X. Extreme Event

PROBLEM X.1: Analysis of Annual Maximum River Discharge

This problem deals with the importance of considering an extreme event such as a flood when making decisions about dam construction.

DESCRIPTION

From US Geological Survey Water Resources Data, the annual maximum discharges for the gauging station of the Salt Fork River near St. Joseph, Illinois between 1959 and 1975 are listed below.

Table X.1.1. Annual Maximum Discharge Data

Year	Annual Discharge feet^3/sec (CFS)
1959	6030
1960	1800
1961	2300
1962	3370
1963	2340
1964	5380
1965	1230
1966	950
1967	2640
1968	6860
1969	2630
1970	2600
1971	2350
1972	1350
1973	2990
1974	2750
1975	1920

METHODOLOGY

We use Extreme Event (EE) Analysis to help make the most effective decision.

a) Suppose a Type I Largest value extremal distribution is adopted for the annual maximum discharge at this station. Determine the parameters for the distribution through sample means and variance.

b) The Probable Maximum Flood (PMF) is defined as the level of cubic feet per second (CFS) whose exceedence probability is 10^{-5}. Find the PMF for the above gauging station by filling the missing entries in Columns 2 and 3 of Table X.1.2 where y is the "actual" flood in any year, measured in terms of proportion (or percentage) of PMF.

c) Suppose there is a flood control earth dam of 100 ft. in height at this gauging station. However, the dam was inappropriately designed as considerable damage downstream still occurs even for moderate floods. Column 4 of Table X.1.2 shows potential flood damage, X in 10^6\$, downstream for different levels of flood y. To make the dam more effective, its height, h feet, is to be raised. The damage X is therefore a function of Y and h. For a given level of dam height h, $h \geq 100$, the damage X is a monotone increasing function of Y. For this reason, the probability that an annual damage X exceeds a particular level x (x is the damage caused by a level y flood) for the existing dam is exactly equal to the probability that the flood level y is exceeded. Use this information to i) fill in all entries in Column 4 of Table X.1.2, and ii) to compute the conditional expected damages f_2, f_3, f_4, and expected damage f_5 for the existing dam. (f_2, f_3, and f_4 are expected damages in 10^6\$, conditioned on the exceedence probability of damage being in [0, .01], [.01, .88], and [.88, 1] respectively.) Fill in this information in Table X.1.3.

d) If the cost of raising the height of the dam in 10^6\$ is $f(h) = h - 100$ for $h \geq 100$, fill in the first row of Table X.1.3 in which three other dam heights are being considered. Suppose the values of f_2, f_3, f_4, and f_5 for the three new alternatives have been computed as shown in Table X.1.3. Plot f_1 vs. f_2, f_1 vs. f_3, f_1 vs. f_4, and f_1 vs. f_5 on the same graph using f_1 as the vertical axis. Estimate all relevant trade-offs between various pairs of alternatives and summarize your results in the form of a table. Discuss their implications on the dam height decision.

Table X.1.2. Template for Flood Data

Y Proportion of PMF	Flood in CFS	Exceedence Prob. $P(Y \geq y)$	Potential Damage in 10^6\$ X, when h = 100 ft.	Exceedence Prob. $P(X \geq x)$ when h = 100
.01		.9906	.1	
.05		.9358	.5	
.07		.8821	1.2	
.10			2.0	
.20		.3155	3.0	
.30		.0972	4.5	
.40		.0270	7.0	
.50		.0073	10.0	
.60		.0020	14.0	
.80		.000167	25.0	
1.0		.00001	50.0	

Table X.1.3. Cost of Dam Improvement Options

Cost of Expected Damage in 10^6\$	Option 0 h_0=100 ft.	Option 1 h_1=102 ft.	Option 2 h_2=105 ft.	Option 3 h_3=110 ft.
f_1: cost				
f_4: low prob. Damage		8.5	5.5	3.0
f_3: med. Prob. Damage		2.4	1.8	1.5
f_2: high. prob. damage		.35	.10	0.0
f_5: overall expected damage		2.0	1.5	1.0

