Module-2: Benefit-Cost Analysis of Dam Construction Projects

By: Nithin Reddy.P (NEU ID:002896440)

Subject:ALY6050: Introduction to Enterprise Analytics
Under the guidance of Shahram Sattar
Submission Date:03-Mar-2024

Introduction:

In this assignment, we are about to conduct cost benefit analysis of two dams (one in Southwest Georgia and other in North Carolina) and recommend one of the two projects to JET Corp. The associated benefits and cost for 30 years annualized period of respective dams are provided in millions, they identified six areas of benefits and two categories associated with cost.

Problem:

The aim is to perform cost benefit analysis, which is obtained by calculating benefit to cost ratio. If its greater then 1, it indicates that there are more benefits and less benefits in the other case. The case study is segregated into three parts. In the initial part, we calculate the profitability ratio of two projects, then conduct frequency distribution and statistics for observed and theoretical values. Secondly, we perform Chi-Square goodness of fit to determine whether there triangular distribution represents uncertainty in the costs and benefits. Finally, we compare findings from both projects and recommend the best project idea.

Part-1:

(i): Perform 10,000 simulations on Dam1 and Dam2, calculate the Benefit to cost ratio and display then as variable Alpha1 and Alpha2.

The given data, for Benefits and Costs for Dam1 and Dam2 have been stored in a matrix for further analysis, please refer to Fig-1. In order to calculate the Total Benefit Cost Ratio for each Dam, I have created a Cost function (Fig-2), it simulates the ratio for mentioned number of iterations. The two arguments **benefits_dam** and **costs_dam** store Minimum, Maximum and Mode respectively. The **IF** loop conducts the simulation for 10,000 times and creates random sales for the mentioned arguments. It then finds the Total Benefits and Total Costs for each instance and lastly computes the requested ratio and stores the variables in a vector of benefit-to-cost ratios (**sampels**), the total benefits (**total_benefits**), and the total costs (**total_cost**).

Cost Benefit Analysis:

As we now have the required data, we have conducted the analysis of cost-benefit and assigned to respective alphas (Fig-3). Moreover, the respective max, min and mode values have been extracted, by unlisten alpha values. To find the mode, we have created a mode_function (Fig-4 and 5).

(ii): Tabular and a graphical frequency distribution for alpha1 and 2

Further by using Table and breaks (40), we have created tables for respective alphas (Fig-6). Furthermore, by using Hist function we have created Frequency distribution of alpha-1 and alpha-2 (Fig-7 and 8). Form the figures, we can see that the distribution follows a Triangle with the minimum, maximum and mean values. Using this assumption that they follow this distribution, we move forward in our analysis to find the theoretical and observed values.

Form Fig-9, we could observe that the mean and standard deviation have been assigned to the respective variables. To find the theoretical values, we have created a function (Fig-10) and the values have been allocated to the respective variables and the values are stored in data frame (Fig-11). The similar approach has been implemented to find the observed and theoretical values of Dam-2 (Fig-12).

Part-2: Triangle Distribution and Chi Square Goodness of Fit

To populate the triangle distribution plot, we have extracted the respective min, max and mode values earlier (Fig-5). We built the pot by creating a **dtriangular** function (Fig-13), by using the earlier derived values (Fig-5), we have built a triangle distribution plot (Fig-14) for alpha-1. From the plot, we could clearly see that the mean, minimum and mode are placed to the left of the peak.

1. **Mean to the Left of the Peak**: In a symmetric triangular distribution (where the peak is at the midpoint between the minimum and maximum), the mean coincides with the peak. However, if the mean is shifted to the left of the peak, it suggests that the distribution is skewed towards the lower end of the range. This skewness indicates that there is a higher probability of observing values below the mean compared to values above the mean.
2. **Minimum to the Left of the Peak**: Similarly, if the minimum value is positioned to the left of the peak, it means that the distribution extends further towards the lower end, and there is a non-zero probability of observing values below the minimum.
3. **Mode to the Left of the Peak**: The mode being to the left of the peak indicates that the most likely value is also skewed towards the lower end of the distribution. This could happen when the underlying phenomenon has a bias or preference towards lower values.

Chi-Square Hypothesis Test:

To explore whether triangular distribution adequately represents the uncertainty in the costs and benefits we have conducted goodness of fit test with the below hypothesis.

H0: The use of a triangular distribution adequately represents the uncertainty in the costs and benefits associated with cost-benefit analysis.

H1: The use of a triangular distribution does not adequately represents the uncertainty in the costs and benefits associated with cost-benefit analysis.

From the test (Fig-15) we could clearly see that the p-value (0.00000000000000022) is less than 0.05, hence we reject the null hypothesis, suggesting that the triangular distribution does not adequately represent the uncertainty in the costs and benefits associated with cost-benefit analysis.

Similarly, the triangular plot and chi-square text is conducted on Dam-2 and we have observed the similar graph and same results as of Dam-1 (Fig-16 and Fig-17).

