

**Project 2**

**Benefit-Cost Analysis**

Sanchi Gupta

College of Professional Studies, Northeastern University Boston

ALY6050: Introduction to Enterprise Analytics

Professor Roy Wada

March 5, 2024

**Introduction**

In this report, we explore the use of Monte Carlo simulations to forecast benefit-cost ratios for two proposed dam projects, highlighting the significance of stochastic methods in strategic decision-making. These simulations enable us to model and predict outcomes where direct calculation is impractical, due to the random variability inherent in the project costs and benefits. In the business world, these simulations are a cornerstone for decision-making under uncertainty, allowing companies to model the probability of different outcomes in a process that cannot easily be predicted due to the intervention of random variables.

**Analysis**

**Part 1:**

We have run a 10,000 benefit-cost ratio simulation for both Dam #1 and Dam #2. Using the RAND() function, which generates uniform random values between 0 and 1, 10,000 random values were generated for each of the two dams' cost and benefit areas. The entire benefit divided by the total costs gives the benefit-cost ratio.

After simulating the benefit-cost ratios, we create frequency distributions for 𝛼1 (Dam #1) and 𝛼2 (Dam #2). The tabular frequency distribution is calculated using COUNTIF() function corresponding to its frequencies.

|  |  |
| --- | --- |
| min | 0.972 |
| max | 1.960 |
| range | 0.988 |
| classes | 100 |
| class width | 0.010 |

**Interpretation for Figure 1:**

The above graph is a representation of the observed frequencies of benefit-cost ratios for Dam #1. The histogram has a bell-like shape, which suggests that the benefit-cost ratios for Dam #1 are approximately normally distributed. This is indicated by the symmetric pattern where most of the data points cluster around a central peak and the frequencies taper off towards both ends of the x-axis. There is a spread of benefit-cost ratios on both sides of the peak, which shows variability in the simulated outcomes. The width of the bell curve suggests a range of possible outcomes from the simulation.

|  |  |
| --- | --- |
| min | 0.942 |
| max | 2.004 |
| range | 1.062 |
| classes | 100 |
| class width | 0.011 |

**Interpretation for Figure 2:**

The above graph displays the observed frequencies of benefit-cost ratios for DAM 2. The histogram displays a bell-like shape, indicating that the benefit-cost ratios for Dam #2 are approximately normally distributed. This distribution is symmetric around a central peak, suggesting the average benefit-cost ratio occurs most frequently. There is a spread around the central peak, indicating a range of potential benefit-cost ratios from the simulation. This histogram also seems to taper off symmetrically.

When comparing the histograms for Dam #1 and Dam #2, both histograms have similar bell-shaped curves, which suggests that both sets of benefit-cost ratios follow a normal distribution, with different means and standard deviations.

|  |  |  |
| --- | --- | --- |
| **Table 1** | | |
| **Dam 1** | **Observed** | **Theoretical** |
| Mean of the Total Benefits | 29.448 | 29.467 |
| SD of the Total Benefits | 2.292 | 4.504 |
| Mean of the Total Cost | 20.751 | 20.767 |
| SD of the Total Cost | 1.530 | 5.117 |
| Mean of the Benefit-cost Ratio | 1.427 | X |
| SD of the Benefit-cost Ratio | 0.152 | X |

|  |  |  |
| --- | --- | --- |
| **Table 2** | | |
| **Dam 2** | **Observed** | **Theoretical** |
| Mean of the Total Benefits | 30.732 | 30.700 |
| SD of the Total Benefits | 2.390 | 4.025 |
| Mean of the Total Cost | 22.039 | 22.067 |
| SD of the Total Cost | 1.715 | 5.200 |
| Mean of the Benefit-cost Ratio | 1.403 | X |
| SD of the Benefit-cost Ratio | 0.156 | X |

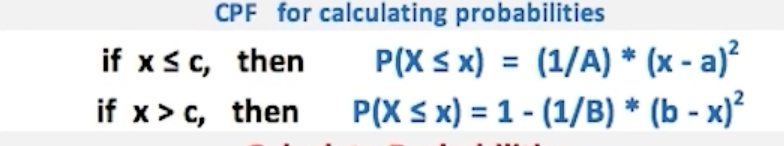
### Above Table 1 and Table 2 provide a list of both observed and theoretical statistical measures for the total benefits, total costs, and benefit-cost ratio for Dam 1 and Dam 2, respectively.

The mean of the total benefits and total costs for both DAMs are very close to their theoretical values, which suggests that the model used to calculate the theoretical values is consistent with the observed data. However, there is a notable difference in the standard deviations (SD) for both benefits and costs, with the theoretical SDs being almost twice as large as the observed SDs. This could suggest that the theoretical model predicts more variability than what is observed, or it might reflect the impact of certain assumptions in the theoretical model that do not perfectly match the observed data.

