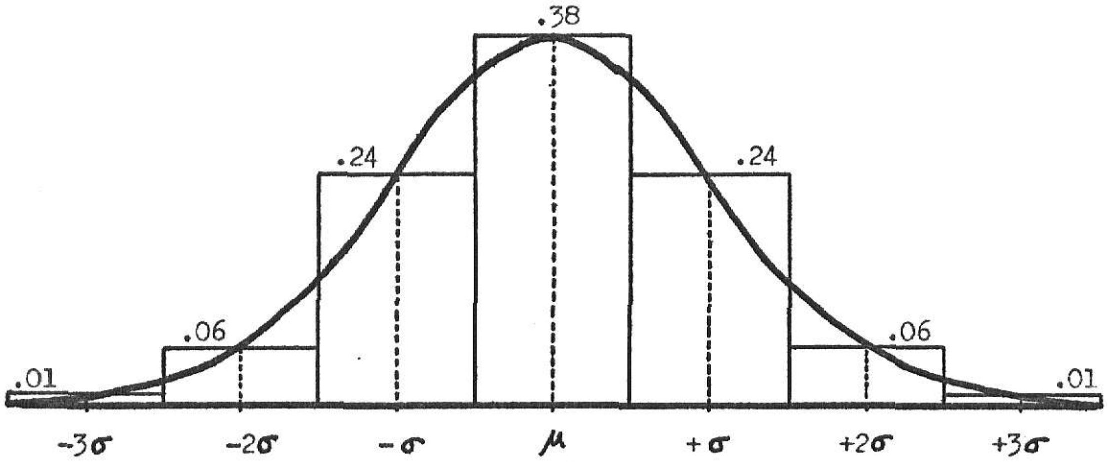
A seven-step function was chosen to approximate the normal distribution. See Fig. 1 for a graphical representation of the normal curve and the seven-step approximation function.

Calculations

For every combination of mean life, degree of variability, and interest rate, there is a value for annual cost (computed by traditional analysis), a value for

Normal Distribution and Step Function Approximation

(Figure 1.)



expected value of annual cost, and a value for variance (or standard deviation) of annual cost. The intention of this example analysis was to investigate combinations of mean life, degree of variability, and interest rate such that the entire range of possible outcomes was sampled.

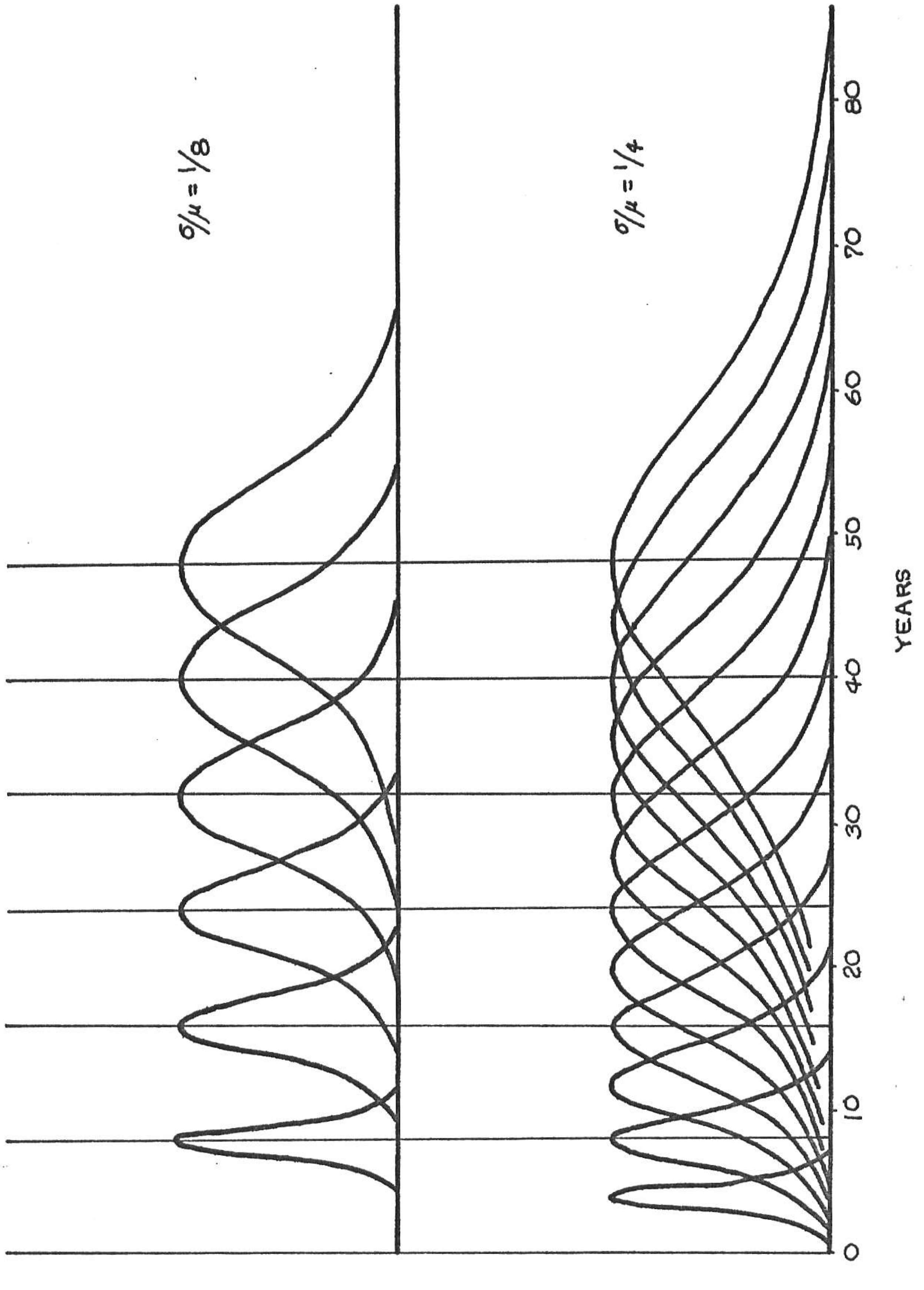
σ/μ was used as the measure of the degree of variability. $\sigma/\mu = 1/4$ was found to be the widest distribution which could be used with the seven-step function and not have the extreme left "tail" of the distribution extending into the range of negative project life. In addition to $\sigma/\mu = 1/4$, a degree of variability exactly half that was used, $\sigma/\mu = 1/8$. It was found that "tighter" distributions than this resulted in almost no difference between the traditional annual cost figures and the expected value of annual cost figures. The graphical comparison of the two degrees of variability, $\sigma/\mu = 1/4$ and $\sigma/\mu = 1/8$, is shown in Plate I.

Results

Shown in Plates II and III are the combinations of mean life, degree of

Graphical Comparison of the Two Degrees of Variability

(Plate I.)


σ/μ = 1/8
σ/μ = 1/4
0
10
20
30
40
50
60
70
80
YEARS

variability, and interest rate which were considered, along with the results computed for each combination.

Appendix D shows an example work sheet and the calculations involved with each combination of life, variability, and interest rate. Appendix E contains graphs of annual cost versus mean life for various interest rates and degrees of variability.

Discussion

It can be seen from Plates II and III that the maximum difference between the expected value of annual cost and the traditional value of annual cost (expressed as a percentage of expected annual cost) was 8.9 percent. This occurred with $i = 0\%$ and $\sigma/\mu = 1/4$ (the lowest possible interest rate and the widest possible distribution). Furthermore, as interest rate increased and/or mean life increased, the percent difference in annual costs decreased. This fact is significant because in the range of higher interest rates and longer mean lives which would be appropriate for long-term capital investments, the distinction between traditional annual cost and expected annual cost becomes merely academic and insignificant.

The standard deviation of annual cost as a percent of expected annual cost was a maximum of 39.2 percent at the same combination of $i = 0\%$ and $\sigma/\mu = 1/4$ as was just described above. This standard deviation as a percent of expected annual cost decreased in a manner similar to that just described for percent difference in annual cost figures.

The + or - one standard deviation interval plotted on the graphs of Appendix E can only be interpreted subjectively. They indicate generally that larger percentage differences in annual costs and larger percentage variance in annual cost go hand in hand. This further serves to discourage the idea that traditional

Analysis Results ($\sigma/\mu = 1/4$)

(Plate II.)