SOLUTION

a) Sample mean $\bar{x} = 2911$

Sample variance $SD^2 = (1663)^2$ (using SD_{n-1} rather than SD_n), hence, SD = 1663

$$\therefore \quad \hat{\alpha}_y = \frac{\pi}{\sqrt{6}(SD)} = 7.7 \text{ x } 10^{-4}$$

and

$$\hat{u}_y = \bar{x} - \frac{\gamma}{\hat{\alpha}_y} = 2911 - \frac{0.577216}{7.7 \text{ x } 10 - 4} = 2161.37$$

b) $0.00001 = P(Y \geq PMF) = P(S \geq \hat{\alpha}_y (PMF - \hat{u}_y)) = P(S \geq 11.50)$

Thus $\hat{\alpha}_y (PMF - \hat{u}_y) = 11.50 \Rightarrow PMF = \dfrac{11.50}{\hat{\alpha}_y} + 2161.37 = 17096.44$

$\therefore$ 0.1PMF = 1709.64

$P(Y \geq 0.1PMF) = P(Y \geq 1709.64) = P(S \geq 7.7 \text{ x } 10^{-4}(1709.64 - 2161.37))$

$= P(S \geq -0.35) = 1 - P(S \leq 0.35) = 1 - 0.2419 = .7581$

c) With the information given, Table X.1.2 becomes:

Table X.1.4. Flood Data with Completed Exceedance Probabilities

Y Proportion of PMF	Flood in CFS	Exceedence Prob. $P(Y \geq y)$	Potential Damage in 10^6\$ X, when h = 100 ft.	Exceedance Prob. $\hat{F}(x) = 1 - F(x)$ $P(X \geq x)$ when h=100
.01	170.96	.9906	.1	.9906
.05	854.8	.9358	.5	.9358
.07	1196.72	.8821	1.2	.8821
.10	1709.64	.7581	2.0	.7581
.20	3419.20	.3155	3.0	.3155
.30	5128.8	.0972	4.5	.0972
.40	6838.4	.0270	7.0	.0270
.50	8548.0	.0073	10.0	.0073
.60	10257.6	.0020	14.0	.0020
.80	13676.8	.000167	25.0	.000167
1.0	17096	.00001	50.0	.00001

Note: X = X(Y, h) X is a monotone increasing function of Y, given h.
∴ $P(X \geq x: h=100) = P(Y \geq y)$ where $x = X(y: h=100)$

For simplicity, let the partitions in the probability scale IP1, IP2, and IP3, be transformed into partitions in the damage scales IX1, IX2, and IX3 respectively, as shown in Table X.1.2. Thus:

$$f_2 = E(X|IX_1) = \frac{\sum_{x_{i+1}, x_i \in IX_1} |\hat{F}(x_i) - \hat{F}(x_{i+1})| \frac{(x_{i+1} + x_i)}{2}}{IP_1}$$

$$= \frac{|1.0 - 0.9906| \frac{(0+0.1)}{2} + |0.9906 - 0.9358| \frac{(0.1+0.5)}{2} + |0.9358 - 0.8821| \frac{(0.5+1.2)}{2}}{1 - 0.8821}$$

$$= 0.53$$

Likewise,

$$f_3 = E(X|IX_2)$$

$$= \frac{|0.8821 - 0.7581| \frac{(1.2+2.0)}{2} + |0.7581 - 0.3155| \frac{(2.0+3.0)}{2}}{0.8821 - 0.0073} +$$

$$\frac{|0.3155-0.0922|\frac{(3.0+4.5)}{2}+|.0922-.0270|\frac{(4.5+7)}{2}+|.027-.0073|\frac{(7+10)}{2}}{.8821-.0073}$$

$$= 3.07$$

$$f_4 = \mathrm{E}\,(X|X_3)$$

$$= \frac{|.0073-.002|\frac{(10+14)}{2}+|.002-.000167|\frac{(14+25)}{2}+|.000167-.00001|\frac{(25+50)}{2}}{.0073-0} +$$

$$\frac{|.00001-0|\frac{(50+100)}{2}}{.0073-0}$$

$$= 18.22$$

$$f_5 = f_4(.0073) + f_3(.8821\text{-}.0073) + f_2(1\text{-}.8821)$$

$$= 2.88$$

d) Table X.1.3 becomes:

Table X.1.5. Cost of Dam Improvement Options with Completed Values of Dam Height Options at Various Partitions

Cost of Expected Damage in 10^6\$	Option 0 h_0=100ft	Option 1 h_1=102ft	Option 2 h_2=105ft	Option 3 h_3=110ft
f_1: cost	0	2	5	10
f_4: low prob. damage (catastrophe)	18.22	8.5	5.5	3.0
f_3: med. prob. damage	3.07	2.4	1.8	1.5
f_2: high prob. damage	.53	.35	.10	0.0
f_5: overall expected damage	2.88	2.0	1.5	1.0

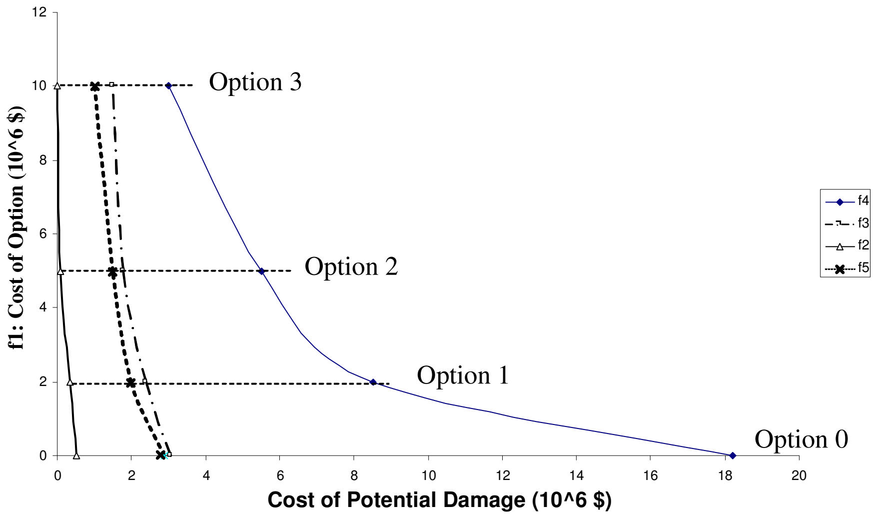

Figure X.1.1. Plot of Table X.1.5

Trade-offs: $\lambda_{ij}(h_p, h_q) \approx \dfrac{f_i(h_p) - f_i(h_q)}{f_j(h_p) - f_j(h_q)}$

	$h_0 - h_1$	$h_0 - h_2$	$h_0 - h_3$	$h_1 - h_2$	$h_1 - h_3$	$h_2 - h_3$
λ_{12}	-11.1	-11.63	-18.87	-12.0	-22.86	-50
λ_{13}	-2.99	-3.94	-6.37	-5.0	-8.89	-16.7
λ_{14}	-0.21	-0.36	-0.66	-1.0	-1.45	-2.0
λ_{15}	-2.27	-3.62	-5.32	-6.0	-8.0	-10

ANALYSIS

The need to consider extreme events as opposed to "average" events is demonstrated clearly in this example. For example, in comparing Option h_1 with h_0, it costs \$2M to construct h_1 while the reduction in damage is \$0.88M in the "average" sense (f_5), and \$9.72M in the extreme-event case (f_4). Considering the average case (f_5) alone may lead to rejection of h_1 in favor of h_0 since the "average" expected benefit will not pay for the "certain" cost. However, considering the *extreme event* case (f_4) will make h_1 very attractive, as the expected benefit in case of an extreme event is about five times that of the cost.

PROBLEM X.2: Integrated Circuit for a Helicopter

Four design options are being considered for an integrated circuit chip computer subsystem for a combat helicopter.