**Part 2:**

Benefit-cost ratios for previously created randomly generated values are used to compute estimates for a, b, and c. Using the following formula, the theoretical probabilities are computed for each class interval for the observations in Part 1(ii).

|  |  |
| --- | --- |
|  | **Estimates** |
| a | 0.972 |
| b | 1.960 |
| c | 1.349 |
| A  = (b-a)\*(c-a) | 0.372 |
| B  =(b-a)\*(b-c) | 0.604 |



Moving on, expected frequencies are obtained by multiplying the theoretical probabilities by the number of observations, which is 10,000 in this case. The chi-squared contribution for each class interval is then calculated using the formula: (Expected – Observed )2 / Expected. To calculate the total chi-squared test statistic, we sum the last column's values for all class intervals. This will determine whether or not the triangular distribution fits well.

Degree of freedom = number of bins (100) – number of parameters (3) –1; since a, b, and c have been employed, there are three parameters.

The hypothesis is as follow:  
**Null Hypothesis (H0):** There is a good fit between the triangular distribution.  
**Alternative Hypothesis (H1):** There is insufficient match between the triangular distribution and the data.

|  |  |
| --- | --- |
| **Table 3** | |
| **Degrees of Freedom** | **96.0000** |
| **Chi-squared Test Statistic** | **1640.042** |
| **Chi-squared P-value** | **0.0000** |

The p-value is less than 0.05, therefore, we reject the null hypothesis. This suggests that there is a statistically significant difference between the observed frequencies and the expected frequencies.

**Part 3:**

We can analyze the performance and risk associated with each project using table 4 below. α1 represents benefit-cost ratio for DAM 1 and α2 represents benefit-cost ratio for DAM 2.

|  |  |  |
| --- | --- | --- |
| **Table 4** | | |
|  | **𝛂1** | **𝛂2** |
| **Minimum** | 0.97181 | 0.94192 |
| **Maximum** | 1.95993 | 2.00433 |
| **Mean** | 1.42682 | 1.40298 |
| **Median** | 1.42375 | 1.39263 |
| **Variance** | 0.02322 | 0.02418 |
| **Standard Deviation** | 0.15239 | 0.15550 |
| **SKEWNESS** | 0.14676 | 0.29943 |
| **P(𝛂i > 2)** | 0.00000 | 0.00010 |
| **P(𝛂i > 1.8)** | 0.00780 | 0.00840 |
| **P(𝛂i > 1.5)** | 0.31530 | 0.25970 |
| **P(𝛂i > 1.2)** | 0.93430 | 0.91180 |
| **P(𝛂i > 1)** | 0.99970 | 0.99930 |

|  |  |
| --- | --- |
| **P(𝛂1 > 𝛂2)** | **0.550** |

**Performance Analysis:**

* **Central Tendency**: The mean and median of a1 are slightly higher than those of 𝛂2, suggesting that Dam 1 tends to have a better benefit-cost ratio than Dam 2.
* **Variability**: The variances and standard deviations are very similar, with Dam 2 being slightly higher, indicating a marginally greater spread in the benefit-cost ratio outcomes for Dam 2.
* **Risk and Distribution Shape**: The skewness of 𝛂1 is lower than 𝛂2, suggesting that Dam 1 has a distribution that is closer to normal, while Dam 2 has a distribution that is more skewed to the right. This might mean Dam 2 has a slightly higher chance of yielding very high benefit-cost ratios, despite a similar mean.

**Recommendation to Management:**

Dam 1 (𝛂1) shows a slightly higher mean and median benefit-cost ratio, a comparable level of risk and less skewness compared to Dam 2. These factors suggest that Dam 1 has a slightly better expected performance with a similar risk profile. Given the probabilities associated with higher benefit-cost ratios, Dam 1 also shows a marginally better chance of achieving higher ratios, further supporting its recommendation.

**Conclusion**

After conducting extensive Monte Carlo simulations and statistical analysis, we recommend Dam 1 for investment over Dam 2. This decision is because in Dam 1 has higher average benefit-cost ratio, lower skewness, and slightly better probabilities of exceeding higher benefit-cost thresholds, indicating a marginally better expected return for a similar level of risk. Our findings demonstrate a robust analytical approach that underscores the power of enterprise analytics in evaluating complex investment opportunities under uncertainty.

**References**

1. Frost, J. (2023, March 21). *How to Interpret P-values and Coefficients in Regression Analysis*. Statistics by Jim. <https://statisticsbyjim.com/regression/interpret-coefficients-p-values-regression/>
2. Kenton, W. (2023, November 2). *Monte Carlo Simulation: History, How it Works, and 4 Key Steps*. Investopedia. <https://www.investopedia.com/terms/m/montecarlosimulation.asp>
3. Frasca. (n.d) Lab: Triangular Probability Distribution. [Video].Panopto. https://northeastern.hosted.panopto.com/Panopto/Pages/Viewer.aspx?id=6f9e314d-3ab0-49bc-8f06-ac9300e649cf&start=2145.860645