σ/μ	i (PERCENT)	μ (YEARS)	R TRADITIONAL ANNUAL COST OF CAPITAL RECOVERY (DOLLARS PER YEAR)	E(R) EXPECTED VALUE OF ANNUAL COST (DOLLARS PER YEAR)	$\frac{E(R)-R}{E(R)} \cdot 100$ (PERCENT)	S.D. STANDARD DEVIATION OF ANNUAL COST (DOLLARS PER YEAR)	$\frac{S.D.}{E(R)} \cdot 100$ (± PERCENT)
1/4	0	4	250,000	274,428	8.9	107,592	39.2
		8	125,000	137,215	8.9	53,797	39.2
		12	83,330	91,475	8.9	35,864	39.2
		16	62,500	68,606	8.9	26,899	39.2
		20	50,000	54,887	8.9	21,518	39.2
		24	41,670	45,733	8.9	17,945	39.2
		28	35,710	39,204	8.9	15,371	39.2
		32	31,250	34,305	8.9	13,448	39.2
		36	27,780	30,494	8.9	11,955	39.2
		40	25,000	27,442	8.9	10,759	39.2
		44	22,730	24,947	8.9	9,781	39.2
		48	20,830	22,868	8.9	8,967	39.2
		52	19,230	21,108	8.9	8,277	39.2
1/4	3	4	269,030	293,550	8.4	108,584	37.0
		8	142,460	154,855	8.0	54,466	35.2
		12	100,460	108,726	7.6	36,197	33.3
		16	79,610	85,809	7.2	27,033	31.5
		20	67,220	72,176	6.9	21,511	29.8
		24	59,050	63,178	6.5	17,807	28.2
		28	53,290	56,830	6.2	15,146	26.7
		32	49,050	52,142	5.9	13,137	25.2
		36	45,800	48,548	5.7	11,561	23.8
		40	43,260	45,729	5.4	10,294	22.5
		44	41,230	43,469	5.2	9,251	21.3
		48	39,580	41,621	4.9	8,375	20.1
		52	38,220	40,099	4.7	7,621	19.0
1/4	10	4	315,470	341,099	7.5	112,170	32.9
		8	187,440	200,247	6.4	55,043	27.5
		12	146,760	155,266	5.5	35,603	22.9
		16	127,820	134,138	4.7	25,649	19.1
		20	117,460	122,434	4.1	19,542	16.0
		24	111,300	115,337	3.5	15,411	13.4
		28	107,450	110,787	3.0	12,427	11.2
		32	104,970	107,753	2.6	10,205	9.5
		36	103,340	105,675	2.2	8,484	8.0
		40	102,260	104,220	1.9	7,134	6.8
		44	101,530	103,181	1.6	6,051	5.9
		48	101,040	102,428	1.4	5,173	5.1
		52	100,710	101,878	1.1	4,438	4.4

CONTINUED

1/4	20	4	386,290	413,066	6.5	115,343	27.9
		8	260,610	273,855	4.8	54,049	19.7
		12	225,260	233,772	3.6	32,712	14.0
		16	211,440	217,360	2.7	21,827	10.0
		20	205,360	209,586	2.0	15,364	7.3
		24	202,550	205,589	1.5	11,215	5.5
		28	201,220	203,405	1.1	8,409	4.1
		32	200,590	202,164	0.8	6,454	3.2

Analysis Results ($\sigma/\mu = 1/8$)

(Plate III.)

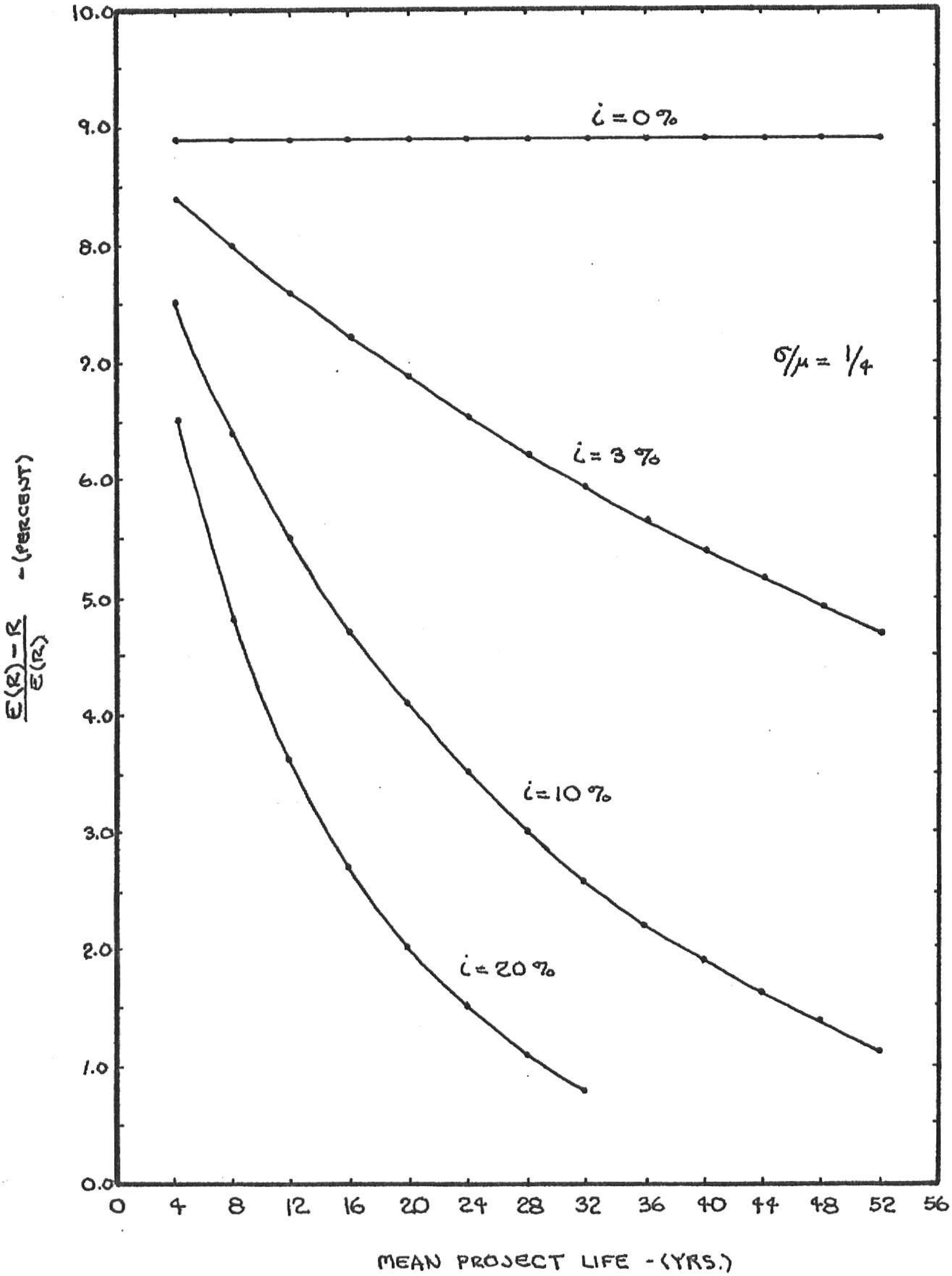
σ/μ	i (PERCENT)	μ (YEARS)	R TRADITIONAL ANNUAL COST OF CAPITAL RECOVERY (DOLLARS PER YEAR)	E(R) EXPECTED VALUE OF ANNUAL COST (DOLLARS PER YEAR)	$\frac{E(R)-R}{E(R)} \cdot 100$ (PERCENT)	S.D. STANDARD DEVIATION OF ANNUAL COST (DOLLARS PER YEAR)	$\frac{S.D.}{E(R)} \cdot 100$ (± PERCENT)
1/8	0	8	125,000	127,361	1.9	18,082	14.2
		16	62,500	63,682	1.9	9,037	14.2
		24	41,670	42,458	1.9	6,020	14.2
		32	31,250	31,839	1.9	4,522	14.2
		40	25,000	25,472	1.9	3,612	14.2
		48	20,830	21,226	1.9	3,013	14.2
		56	17,860	18,196	1.9	2,583	14.2
1/8	3	8	142,460	144,855	1.7	18,270	12.6
		16	79,610	80,810	1.5	9,020	11.2
		24	59,050	59,846	1.3	5,888	9.8
		32	49,050	49,644	1.2	4,290	8.6
		40	43,260	43,740	1.1	3,302	7.5
		48	39,600	39,983	1.0	2,638	6.6
		56	37,080	37,415	0.9	2,162	5.8
1/8	10	8	187,440	189,919	1.3	18,158	9.6
		16	127,820	129,037	0.9	7,991	6.2
		24	111,300	112,066	0.7	4,360	3.9
		32	104,970	105,483	0.5	2,498	2.4
		40	102,260	102,603	0.3	1,429	1.4
		48	101,040	101,266	0.2	882	0.9
		56	100,480	100,628	0.1	459	0.5
1/8	20	8	260,610	263,161	1.0	16,964	6.4
		16	211,440	212,535	0.5	5,532	2.6
		24	202,550	203,042	0.2	2,801	1.4
		32	200,590	200,805	0.1	695	0.3

annual cost and expected annual cost are significantly different because increasing difference is accompanied by increasing uncertainty and variability.