DESCRIPTION

The objective of this problem is to maximize the reliability of the chip subsystem for a mission of a three-hour duration. The reliability of each design must be weighed against the cost.

METHODOLOGY

We solve this problem using Extreme Event Analysis.

Suppose N identical components are tested for one time period. Let $N_f(t)$ be the number of components that have failed, and $N_0(t)$ be the number of components that are operating. The failure rate (λ) of the components is given by:

$$\lambda = \frac{N_f(t)}{N} = \frac{N_f(t)}{N_o(t) + N_f(t)}$$

The reliability $R(t)$ of a system is defined as the conditional probability that a system performs correctly throughout an interval of time $[t_0, t]$, given that the system was performing correctly at time t_0. For components having exponential time to failure distribution, the reliability is given by:

$$R(t) = e^{-\lambda t}$$

You may want to use the following equation:

$$f_4^N(\cdot) = \mu + \sigma\sqrt{2\ln(n)}$$

An integrated circuit chip for use in the computer for a combat helicopter has a mean failure rate of 0.05 (λ = 0.05) per hour, and a standard deviation of 0.02 (assuming normal distribution). The cost of each chip is $100. Maximizing the reliability of the chip subsystem for a three-hour mission can be done by placing the chips in parallel.

The failure rate for such a parallel system is given by λ^n, where n is the total number of parallel components used. Assume that the standard deviation of the parallel system is the same for each individual chip.

SOLUTION

i) The 4 design options for the chip subsystem each have 1, 2, 3, and 4 chips in parallel. For each of the options, compute the mean reliability of the subsystem for the 3-hour mission.

Design Option 1:

$\lambda_{sys} = (0.05)^{1} = 0.05$

$R(3) = e^{(-0.05)3} = 0.86071$

Design Option 2:

$\lambda_{sys} = (0.05)^{2} = 0.0025$

$R(3) = e^{(-0.0025)3} = 0.99253$

Taking similar steps for Design Options 3 and 4, the complete results are as follows:

Table X.2.1. Failure and Reliability Rate by Design Option

Design Option	λ_{sys}	R(3)
1	0.05	0.86071
2	0.0025	0.99253
3	0.000125	0.99963
4	0.00000625	0.99998

ii) Calculate $f_4(\cdot)$ for each design option for a partition point of 85% on the reliability axis.
Hint: To compute the partition point on the probability axis from the partition point on the reliability axis, calculate the corresponding value of the failure rate using:

$$R(t) = e^{-\lambda t}$$

Design Option 1:

$$R(3) = e^{-3\lambda_1} = 0.85$$

where λ_1 is the partition point on the failure rate axis.

$$\lambda_1 = 0.05417$$

Converting to standard normal:

$$\Phi^{-1}(\alpha_1) = \frac{\lambda_1 - \mu_\lambda}{\sigma} = 0.2085 \approx 0.21$$

where μ_λ is the mean failure rate of the system, σ is the standard deviation of the failure rate, and α_1 is the partition point of the probability axis.

From standard normal tables we have:

$$\alpha_1 = 0.5826$$

$$\begin{aligned} f_4^n(\cdot) &= \mu + \sigma\sqrt{2\ln(n)} \\ &= 0.05 + 0.02\sqrt{2\ln\left(\frac{1}{1-0.5826}\right)} \\ &= 0.07644 \end{aligned}$$

Taking similar steps for Design Options 2, 3, and 4, the complete results are shown in Table X.2.2:

Table X.2.2. f_4 with a Partition Point of 85% on the Reliability Axis

Design Option i	λ_i	$\Phi^{-1}(\alpha_i)$	α_i	n	$f_4^N(\cdot)$
1	0.05417	0.21	0.5826	2.3960	0.07644
2	0.05417	2.58	0.9951	204.5787	0.06774
3	0.05417	2.70	0.9966	290.5198	0.06748
4	0.05417	2.71	0.9966	295.7597	0.06747

iii) Calculate $f_4(\cdot)$ for each design option using a partition point of 0.99 on the probability axis.