Part-3: Results Simulation

In the final step, we have used **moments** library and found the skewness of alpha-1 &2 along with the respective min, max, mean, median, variance and respective probabilities being greater than certain values (2, 1.8, 1.5, 1.2, and 1) using mean function (Fig-18). We tried to check whether probability of C-B ratio of Dam 1 will be higher and observed that Dam-1 outperform Dam-2 (Fig-19).

Summary:

Triangle Plot Analysis: A triangle plot was constructed, where the mean, minimum, and mode were observed to be positioned to the left side of the peak. This positioning suggests a skewness towards lower values in the distribution.

Hypothesis Testing: The p-value obtained from a chi-squared goodness-of-fit test for both Dam-1 and Dam-2 was less than 0.05. Therefore, the null hypothesis was rejected, indicating that the use of a triangular distribution does not adequately represent the uncertainty in the costs and benefits associated with the cost-benefit analysis for both dams.

Descriptive Statistics: Descriptive statistics were computed for both Dam-1 (Alpha1) and Dam-2 (Alpha2) projects, including minimum, maximum, mean, median, variance, standard deviation, skewness, and probabilities of alpha values exceeding certain thresholds.

Comparison of Projects: It was estimated that Alpha1 consistently outperforms Alpha2 in terms of their benefits and costs, with a probability of approximately 0.6002.

Conclusion:

In conclusion, it can be concluded that despite the skewness observed in the triangle plot and the rejection of the null hypothesis in the chi-squared test, Alpha1 consistently outperforms Alpha2 across various metrics. This suggests that, despite the limitations of the triangular distribution model, Alpha1 appears to be a more favourable option compared to Alpha2 in terms of its benefits and costs.

Appendix:

Scenario-1

Fig-1: Given Data

```
## Given Data ##

# Matrix for Dam 1
benefits_dam1 <- matrix(c(1.1, 2, 2.8,
                          8, 12, 14.9,
                          1.4, 1.4, 2.2,
                          6.5, 9.8, 14.6,
                          1.7, 2.4, 3.6,
                          0, 1.6, 2.4), nrow = 6, byrow = TRUE)
cost_dam1 <- matrix(c(13.2, 14.2, 19.1,
                       3.5, 4.9, 7.4), nrow = 2, byrow = TRUE)

# Matrix for Dam 2
benefits_dam2 <- matrix(c(2.1, 3, 4.8,
                          8.7,12.2, 13.6,
                          2.3, 3, 3,
                          5.9, 8.7, 15,
                          0, 3.4, 3.4,
                          0, 1.2, 1.8), nrow = 6, byrow = TRUE)
cost_dam2 <- matrix(c(12.8, 15.8, 20.1,
                       3.8, 5.7, 8), nrow = 2, byrow = TRUE)
```

Fig-2: Function for calculating Cost

```
##Creating a Cost function to perform simulation


cost <- function(benefits_dam, costs_dam)
{
  # Seed to generate the same pseudo random values
  set.seed(20)
  n = 10^4  # 10,000 simulations
  sampels <- numeric(n)
  sum_of_benefits <- numeric(n)
  sum_of_costs <- numeric(n)
  for(i in 1:n)
  {
    total_benefits <- apply(benefits_dam, 1, function(x) runif(1, min = x[1], max = x[3]))
    total_cost <- apply(costs_dam, 1, function(x) runif(1, min = x[1], max = x[3]))
    benefit_to_cost_ratio <- sum(total_benefits) / sum(total_cost )
    sampels[i] <- benefit_to_cost_ratio
    sum_of_benefits[i] <- sum(total_benefits)
    sum_of_costs[i] <- sum(total_cost)
  }
  output <- list(sampels, total_benefits, total_cost)
  return(output)
}
```

Fig-3: Cost and Benefit analysis:

```
> #Cost Benefit analysis
>
> alpha1 <- cost(benefits_dam1, cost_dam1)
> alpha1
[[1]]
 [1] 1.9518616 1.5464168 1.4223644 1.4773291 1.3330816 1.1359619 1.5806858 1.0388007 1.4991037
[10] 1.3380641 1.7009202 1.0439314 1.2720547 1.6902935 1.5593140 1.4465602 1.3390621 1.4254664
[19] 1.5687651 1.0716763 1.3533793 1.4546239 1.9196730 1.1478948 1.5019682 1.4874427 1.6509213
[28] 1.4146913 1.6688956 1.3304337 1.4403641 1.4934761 1.1295149 1.6196381 1.1685288 0.9687547


> alpha2 <- cost(benefits_dam2, cost_dam2)
> alpha2
[[1]]
 [1] 2.0032200 1.5292950 1.4793183 1.4908854 1.3399073 1.1752036 1.5473816 0.9819880 1.5761017
[10] 1.2903491 1.6781367 1.0389654 1.2407292 1.6720661 1.5101035 1.4948899 1.4181339 1.4900325
[19] 1.5682303 1.0240276 1.3403491 1.4111455 2.0253282 1.1911676 1.4993488 1.4491987 1.6396633
[28] 1.3835357 1.6471585 1.3281394 1.3959324 1.5846027 1.0712417 1.6151785 1.1028387 0.9687213
[37] 1.1126141 1.6757693 1.4720225 1.2419215 1.2619780 1.0805258 1.2269880 1.5024457 1.4492646
```