This whole discussion of changing percentages is summarized and illustrated in Plates IV, V, VI, and VII. From these graphs one can determine approximately the percent difference between the traditional and expected annual costs, or the percent the standard deviation is of the expected annual cost; either one may be found if it is known what the mean life is, what interest rate is being used, and what degree of variability applies ($\sigma/\mu = 1/4$ or $\sigma/\mu = 1/8$).

Percent Difference in Annual Costs vs. Mean Life ($\sigma/\mu = 1/4$)

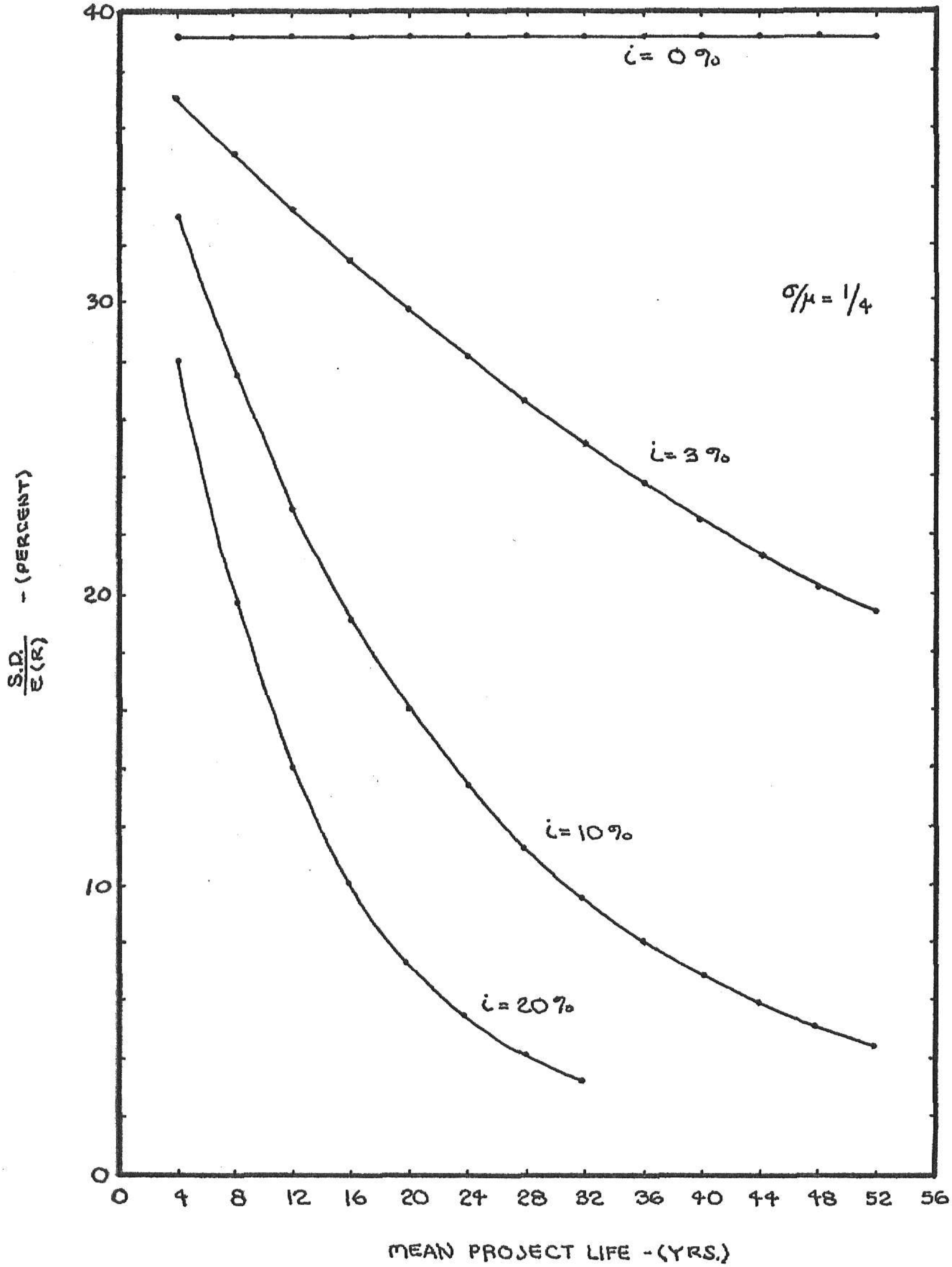
(Plate IV.)



10.0
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
0.0
0
4
8
12
16
20
24
28
32
36
40
44
48
52
56
$i = 0\%$
$i = 3\%$
$i = 10\%$
$i = 20\%$
$\sigma/\mu = 1/4$
$\frac{E(R) - R}{E(R)}$ - (PERCENT)
MEAN PROJECT LIFE - (YRS.)

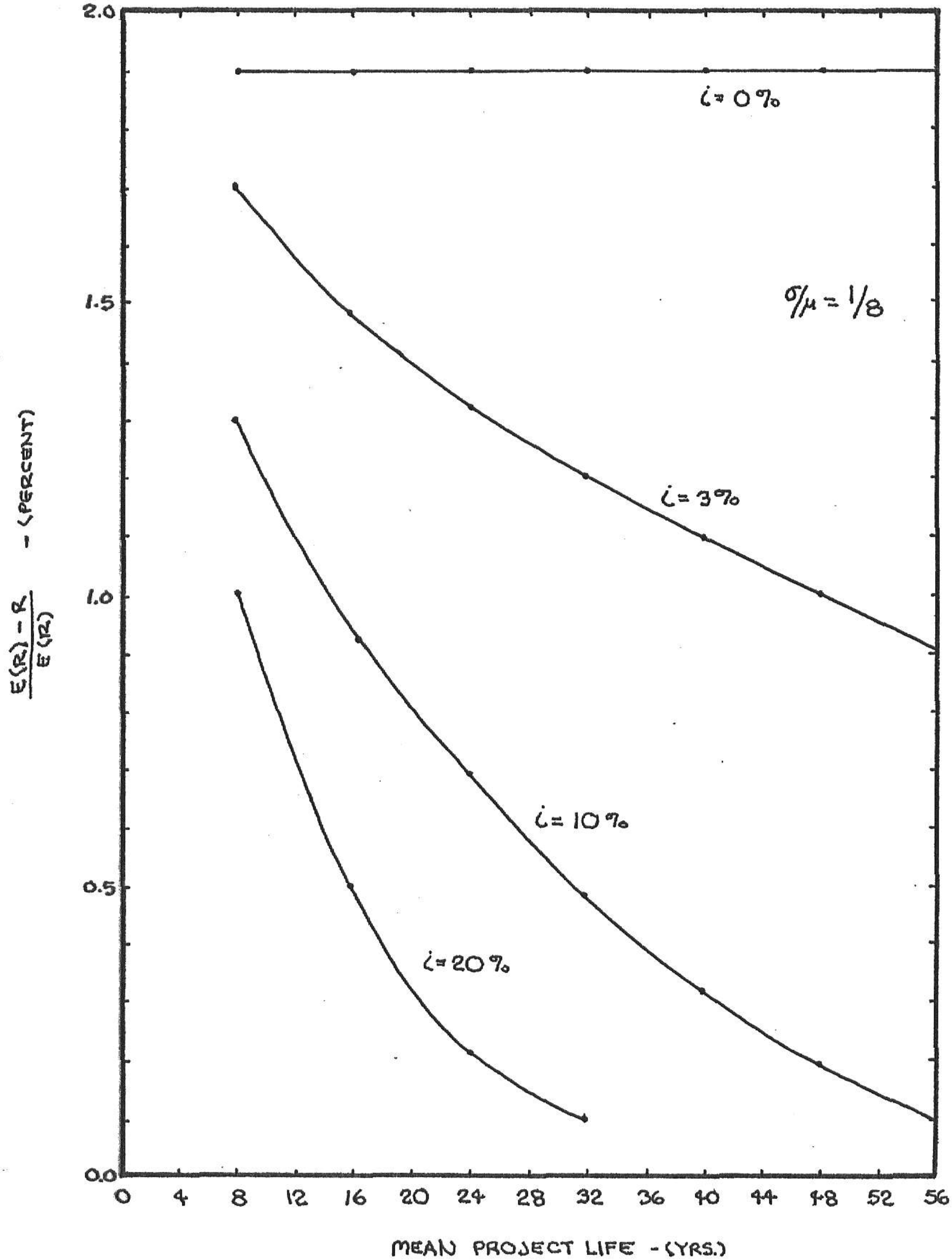
Standard Deviation as a Percent vs. Mean Life ($\sigma/\mu = 1/4$)

(Plate V.)