<u>Design Option 1:</u>

$$\begin{aligned} f_4^n(\cdot) &= \mu + \sigma\sqrt{2\ln(n)} \\ &= 0.05 + 0.02\sqrt{2\ln\left(\frac{1}{1-0.99}\right)} \\ &= 0.11070 \end{aligned}$$

Taking similar steps for Design Options 2, 3, and 4, the complete results are presented in Table X.2.3:

Table X.2.3. f_4 with a Partition Point of 99% on the Probability Axis

Design Option	α	n	$f_4^N(\cdot)$
1	0.99	100	0.11070
2	0.99	100	0.06320
3	0.99	100	0.06082
4	0.99	100	0.06070

iv) Plot your results in terms of cost vs. damage, where the damage is given by the calculated value of the failure rate of the chip subsystem. Below is the graph for these results:

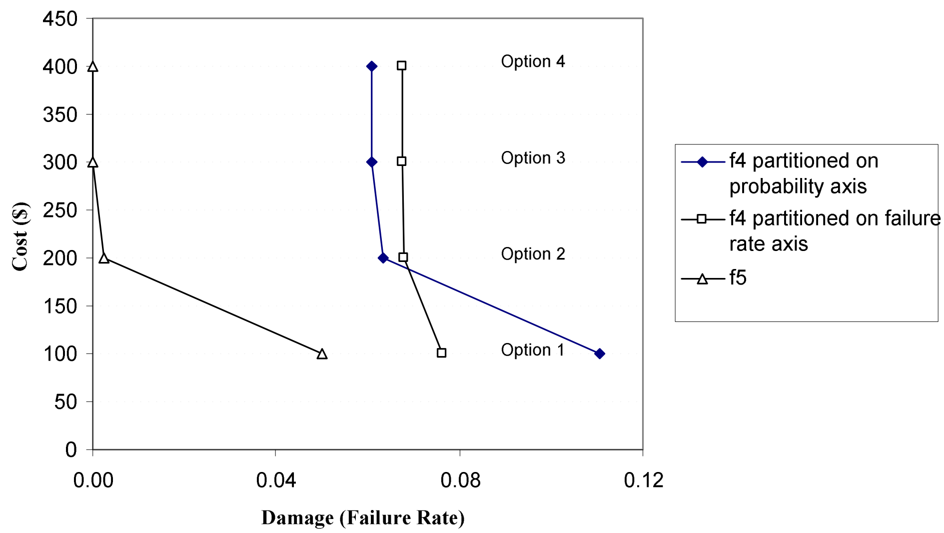

Figure X.2.1. Unconditional and Conditional Damage Rates versus Cost

ANALYSIS

Figure X.2.1 shows the performance of the four design options relative to two objectives: minimizing failure rate and minimizing cost. Several Pareto-optimal solutions are presented based on expected value (f_5), and two versions of conditional expected value (f_4) based on partitioning either the failure-rate axis or the probability axis. It should be observed that Option 1 has a low cost but the failure rate is high. Option 2 significantly reduces the failure rate at an additional unit cost of $100. It should be observed that the reduction in failure rate between Options 3 and 4 is marginal. Nevertheless, the small change in reliability could be crucial, especially when dealing with safety-critical systems such as the computer used in operating a combat helicopter.

PROBLEM X.3: Overpopulation

Overpopulation is becoming a threat to many developing countries. This problem addresses this through the perspective of conception.

DESCRIPTION

A developing country is considering trying to decrease the number of new babies, so that it can control the growth of the overall population. It must decide which of the following four birth control options to subsidize in order to control the number of conceptions:

- Contraceptive pills
- Contraceptive patches
- Condoms
- Diaphragms

METHODOLOGY

We try to find the best solution using the extreme event analysis

The number of theoretical conceptions is assumed to be of a normal distribution with a mean μ of 0 and standard deviation σ of 1.

Key Assumptions:

- The number of theoretical conceptions is statistically independent between days.
- We conduct the birth control measurement experiment for an initial period of 1 month (30 days).
- We take a sample of 100 couples in the given geographical area.

