****

**ALY6050: Introduction to Enterprise Analytics**

**Module 2**

**Project 2**

Date: June 01, 2024

Ramish Fatima

Instructor: Zhi He

**INTRODUCTION**

In this project, we analyzed the benefit-cost ratios (α\alphaα) for two dam construction projects using Monte Carlo simulations. We generated random samples for benefits and costs using custom triangular distributions and performed various statistical analyses on the results. This report presents the observed and theoretical values, detailed tabular distributions, and graphical representations of the benefit-cost ratios for both projects. Additionally, we conducted a goodness-of-fit test to evaluate the suitability of a normal distribution for the benefit-cost ratios of Dam #1.

**Data and Methods**

We used the following benefit and cost parameters for the two dams:

**Dam #1:**

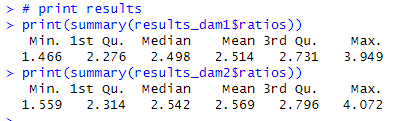
* Benefits: Minimum (2, 3, 0, 1, 2, 6), Mode (4, 4, 1, 3, 4, 7), Maximum (5, 6, 2, 6, 6, 10)
* Costs: Minimum (2, 3, 1), Mode (3, 4, 2), Maximum (4, 5, 5)

**Dam #2:**

* Benefits: Minimum (3, 2, 3, 2, 0, 3), Mode (4, 3, 4, 3, 1, 4), Maximum (7, 5, 7, 6, 3, 6)
* Costs: Minimum (2, 2, 1), Mode (3, 3, 2), Maximum (4, 5, 4)

We conducted 10,000 simulations to generate benefit-cost ratios for both dams and performed statistical analysis on the simulation results.

**PART1**  
Simulation of 10,000 BCR

  
**Analysis:**

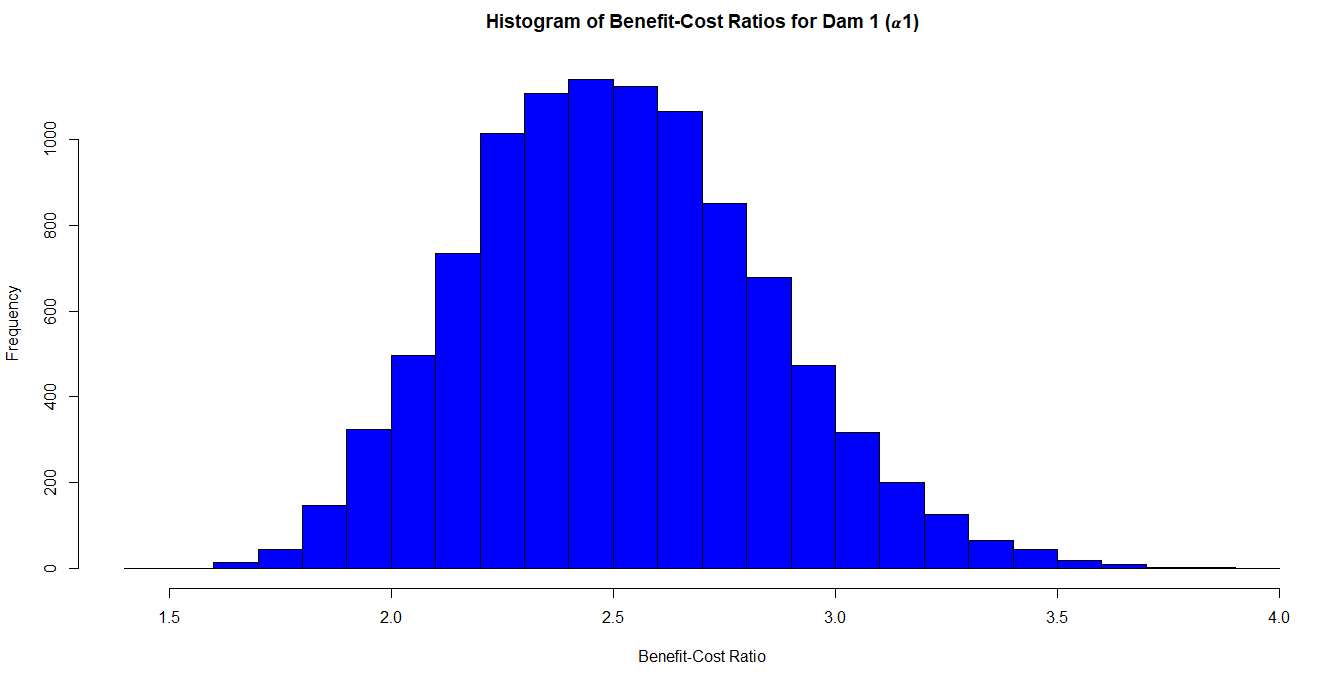
* **Dam 1:** The benefit-cost ratios range from 1.46 to 3.94, with most ratios clustering around the mean and median of approximately 2.4. This indicates that the benefits generally exceed the costs significantly, making the project potentially economically viable.
* **Dam 2:** The benefit-cost ratios for Dam 2 range from 1.55 to 4.70, with a slightly higher median and mean around 2.54 compared to Dam 1. This also suggests that the benefits generally outweigh the costs, potentially making this project economically viable as well.