40
30
20
10
0
0 4 8 12 16 20 24 28 32 36 40 44 48 52 56
i = 0 %
i = 3 %
i = 10 %
i = 20 %
σ/μ = 1/4
S.D./E(R) - (PERCENT)
MEAN PROJECT LIFE - (YRS.)

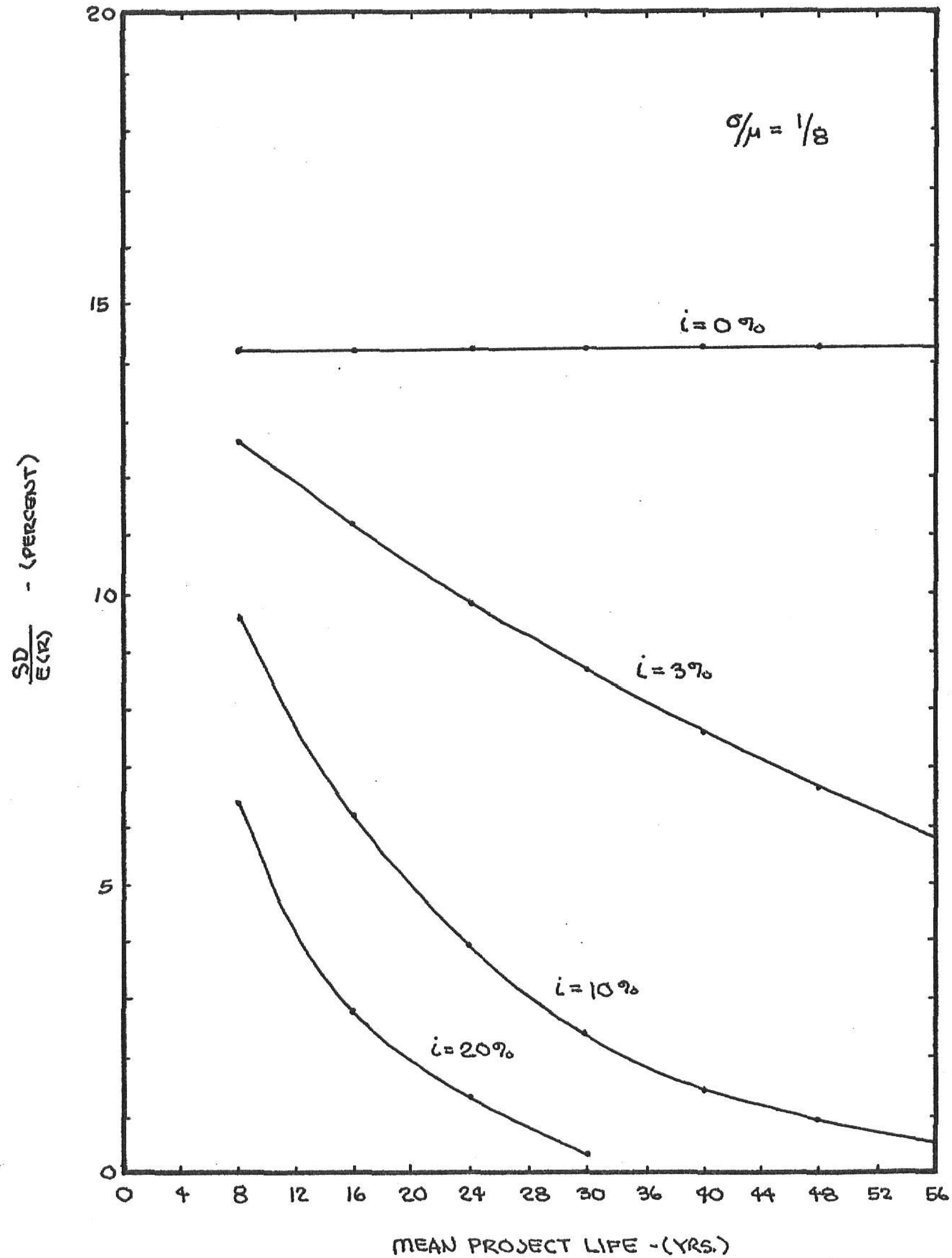
Percent Difference in Annual Costs vs. Mean Life ($\sigma/\mu = 1/8$)

(Plate VI.)


2.0
1.5
1.0
0.5
0.0
$\frac{E(R) - R}{E(R)}$ - (PERCENT)
0
4
8
12
16
20
24
28
32
36
40
44
48
52
56
MEAN PROJECT LIFE - (YRS.)
$i = 0\%$
$i = 3\%$
$i = 10\%$
$i = 20\%$
$\sigma/\mu = 1/8$

Standard Deviation as a Percent vs. Mean Life ($\sigma/\mu = 1/8$)

(Plate VII.)



20
15
10
5
0
$\sigma/\mu = 1/8$
$i = 0\%$
$i = 3\%$
$i = 10\%$
$i = 20\%$
$\frac{SD}{E(R)}$ - (PERCENT)
0 4 8 12 16 20 24 28 30 36 40 44 48 52 56
MEAN PROJECT LIFE - (YRS.)

CONCLUSION

The theoretical analysis for which the development is traced out in the first part of this paper was conceived as a means by which to consider uncertainty in the estimates of the variables involved in engineering economic analysis. The example analysis has shown that the effect of uncertainty is noticable and in some cases significant.

It was shown that for the particular case in which project life was the only uncertain variable and that uncertainty was described by a normal probability distribution, that the expected value of annual cost was always slightly higher than the annual cost computed by the traditional approach. This is entirely reasonable because the normal distribution is spread symmetrically about the mean, and, due to the nature of the annual cost method, project lives below the mean are more influential than the corresponding project lives above the mean, thus increasing the annual cost. Any time-value-of-money calculation discounts distant future costs and happenings and gives more relative importance to the present and the immediate future.

In the range of high interest rates and long project lives, which is the situation encountered in most capital investment proposals, it appears that it is unnecessary and time-wasteful to attempt to consider uncertainty in the estimate of project life; the final difference is insignificant.

As a general guide to engineers performing economic analysis, if (a) the only variable having considerable uncertainty is project life and (b) if a normal probability distribution is a reasonable choice to describe that uncertainty, then:

1. Estimate the mean project life;

2. Choose the degree of variability (σ/μ = 1/4 or 1/8) which seems most appropriate;

3. Determine from Plate IV or VI (depending on σ/μ) the approximate percent difference between traditional annual cost and expected annual cost. (Remember that expected annual cost is always the higher of the two.)

The percent difference determined in step 3 above may be small and may be judged to be insignificant, but the analyst must be aware of the possibility that when choosing between two proposed projects which are very nearly equal in annual cost, there may be enough uncertainty so that choosing the project with the slightly lower annual cost cannot be considered to be absolutely right. In this special case a more detailed analysis may be justified.

SUMMARY

The stated purpose of this paper was twofold: (1) to suggest an analysis technique which would allow engineers to quantitatively consider uncertainty in the variables which must be estimated when performing an engineering economic analysis, (2) to demonstrate an analysis technique by means of an example analysis in which project life is taken to be the only variable with uncertainty in its estimate.

The discussion and development in this paper began with the traditional economic analysis approach. First one, then two, then three variables were allowed to contain uncertainty, and finally the general case was discussed in which uncertainty was allowed in any number of variables.

An approximation technique was reviewed in which a random number generator was used as a tool to simulate real world experience.

The results of the example analysis described above were presented and discussed and general conclusions drawn.