The following table summarizes the corresponding cost, mean, and standard deviation for each of the birth control strategies:

Table X.3.1. Birth Control Strategies with Associated Cost, Mean, and Standard Deviation

Birth Control Strategies	Cost ($)	Mean	Std. Deviation
1. Contraceptive Patch	2000	0	1
2. Contraceptive Pills	1500	3	4
3. Condoms	100	10	9
4. Diaphragms	2400	18	12

SOLUTION

The four required steps are as follows:

A) Determine the most probable one-month maximum number of conceptions.

We calculate as follows:

$\mu = 0$
$\sigma = 1$

$u_n = \mu + \sigma \Phi^{-1}\,(1\text{-}1/n)$

Model value for n = 30:
$u_{30} = 0 + \Phi^{-1}\,(1\text{-}1/30)$
$= 1.834$

Follow these same steps for all four strategies.

B) Determine the probability that the maximum number of conceptions will exceed 20 in the given month and determine the corresponding return period.

<u>Strategy 1—Contraceptive patch:</u>

$P_Y(y) = [P_X(y)]^{30}$
$= [\Phi((y\text{-}\mu)/\sigma)]^{30}$
Pr (max # of conceptions > 20) = $1\text{-}\ P_Y\,(20)$
$= 1\text{-}\ [\Phi\ ((20\text{-}0)/1)]^{30}$

Return period of maximum number of conceptions of 20 = (1/0), which implies it would seldom exceed 20.

<u>Strategy 2—Pill:</u>

$P_Y(\mathrm{y}) = [P_X(y)]^{30}$
$= [\Phi((\mathrm{y}\text{-}\mu)/\sigma)]^{30}$
Pr (max # of conceptions > 20) = $1\text{-}\ \mathrm{P_Y}\,(20)$
$= 1\text{-}\ [\Phi((20\text{-}3)/4)]^{30}$

Return period of maximum number of conceptions of 20 = 3117 months.

<u>Strategy 3—Condoms:</u>

$P_Y(y) = [P_X(y)]^{30}$
$= [\Phi\ ((y\text{-}\mu)/\sigma]^{30}$
Pr (max # of conceptions > 20) = $1\text{-}\ P_Y\,(20)$
$= 1\text{-}\ [\Phi\ ((20\text{-}10)/9)]^{30}$

Return period of maximum number of conceptions of 20 = 1.0139 months.

<u>Strategy 4—Diaphragms:</u>

$P_Y(y) = [P_X(y)]^{30}$
$= [\Phi\,((y\text{-}\mu)/\sigma)]^{30}$
Pr (max # of conceptions > 20) = 1- $P_Y\,(20)$
$= 1\text{-}\,[\Phi((20\text{-}18)/12)]^{30}$

Return period of maximum number of conceptions of 20 = every month.

C) Determine the expected (f_5) and conditional expected values (f_4) for the above return period for each birth control strategy.

Strategy 1—Patch

$$f_4(30) = u_n + (1/\delta_n)$$

$$\delta_{30} = \frac{30}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\Phi^{-1}(1-\frac{1}{30})\right)^2\right]$$
$$= 2.227$$

Therefore, $f_4(30) = 1.834 + (1/2.227) = 2.283$
$f_5(30) = 0$

Strategy 2— Pills

$$f_4(30) = u_n + (1/\delta_n)$$

$$\delta_{30} = \frac{30}{\sqrt{2\pi 4^2}}\exp\left[-\frac{1}{2}\left(\Phi^{-1}(1-\frac{1}{30})\right)^2\right]$$
$$= 0.5567$$

Therefore, $f_4(30) = 10.336+(1/0.5567) = 12.13$
$f_5(30) = 3$

Strategy 3—Condoms

$$f_4(30) = u_n + (1/\delta_n)$$

$$\delta_{30} = \frac{30}{\sqrt{2\pi 9^2}}\exp\left[-\frac{1}{2}\left(\Phi^{-1}(1-\frac{1}{30})\right)^2\right]$$
$$= 0.2474$$