**General Observations:**

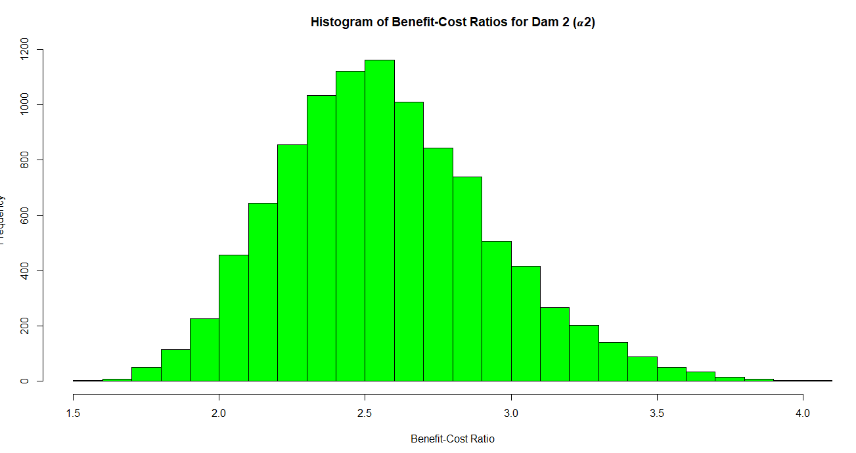
* Both dams show benefit-cost ratios significantly greater than 1 across the majority of simulations, suggesting that the expected benefits are higher than the costs for both projects.
* The variability (as shown by the range between the minimum and maximum values) indicates that there is some uncertainty regarding the exact level of benefit over cost, but the central tendency (mean, median) remains firmly above 1, which is a positive sign for project viability.