REFERENCES

(1) Smith, Gerald W., Engineering Economy: Analysis of Capital Expenditures, Iowa State University Press, 1968, page 340

(2) Freund, John E., Mathematical Statistics, Prentice-Hall, Inc., 1962, page 91

(3) Freund, John E., Mathematical Statistics, Prentice-Hall, Inc., 1962, page 47

(4) Smith, Gerald W., Engineering Economy: Analysis of Capital Expenditures, Iowa State University Press, 1968, page 344

(5) Grant, Eugene L., Ireson, W. Grant, Principles of Engineering Economy, fourth edition, Ronald Press, 1964, page 45

(6) Meyer, Paul L., Introductory Probability and Statistical Applications, second edition, Addison-Wesley Publishing Company, 1970, page 135

BIBLIOGRAPHY

1. Freund, John E., Mathematical Statistics, Prentice-Hall, Inc., 1962

2. Grant, Eugene L., Ireson, W. Grant, Principles of Engineering Economy, fourth edition, Ronald Press, 1964

3. Meyer, Paul L., Introductory Probability and Statistical Applications, second edition, Addison-Wesley Publishing Company, 1970

4. Smith, Gerald W., Engineering Economy: Analysis of Capital Expenditures, Iowa State University Press, 1968

5. Spiegel, Murray R., Theory and Problems of Statistics, Schaum's Outline Series, McGraw-Hill Book Company, 1961

6. Woods, Baldwin M., Degarmo, E. Paul, Introduction to Engineering Economy, second edition, MacMillan Company, 1953

APPENDIX A - Notation

R = Annual cost of capital recovery ($/yr.)

n = Project life (yrs.)

i = Interest rate (%/yr.)

P = Initial cost ($)

CRF = Capital recovery factor

SV = Salvage value ($)

OC = Operating cost ($/yr.)

MC = Maintenance cost ($/yr.)

SFF = Sinking fund factor

p = Probability

V = Variance

SD = Standard deviation

APPENDIX B - Derivation of CRF and SFF

Reference (5)

Sinking Fund Factor

If P is invested at interest rate i, the interest for the first year is iP and the total amount at the end of the first year is $P + iP = P(1+i)$. The second year the interest on this is $iP(1+i)$, and the amount at the end of this year is $P(1+i) + iP(1+i) = P(1+i)^2$. Similarly, at the end of the third year the amount is $P(1+i)^3$; at the end of n years it is $P(1+i)^n$.

This is the formula for the compound amount, S, obtainable in n years from principal, P, which is shown below:

$$S = P(1+i)^n \quad . \tag{1}$$

If R is invested at the end of each year for n years, the total amount at the end of n years will obviously be the sum of the compound amounts of the individual investments. The money invested at the end of the first year will earn interest for (n-1) years; its amount will thus be $R(1+i)^{n-1}$. The second year's payment will amount to $R(1+i)^{n-2}$; the third year's to $R(1+i)^{n-3}$; and so on until the last payment, made at the end of n years, which has earned no interest. The total amount S is $R(1 + (1+i) + (1+i)^2 + (1+i)^3 + \ldots + (1+i)^{n-1})$.

This expression for S in terms of R may be simplified to its customary form by the following algebraic manipulations:

$$S = R(1 + (1+i) + (1+i)^2 + (1+i)^3 + \ldots + (1+i)^{n-2} + (1+i)^{n-1})$$

Multiplying both sides of the equation by (1+i), gives

$$(1+i)S = R(\ (1+i) + (1+i)^2 + (1+i)^3 + \ldots + (1+i)^{n-1} + (1+i)^n\ ) \quad .$$

Subtracting the original equation from this second equation gives

$$iS = R(\ (1+i)^n - 1) \quad ,$$

Then resulting in

$$R = S\frac{i}{(1+i)^n - 1} \quad . \qquad (2)$$

Capital Recovery Factor

To find the uniform end-of-year payment, R, which can be secured for n years from a present investment, P, substitute in equation (2) the value given for S in equation (1) ; thus

$$R = S\frac{i}{(1+i)^n - 1} \quad ,$$

$$R = P(1+i)^n\frac{i}{(1+i)^n - 1} \quad ,$$

$$R = P\frac{i(1+i)^n}{(1+i)^n - 1} \quad . \qquad (3)$$

APPENDIX C - Derivation of the Variance Formula

Reference (6)

Let X be a random variable. Variance of X, denoted V(X), is defined as

$$V(X) = E(X - E(X))^2 \tag{1}$$

The positive square root of V(X) is called the standard deviation of X and is denoted by S.D. .

Expanding equation (1) and using the properties for expectation gives

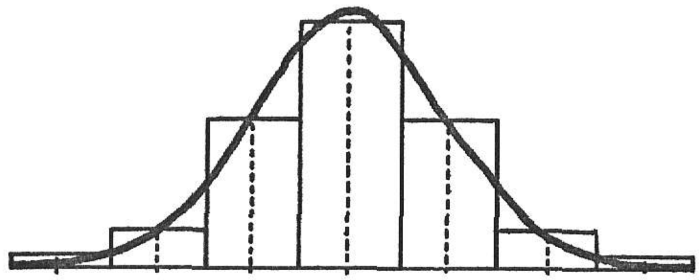
$$V(X) = E(X - E(X))^2$$

$$V(X) = E(X^2 - 2XE(X) + (E(X))^2)$$

$$V(X) = E(X^2) - 2E(X)E(X) + (E(X))^2$$

$$V(X) = E(X^2) - (E(X))^2$$

APPENDIX D - Example Work Sheet



i = 3 %

σ/μ = 1/4

PROBABILITY OF OCCURRENCE MATRIX

6	12	18	24	30	36	42
.01	.06	.24	.38	.24	.06	.01
Σ = 1.00						

ANNUAL COST MATRIX

.18460	.10046	.07271	.05905	.05102	.04580	.04219
6	12	18	24	30	36	42
184,600	100,460	72,710	*59,050	51,020	45,800	42,190

*Traditional Value of Annual Cost

EXPECTED VALUE OF ANNUAL COST MATRIX

6	12	18	24	30	36	42
1,846	6,028	17,450	22,439	12,245	2,748	422
Σ = E(R) = 63,178						

VARIANCE OF ANNUAL COST MATRIX

6	12	18	24	30	36	42
340,772,000	605,573,000	1,268,790,000	1,325,023,000	624,740,000	125,858,000	17,804,000
$\Sigma = E(R^2)$ = 4,308,560,000						