Therefore, $f_4(30) = 26.505+(1/0.2474) = 30.55$
$f_5(30) = 10$

Strategy 4—Diaphragm

$f_4(30) = u_n + (1/\delta_n)$

$$\delta_{30} = \frac{30}{\sqrt{2\pi 12^2}} \exp\left[-\frac{1}{2}\left(\Phi^{-1}(1-\frac{1}{30}) \right)^2 \right]$$

$= 0.1856$

Therefore, $f_4(30) = 40.007+(1/0.1856) = 45.40$

$f_5(30) = 18$

D) Plot the cost of each strategy versus both the expected value (f_5) and the conditional expected value (f_4) of the number of conceptions.

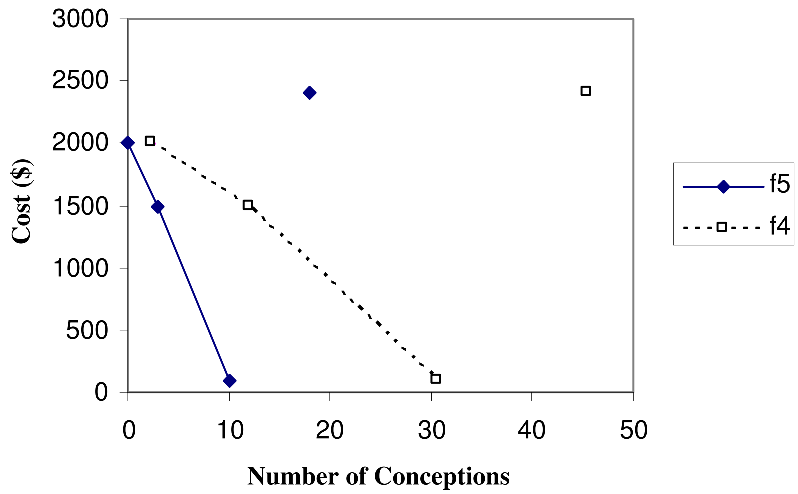

Figure X.3.1. Cost versus conditional expected value (f_4) and expected value (f_5) for 30 days

ANALYSIS

Obviously, from the above chart we may conclude that Strategy 4 (diaphragms) is an inferior solution. It is an older technology that is not only expensive to implement, but more invasive and less effective. However, it doesn't manipulate the hormones in the way Strategies 1 (the patch) and 2 (the pill) do.

Strategy 1 has the greatest likelihood of working with the lowest risk, but comes at a price. To control overpopulation, it is likely worth the extra $500/month/100 people to choose Strategy 1 over Strategy 2, or choose the birth control patch over the pill. The reason for the great change in extreme value is the complication and consistency required in using the pill (Strategy 2), that is not required in applying the patch (Strategy 1) for the same effectiveness.

Strategy 3, issuing condoms for population control, appears to be a less effective method. However because of its nature, the number of theoretical conceptions can be decreased simply by purchasing a greater number. A future study may compare the greater distribution of condoms (Strategy 3) using the patch (Strategy 1) as a benchmark.

By understanding the most effective tools to aid couples in family planning, the government can have a means of knowing where to subsidize these efforts. With government subsidies the population growth can become better controlled, which will lead to greater efficiency in existing education, transportation, and related infrastructures. Ultimately, population control will lead to a developing country's economic stability and growth and enhance the population's standard of living.

PROBLEM X.4: Effective Snow Removal in a City

The goal of this problem is to use extreme event analysis derived from historical snow precipitation data to analyze different policies for efficient removal of roadway snow accumulation.

The Department of Transportation in a particular city is concerned with keeping a city's roadways free of snow and ice during the winter months, especially in January. They have 4 different policy options on snow removal, and wish to improve efficiency.