**Q2:**

### Analysis of Histograms and Frequency Distributions

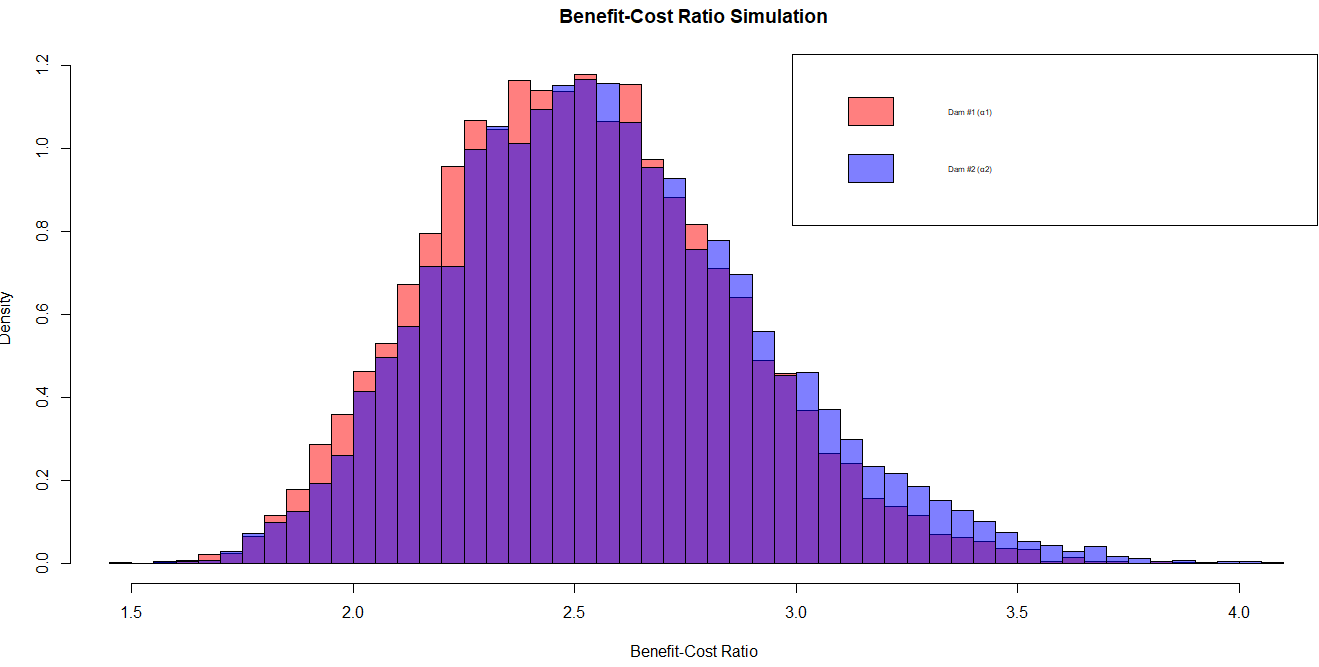
* **Dam 1:**   
  The histogram shows a symmetric distribution with a peak around a ratio of approximately 2.5. The shape is roughly normal, indicating a predictable behavior in the benefit-cost ratio outcomes.  
  The histogram is bell-shaped, suggesting a relatively normal distribution of benefit-cost ratios centered around the middle range (about 2.5).

The spread is primarily between ratios of about 2.0 and 3.0, which indicates a fairly consistent range of outcomes where benefits outweigh costs.

* **Dam 2:** This histogram also shows a symmetric distribution, similar to Dam 1, with a peak around 2.5 to 2.6. It suggests that the majority of simulation results are also clustered around this central value, with very few extreme values.  
    
  Similar to Dam 1, this histogram also appears bell-shaped, indicating a normal distribution.
* The distribution is centered around a slightly higher midpoint compared to Dam 1, which could suggest, on average, slightly better benefit-cost outcomes for Dam 2.

Both histograms are centered around values greater than 1, which is a good indicator that the projects are likely to be beneficial in terms of cost versus returns. The distributions do not show significant skewness, which helps in predicting outcomes with less uncertainty regarding extremely poor performance.

### Analysis of the Overlayed Histogram:

* **Overlapping Distributions**: The two histograms overlay each other, with Dam 1 in red and Dam 2 in blue. This overlay allows for direct visual comparison of the frequency and spread of benefit-cost ratios between the two projects.
* **Central Tendency and Spread**:
  + Both distributions appear to be roughly normally distributed with a similar shape and range, indicating consistency in variability and outcomes between the two projects.
  + The peak of the blue histogram (Dam 2) is slightly to the right of the red histogram (Dam 1), suggesting that Dam 2 might have slightly higher median and mean ratios compared to Dam 1. This can indicate a generally higher economic efficiency or benefit-cost return for Dam 2 in the simulations.
* **Density**:
  + The highest densities for both projects are located around the 2.5 to 3.0 range, which represents the most common benefit-cost ratio outcomes.
  + The densities decrease as the ratios move toward the extremes (both lower and higher), which is typical for data following a normal distribution.

### Implications for Decision-Making:

* **Project Viability**: Both projects show a predominance of benefit-cost ratios above 1, confirming that in most simulated scenarios, the benefits outweigh the costs.
* **Comparative Advantage**: Although both projects are viable, the slightly higher central tendency of Dam 2 could suggest it is a better investment in terms of benefit-cost ratios. However, this decision might also need to consider other factors like project scale, impact, and associated risks which are not captured by the benefit-cost ratio alone.