$$V_R = E(R^2) - (E(R))^2$$

$$E(R^2) = 4\,308\,560\,000$$

$$(E(R))^2 = 3\,991\,460\,000$$

$$V_R = 317\,100\,000$$

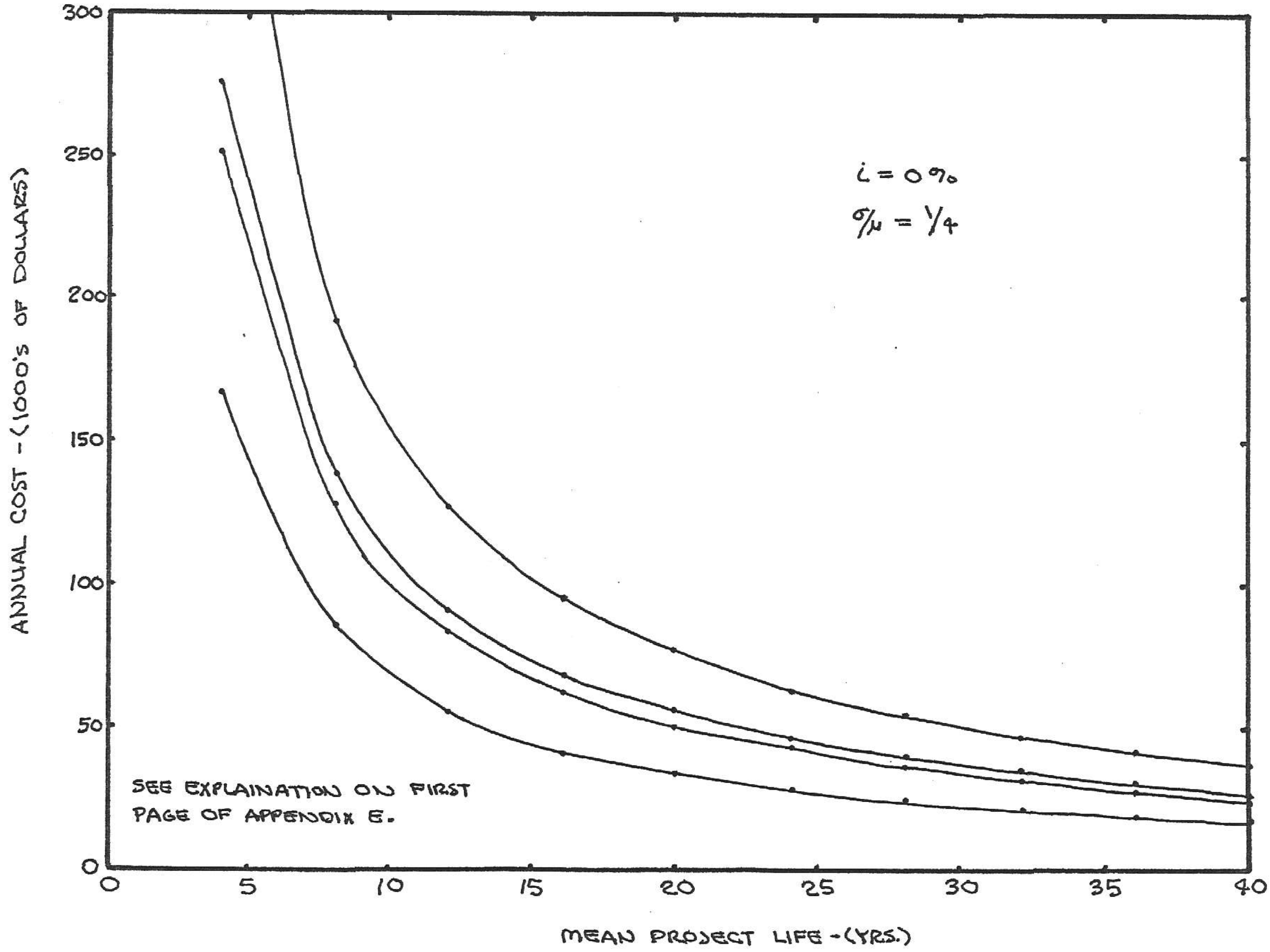
$$SD = 17{,}807$$

$$\frac{SD}{E(R)} = \frac{17{,}807}{63{,}178} = 28.2\,\%$$

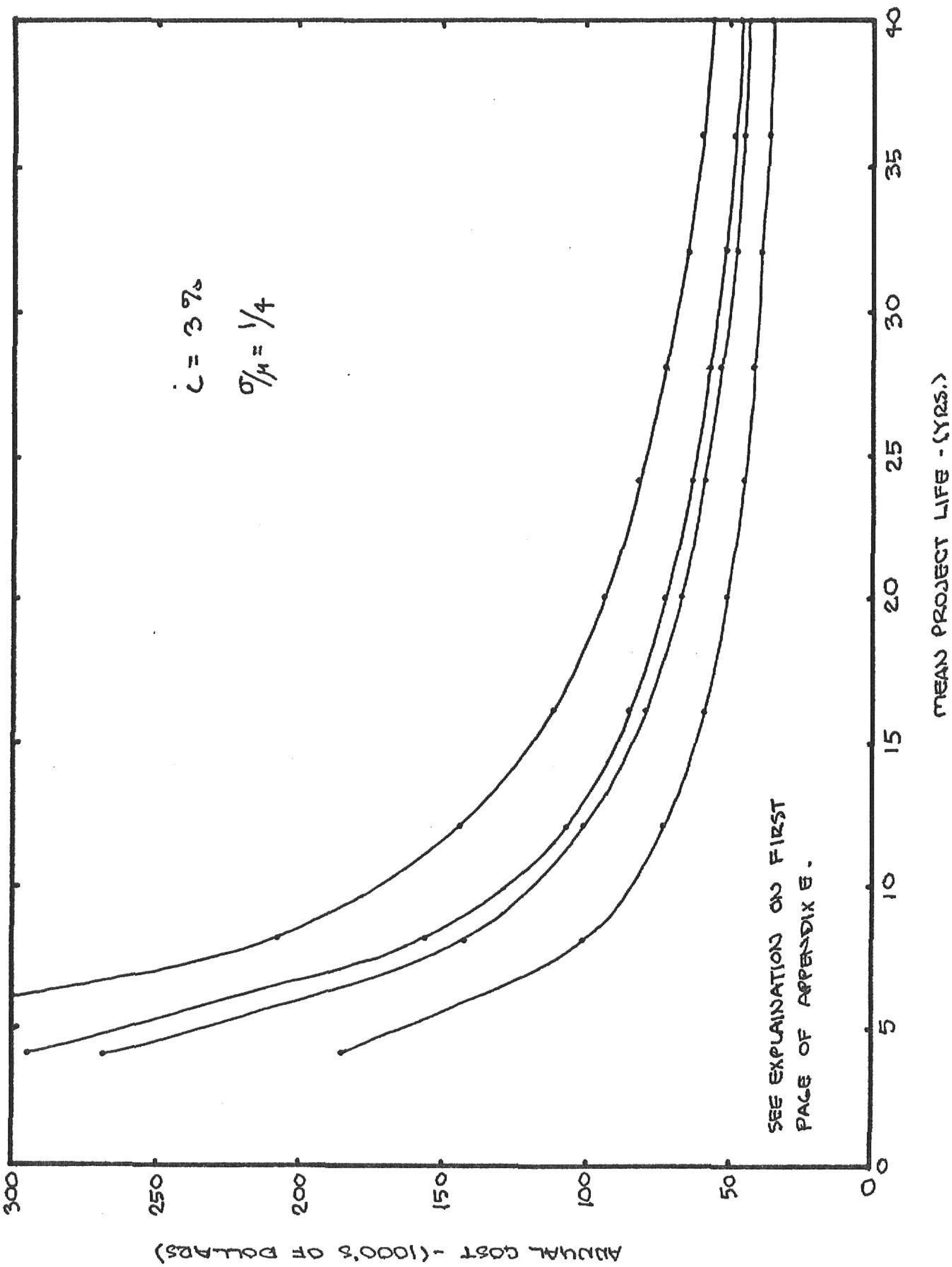
$$\frac{E(R)-R}{E(R)} = \frac{63{,}178 - 59{,}050}{63{,}178} = 6.5\,\%$$

APPENDIX E - Annual Cost vs. Mean Life for Various i and σ/μ

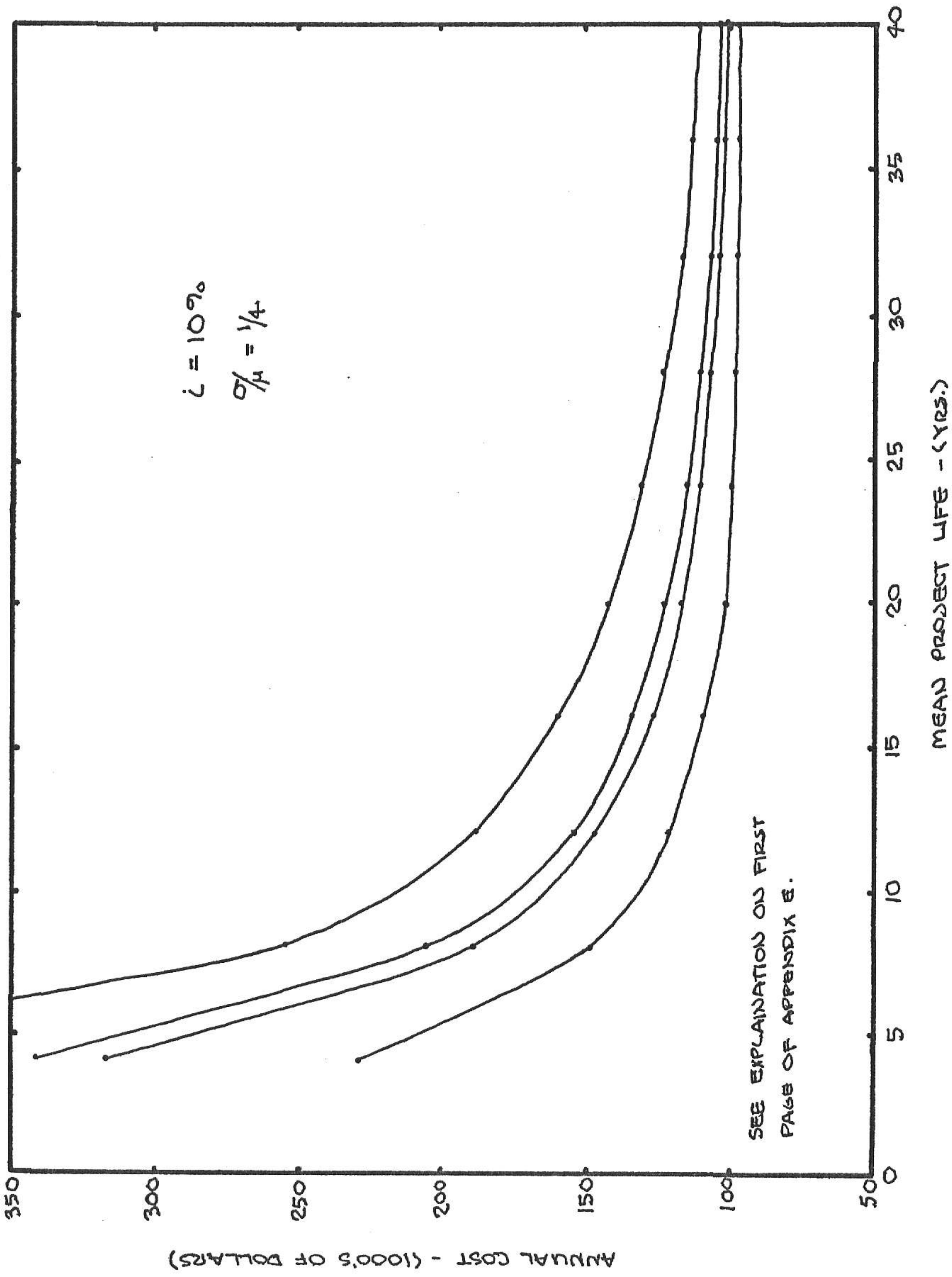
These curves show the relationship between mean life and annual cost. The lower of the two middle lines represents the traditional annual cost. The upper of the middle two lines represents expected annual cost. The extreme top and bottom lines represent the plus or minus one standard deviation interval about the expected annual cost line.