Analyzing the city's historical snowfall data revealed that in January, the amount of daily snowfall (x in cm) has the following distribution:

$$\Pr(X \le x) = \begin{cases} 0.92 & \text{if } x = 0 \\ 0.92 + 0.08 \cdot \int_0^x \frac{1}{us\sqrt{2\pi}} \exp\left(-\frac{(\ln u - k)^2}{2s^2}\right) du & \text{if } x > 0 \end{cases}$$

where $s = 1$, $k = 2.7$ (obtained through simulation).

The analysis showed that each of four options has the same effect of changing s and k values of the above distribution, such that:

- Policy 1 ~ ($s = 1.2$, $k = 2.0$)
- Policy 2 ~ ($s = 1.2$, $k = 1.5$)
- Policy 3 ~ ($s = 1$, $k = 2.0$)
- Policy 4 ~ ($s = 1$, $k = 1.5$)

Using the above specifications:

(i) Calculate u_{30}, the most probable 1-month maximum snowfall (cm) for each of the above policy options.

(ii) Calculate the probability that maximum snowfall of 20cm would be exceeded. What is the corresponding return period of the exceeding 20cm for each of the policy options?

(iii) Using the approximation formula $f_4(\cdot) = u_t + 1/\delta_t$, where $\delta_t = tf_X(u_t)$, perform multiobjective tradeoff analysis of $f_5(\cdot)$ vs. monthly cost, and $f_4(\cdot)$ vs. monthly cost. Assume that the monthly cost for Policies 1, 2, 3, and 4 are \$2M, \$4M, \$5M, and \$8M, respectively.

PROBLEM X.5: Analyzing Investment Opportunities

An investor wants to decide between four investment opportunities. A market theory asserts that investment returns, denoted by X, are normally distributed. For this problem, we interpret investment returns X as "opportunity losses." Therefore, the upper-tail region in a distribution of such investment returns corresponds to events that have low likelihoods of occurrence, but high opportunity losses.

An investor who has faith in this market theory wants to conduct extreme-event analysis for the following four long-term bond investment alternatives. For a given investment i, the notation $X_i \sim N_i(\mu_i, \sigma_i)$ is used to refer to a normal distribution with parameters μ_i and σ_i, the mean and standard deviation, respectively, of the underlying random variable X_i. These parameters are estimated from historical *annual* data.

(i) *Investment 1*: $X_1 \sim N_1(0.047, 0.010)$; Unit Cost = \$10
(ii) *Investment 2*: $X_2 \sim N_2(0.048, 0.015)$; Unit Cost = \$8
(iii) *Investment 3*: $X_3 \sim N_3(0.049, 0.020)$; Unit Cost = \$5
(iv) *Investment 4*: $X_4 \sim N_4(0.050, 0.025)$; Unit Cost = \$4

Procedure:

(a) Calculate the f_4 for each investment alternative for $n = 20$ years. Use the exact formulas for f_4 of a normal distribution.

(b) For each of the investment alternatives, plot the f_4 and f_5 values along the x-axis and the corresponding costs along the y-axis. Analyze the resulting graph.

(c) Rework (a) using the approximation formulas for u_n and δ_n and f_4 as follows, and calculate the % error relative to the exact values obtained from (a):

$$f_4 = u_n + 1/\delta_n \tag{X.5.1}$$

$$u_n = \mu + \sigma\left[\sqrt{2\ln n} - \frac{\ln(4\pi \ln n)}{2\sqrt{2\ln n}}\right] \tag{X.5.2}$$

$$\delta_n = \frac{\sqrt{2\ln n}}{\sigma} \tag{X.5.3}$$

PROBLEM X.6: Modeling Stream's Oxygen Concentration

To determine the concentration of dissolved oxygen in a stream. The daily level of dissolved oxygen (DO) concentration for a stream is assumed to be of normal distribution with a mean of 3 mg/L and a standard deviation of 0.5 mg/L. We assume that the DO concentrations between days are statistically independent.

Using extreme event analysis, determine:

a) Determine the one-year most probable maximum DO level.
b) Determine the probability that the maximum DO level will exceed 5 mg/L in a year. Determine the corresponding return period.