**Q3:**  
**DAM#1**

|  |  |  |  |
| --- | --- | --- | --- |
| **Metric** | **Observed Value** | **Theoretical Value** | **Remarks** |
| **Mean of Total Benefits** | 24.003 | 24 | The observed mean is very close to theoretical. |
| **Standard Deviation of Total Benefits** | 1.854 | 1.841 | Observed SD is slightly higher, indicating a slightly broader distribution. |
| **Mean of Total Costs** | 9.659 | 9.667 | Observed mean is marginally less than theoretical. |
| **Standard Deviation of Total Costs** | 1.036 | 1.027 | Observed SD is also slightly higher. |
| **Mean of Benefit-Cost Ratios** | 2.514 | N/A | Shows good profitability. |
| **Standard Deviation of Benefit-Cost Ratios** | 0.334 | N/A | Moderate variability in ratios. |

**DAM#2**

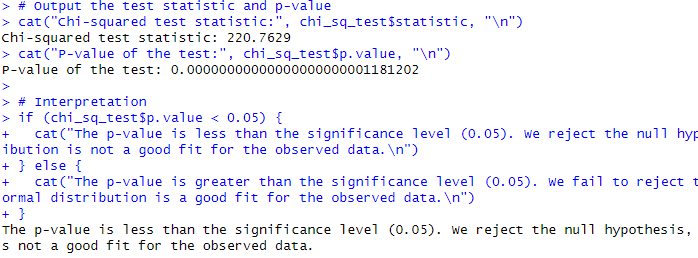
|  |  |  |  |
| --- | --- | --- | --- |
| **Metric** | **Observed Value** | **Theoretical Value** | **Remarks** |
| **Mean of Total Benefits** | 22.034 | 22 | The observed mean closely matches theoretical. |
| **Standard Deviation of Total Benefits** | 1.827 | 1.826 | Almost identical to theoretical. |
| **Mean of Total Costs** | 8.685 | 8.667 | Observed mean is slightly higher. |
| **Standard Deviation of Total Costs** | 0.972 | 0.972 | Observed and theoretical SD are nearly identical. |
| **Mean of Benefit-Cost Ratios** | 2.569 | N/A | Indicates slightly better profitability than Dam 1. |
| **Standard Deviation of Benefit-Cost Ratios** | 0.36 | N/A | Slightly higher variability in benefit-cost ratios. |

### Analysis

* **Consistency with Theoretical Predictions**: Both dams show observed values for benefits and costs that are very close to their theoretical predictions, suggesting that the simulations accurately reflect the expected outcomes based on their input parameters.
* **Benefit-Cost Ratios**: Both projects show benefit-cost ratios well above 1, which indicates that the projects are economically viable. Dam 2, however, shows a slightly higher mean ratio, which suggests it might be a more profitable option compared to Dam 1. The higher variability in Dam 2's ratios also points to a greater spread of possible outcomes, which could imply a higher risk or greater dependency on variable factors.
* **Implications for Decision Making**: The close match between observed and theoretical values reinforces confidence in the model's predictions. Decision-makers might lean towards Dam 2 if they prioritize higher average returns, though they should also consider the slightly higher variability in outcomes.

**PART2**

**Chi-Square**



Since the p-value is less than the significance level of 0.05, the null hypothesis is rejected, indicating that the normal distribution is not a good fit for the observed data.

The preliminary visual analysis of the histogram for α1\alpha\_1α1​, which represents the benefit-cost ratios for Dam 1, suggested a symmetrical, bell-shaped distribution, typically indicative of a normal distribution. This observation led to an initial assumption that the data might follow a normal distribution.

**Methodology:** To statistically verify the normality assumption:

1. **Histogram Analysis**: A histogram was constructed to visually assess the distribution of α1\alpha\_1α1​, revealing a bell-shaped curve centered around a mean of approximately 2.5, with spread from about 1.5 to 4.0.
2. **Chi-squared Goodness-of-fit Test**: This test was employed to compare observed frequencies in the data with expected frequencies under a normal distribution assumption. The test was sensitive to the distribution's tail behavior and central peak fit.