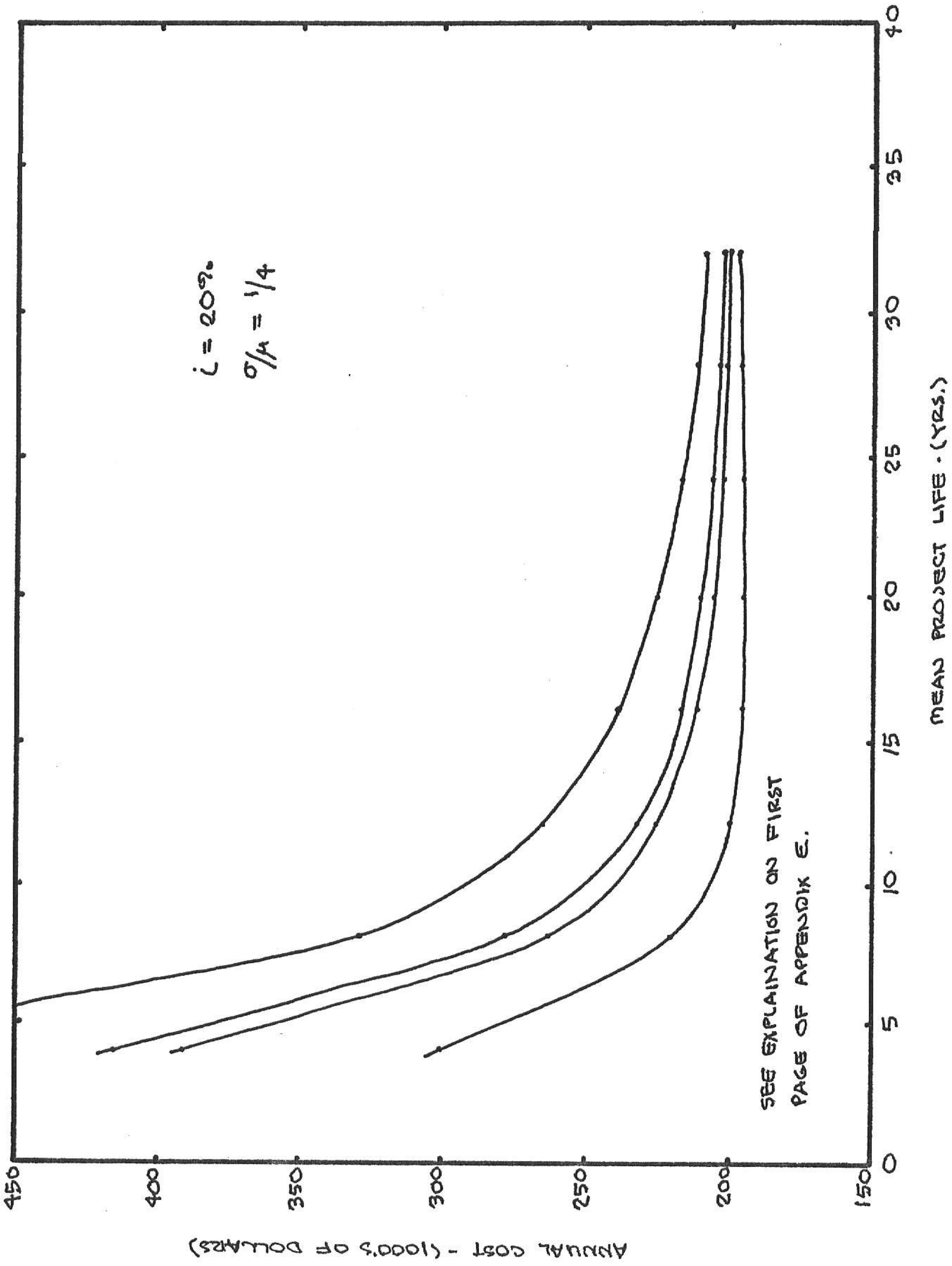


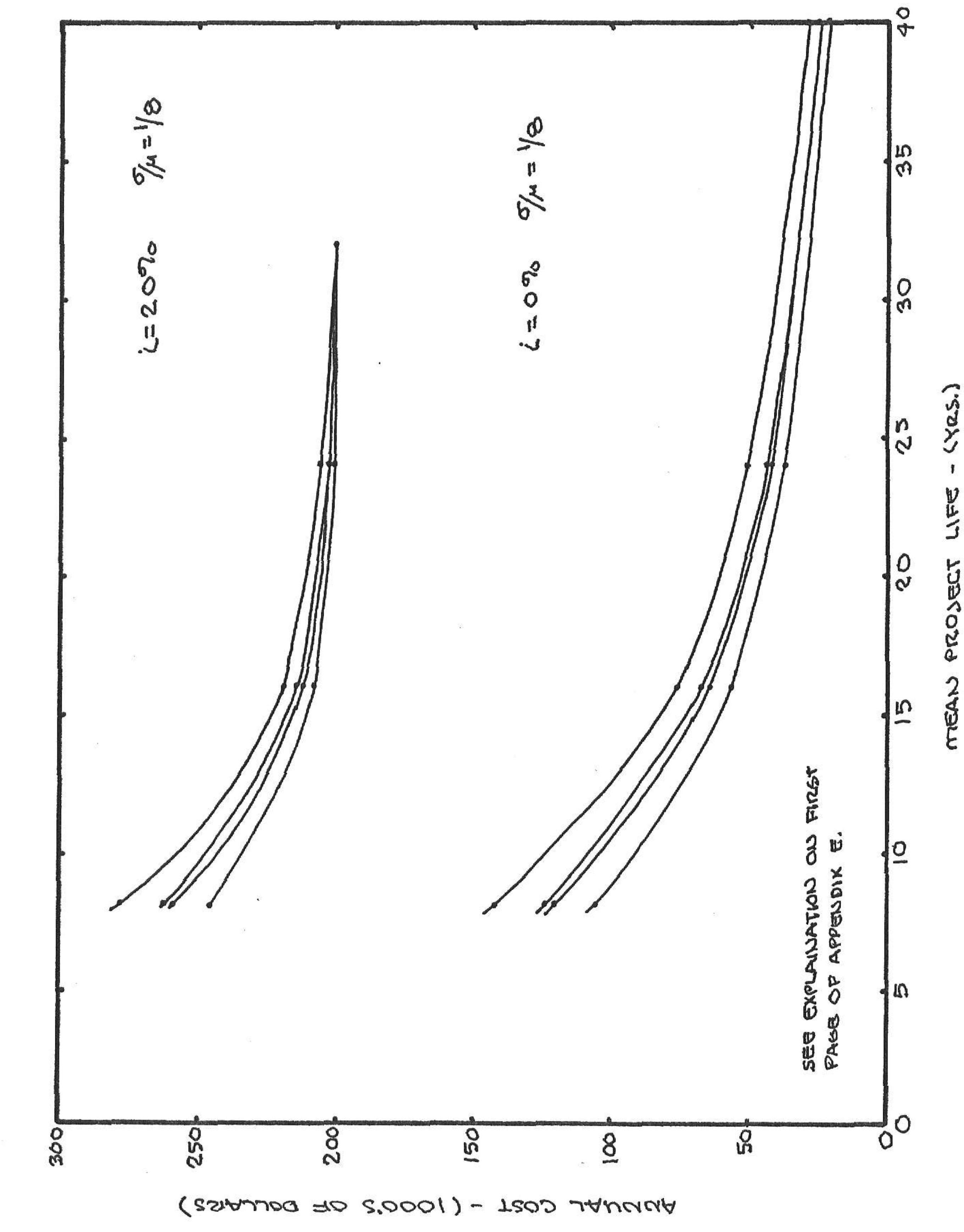
300
250
200
150
100
50
0
0
5
10
15
20
25
30
35
40
ANNUAL COST -(1000's OF DOLLARS)
MEAN PROJECT LIFE -(YRS.)
$i = 0\%$
$\sigma/\mu = 1/4$
SEE EXPLAINATION ON FIRST PAGE OF APPENDIX E.