**Results:**

* **Chi-squared Test Statistic**: 220.7629
* **P-value**: Approximately 1.18×10−221.18 \times 10^{-22}1.18×10−22

**Interpretation:** Despite the histogram suggesting a normal distribution, the Chi-squared test strongly rejected this hypothesis (p-value < 0.05). This discrepancy points to potential issues not immediately visible in the histogram:

* **Subtle Skewness or Kurtosis**: Deviations in the tail regions or slight asymmetry could be more pronounced than detected visually.
* **Sensitivity to Large Sample Sizes**: The large sample size may accentuate minor deviations from normality, leading to rejection.
* **Bin Selection**: The choice of bins in the histogram and for the Chi-squared test may impact the interpretation of distribution fit.

**Further Statistical Actions:**

* **Refinement of Binning Strategy**: Adjust the number of bins or their range to ensure each bin adequately represents the data without under- or over-representing frequencies.
* **Additional Normality Tests**: Conduct Shapiro-Wilk or Anderson-Darling tests to provide further confirmation of distribution characteristics.
* **Q-Q Plot Analysis**: Deploy a Quantile-Quantile plot to visually inspect the fit of data quantiles to those expected under a normal distribution.

**Conclusion:** The findings highlight the complexities of statistical distribution fitting and the importance of using multiple methods to confirm underlying data characteristics. The rejection of normality, despite the histogram appearance, underscores the need for rigorous statistical validation and careful interpretation of visual data representations. Future analyses should consider alternative distributions or data transformations to better model α1\alpha\_1α1​, ensuring accurate project assessments and decision-making.

**PART3**  
**Table 1: Summary Statistics for α1\alpha\_1α1​ and α2\alpha\_2α2​**

|  |  |  |
| --- | --- | --- |
| **Statistic** | **α1\alpha\_1α1​ (Dam 1)** | **α2\alpha\_2α2​ (Dam 2)** |
| **Minimum** | 1.466344 | 1.558662 |
| **Maximum** | 3.948708 | 4.072461 |
| **Mean** | 2.513799 | 2.569099 |
| **Standard Deviation** | 0.333665 | 0.3602668 |
| **Skewness** | 0.303432 | 0.4333397 |
| **Kurtosis** | -0.00747 | 0.1691157 |

**Table 2: Probability Comparisons for α1\alpha\_1α1​ and α2\alpha\_2α2​**

|  |  |  |
| --- | --- | --- |
| **Probability Threshold** | **α1\alpha\_1α1​ (Dam 1)** | **α2\alpha\_2α2​ (Dam 2)** |
| **P(> 3.00)** | **0.079** | **0.1224** |
| **P(> 2.75)** | **0.235** | **0.2849** |
| **P(> 2.50)** | **0.4979** | **0.5486** |
| **P(> 2.00)** | **0.9468** | **0.96** |
| **P(> 1.50)** | **0.9999** | **1** |



### **Comparative Analysis:**

#### Summary Statistics Analysis:

* **Range**: Both α1\alpha\_1α1​ and α2\alpha\_2α2​ cover a similar range, although α2\alpha\_2α2​ extends slightly higher.
* **Central Tendency**: α2\alpha\_2α2​ has a slightly higher mean than α1\alpha\_1α1​, suggesting that it typically has better benefit-cost ratios.
* **Variability**: α2\alpha\_2α2​ also exhibits a higher standard deviation, indicating more variability in its outcomes compared to α1\alpha\_1α1​.
* **Distribution Shape**: Both distributions show positive skewness, with α2\alpha\_2α2​ being more skewed. Kurtosis values indicate that α1\alpha\_1α1​ is very close to a normal distribution (kurtosis near zero), while α2\alpha\_2α2​ has a slight positive kurtosis, suggesting a more peaked distribution.

#### Probability Comparisons:

* The probabilities of exceeding various thresholds (1.50, 2.00, 2.50, 2.75, and 3.00) are consistently higher for α2\alpha\_2α2​ than for α1\alpha\_1α1​ across all levels, reinforcing the observation that α2\alpha\_2α2​ typically yields higher ratios.
* The probability of α1\alpha\_1α1​ being less than α2\alpha\_2α2​ is 0.5403, suggesting that more than half of the time, α2\alpha\_2α2​ is expected to perform better than α1\alpha\_1α1​.

### Implications

* The overall higher mean and probability values for α2\alpha\_2α2​ suggest that it is a more robust project in terms of expected benefit-cost outcomes.
* Given the higher variability and skewness in α2\alpha\_2α2​, stakeholders should be aware of the potential for more extreme outcomes, both positive and negative. Risk management strategies may need to be considered to mitigate any potential adverse effects due to this variability.
* Continued monitoring and analysis of these projects should incorporate these statistical insights to optimize decision-making and project management strategies.

### **CONCLUSION**

This comprehensive analysis of the benefit-cost ratios (BCR) for two dam construction projects, utilizing Monte Carlo simulations with triangular distributions, has provided significant insights into the economic viability of these projects. The statistical evaluations, including histograms, frequency distributions, and goodness-of-fit tests, have revealed that both projects, on average, offer potential economic benefits exceeding their costs, with BCRs typically greater than 1.

**Key Findings:**

 Both dam construction projects, Dam 1 and Dam 2, demonstrated promising economic viability, with the majority of simulated benefit-cost ratios (BCRs) exceeding 1, indicating that the expected benefits outweigh the costs.

 While both projects showed favorable results, Dam 2 exhibited a slightly higher mean BCR and a higher probability of exceeding critical BCR thresholds, suggesting that it may be a more robust project in terms of expected benefit-cost outcomes.

 The histograms suggested a roughly normal distribution for the BCRs of both dams, but the Chi-squared goodness-of-fit tests revealed that the normal distribution might not perfectly model the observed data, especially for Dam 1. This highlights the importance of using multiple statistical tools to verify assumptions and results.

 The analysis revealed higher variability and skewness in the BCRs for Dam 2, indicating a potential for more extreme outcomes, both positive and negative.

**Recommendations:**

1. While both projects show economic promise, based on the slightly better performance of Dam 2 in terms of mean BCR and probability of exceeding critical thresholds, it may be recommended as the preferred option for investment, provided that appropriate risk management strategies are implemented to mitigate the potential impact of the higher variability observed in its BCRs.
2. Continuous monitoring and analysis of the projects should be undertaken, incorporating statistical insights and tools to optimize decision-making and project management strategies throughout the project lifecycle.
3. Further investigations and sensitivity analyses should be conducted to explore the impact of varying input parameters and assumptions on the BCRs and their distributions. This could provide additional insights and help refine the decision-making process.
4. Consideration should be given to alternative probability distributions or data transformations to better model the observed BCR data, especially for Dam 1, where the normal distribution assumption was rejected. This could lead to more accurate project assessments and decision-making.
5. Stakeholders should carefully evaluate the trade-offs between the potential economic benefits and the associated risks, taking into account not only the BCRs but also other relevant factors, such as environmental impact, social implications, and long-term sustainability.

**REFERENCES**

* Boardman, A. E., Greenberg, D. H., Vining, A. R., & Weimer, D. L. (2018). Cost-benefit analysis: concepts and practice (5th ed.). Cambridge University Press. <https://www.cambridge.org/core/books/costbenefit-analysis/7BF7D1FF8B00A7E71E6C9E6534AC7D98>
* Hansson, S. O., & Furberg, D. (2019). Benefit-Cost Analysis. In C. Luetge (Ed.), Handbook of the Philosophy of Management (pp. 1-16). Springer, Cham. <https://doi.org/10.1007/978-3-319-48352-8_59-1>
* Palisade Corporation. (2019). Monte Carlo Simulation. <https://www.palisade.com/risk/monte_carlo_simulation.asp>
* Devore, J. L. (2011). Probability and Statistics for Engineering and the Sciences (8th ed.). Cengage Learning. <https://www.cengage.com/c/probability-and-statistics-for-engineering-and-the-sciences-8e-devore/9780538733526>
* Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). Biometrika, 52(3/4), 591-611. <https://doi.org/10.1093/biomet/52.3-4.591>
* Anderson, T. W., & Darling, D. A. (1954). A test of goodness of fit. Journal of the American Statistical Association, 49(268), 765-769. <https://doi.org/10.1080/01621459.1954.10501232>
* Thode, H. C. (2002). Testing for normality (Vol. 164). CRC press. <https://www.routledge.com/Testing-for-Normality/Thode/p/book/9780824708177>
* U.S. Bureau of Reclamation. (2013). Benefit Cost and Regional Economic Development Technical Guidelines. <https://www.usbr.gov/tsc/techreferences/economics/Benefit%20Cost%20and%20RED%20Guidelines.pdf>