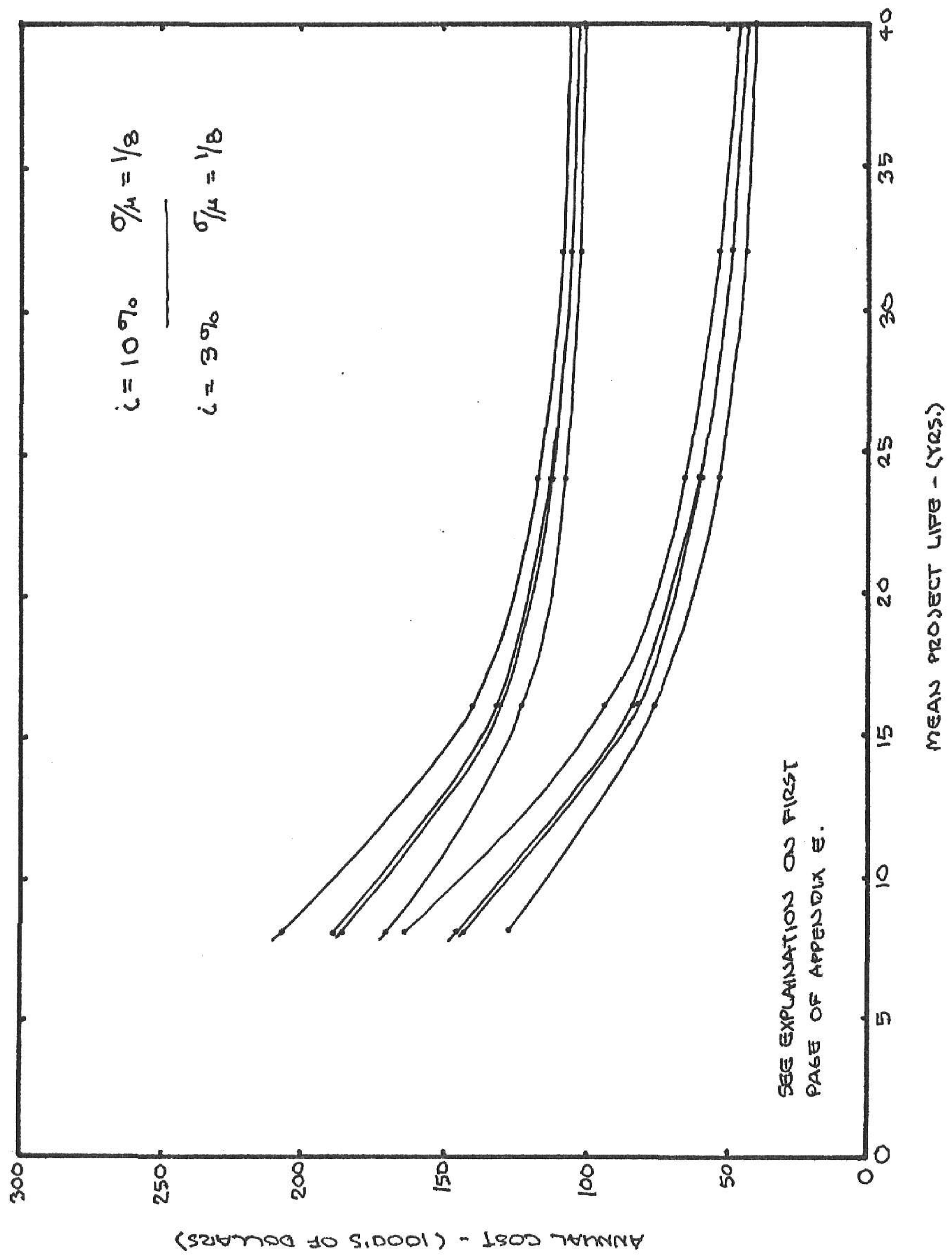
$i = 3\%$
$\sigma/\mu = 1/4$
SEE EXPLAINATION ON FIRST PAGE OF APPENDIX E.
MEAN PROJECT LIFE - (YRS.)
0
5
10
15
20
25
30
35
40
ANNUAL COST - (1000's OF DOLLARS)
0
50
100
150
200
250
300


i = 10%
σ/μ = 1/4
SEE EXPLANATION ON FIRST
PAGE OF APPENDIX E.
MEAN PROJECT LIFE - (YRS.)
ANNUAL COST - (1000'S OF DOLLARS)
0
5
10
15
20
25
30
35
40
50
100
150
200
250
300
350


i = 20%
σ/μ = 1/4
SEE EXPLAINATION ON FIRST PAGE OF APPENDIX E.
0
5
10
15
20
25
30
35
40
MEAN PROJECT LIFE - (YRS.)
150
200
250
300
350
400
450
ANNUAL COST - (1000'S OF DOLLARS)



ANNUAL COST - (1000'S OF DOLLARS)
MEAN PROJECT LIFE - (YRS.)
0
50
100
150
200
250
300
5
10
15
20
25
30
35
40
i = 20%
σ/μ = 1/8
i = 0%
σ/μ = 1/8
SEE EXPLANATION ON FIRST
PAGE OF APPENDIX E.


$i = 10\%$ $\sigma/\mu = 1/8$
$i = 3\%$ $\sigma/\mu = 1/8$
SEE EXPLAINATION ON FIRST
PAGE OF APPENDIX E.
MEAN PROJECT LIFE - (YRS.)
ANNUAL COST - (1000's OF DOLLARS)
0
5
10
15
20
25
30
35
40
0
50
100
150
200
250
300

CONSIDERING THE UNCERTAINTY OF ESTIMATES
IN ENGINEERING ECONOMIC ANALYSIS

by

JON HOWARD ESHELMAN

B.S., Kansas State University, 1970

AN ABSTRACT OF A MASTER'S REPORT

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Civil Engineering

KANSAS STATE UNIVERSITY

Manhattan, Kansas

1971

The traditional engineering economic analysis requires that exact, single-valued estimates be made for such variables as project life, initial cost, interest rate, salvage value, and so forth. Knowledge of these factors is certainly not so complete that we can make these estimates with any degree of confidence, although certainly such things as initial cost can be more accurately estimated than can salvage value. It seems more reasonable that probability distributions could more realistically describe the possible values which could be assumed by a variable than could a single-valued estimate.

A theoretical analysis is developed which allows the engineer to quantatatively consider uncertainty in any number of estimated variables, assuming that the initial probability distributions which describe those variables are known.

An approximation technique is reviewed which employes a random number generator which, over the course of many iterations, approximates the set of all possible outcomes described theoretically by the initial probability distributions, which must again be known.

An example analysis is discussed, results are presented, and general conclusions drawn. The example case is one in which project life is the only variable having uncertainty in its estimate. A normal probability distribution was used to describe the variability in project life.

The conclusion of the example analysis was that although the expected value of annual cost was always slightly higher than the annual cost computed by the traditional method, the difference was insignificant except in very special cases of very low interest rates and/or short mean lives.