Fig-4: Function for Mode :

```
# Define a function to calculate the mode
calculate_mode <- function(x) {
  ux <- unique(x)
  ux[which.max(tabulate(match(x, ux)))]
}
mode_value <- calculate_mode(data)
mode_value
```

Fig-5: Extracting Min, Max, Median and Mode

```
> data <- unlist(alpha1)
> mean_alpha1 <- mean(data)
> mean_alpha1
[1] 1.385684
> min_alpha1 <- min(data)
> min_alpha1
[1] 0.8309753
> mode_alpha1 <- mode(data)
> mode_alpha1
[1] "numeric"
> max_alpha1 <- max(data)
> max_alpha1
[1] 18.6489
> # Define a function to calculate the mode
> calculate_mode <- function(x) {
+   ux <- unique(x)
+   ux[which.max(tabulate(match(x, ux)))]
+ }
> mode_value <- calculate_mode(data)
> mode_value
[1] 1.951862
```

```
> data2 <- unlist(alpha2)
> mean_alpha2 <- mean(data2)
> mean_alpha2
[1] 1.375719
> min_alpha2 <- min(data2)
> min_alpha2
[1] 0.8207727
> max_alpha2 <- max(data2)
> max_alpha2
[1] 19.54186
> mode_value2 <- calculate_mode(data2)
> mode_value2
[1] 2.00322
```

Fig-6: The Alpha tables

```
> alpha1_tabel <- table(cut(alpha1[[1]], breaks = 40))
> alpha2_tabel<- table(cut(alpha2[[1]], breaks = 40))
> alpha1_tabel

 (0.83,0.863] (0.863,0.895] (0.895,0.927] (0.927,0.959] (0.959,0.991]  (0.991,1.02]   (1.02,1.05]
            2            15            26            30            77           111           142
  (1.05,1.09]   (1.09,1.12]   (1.12,1.15]   (1.15,1.18]   (1.18,1.21]   (1.21,1.25]   (1.25,1.28]
          203           304           318           446           516           530           523
  (1.28,1.31]   (1.31,1.34]   (1.34,1.37]   (1.37,1.41]   (1.41,1.44]   (1.44,1.47]    (1.47,1.5]
          602           632           611           645           557           528           491
   (1.5,1.53]   (1.53,1.57]    (1.57,1.6]    (1.6,1.63]   (1.63,1.66]   (1.66,1.69]   (1.69,1.73]
          477           397           360           296           248           208           189
  (1.73,1.76]   (1.76,1.79]   (1.79,1.82]   (1.82,1.85]   (1.85,1.89]   (1.89,1.92]   (1.92,1.95]
          134           101            67            55            48            41            25
  (1.95,1.98]   (1.98,2.01]   (2.01,2.05]   (2.05,2.08]   (2.08,2.11]
           20            14             8             1             2
> alpha2_tabel

(0.819,0.854] (0.854,0.888] (0.888,0.922] (0.922,0.955] (0.955,0.989]  (0.989,1.02]   (1.02,1.06]
            6            19            32            57            95           145           194
  (1.06,1.09]   (1.09,1.12]   (1.12,1.16]   (1.16,1.19]   (1.19,1.22]   (1.22,1.26]   (1.26,1.29]
          262           346           469           498           532           568           594
  (1.29,1.32]   (1.32,1.36]   (1.36,1.39]   (1.39,1.43]   (1.43,1.46]   (1.46,1.49]   (1.49,1.53]
          622           623           587           565           546           506           439
  (1.53,1.56]   (1.56,1.59]   (1.59,1.63]   (1.63,1.66]   (1.66,1.69]   (1.69,1.73]   (1.73,1.76]
          402           372           267           272           222           182           130
   (1.76,1.8]    (1.8,1.83]   (1.83,1.86]    (1.86,1.9]    (1.9,1.93]   (1.93,1.96]     (1.96,2]
           99            96            71            48            41            32            23
     (2,2.03]   (2.03,2.06]    (2.06,2.1]    (2.1,2.13]   (2.13,2.17]
           20             9             7             1             1
```

Fig-7: Frequency distribution

```
> hist(alpha1[[1]], breaks = 40, main = "Histogram of Alpha-1", xlab = "Benefit to Ratio")
> hist(alpha2[[1]], breaks = 40, main = "Histogram of Alpha-2", xlab = "Benefit to Cost Ratio")
```

Fig-8: Histogram of Alpha-1 and Alpha-2:

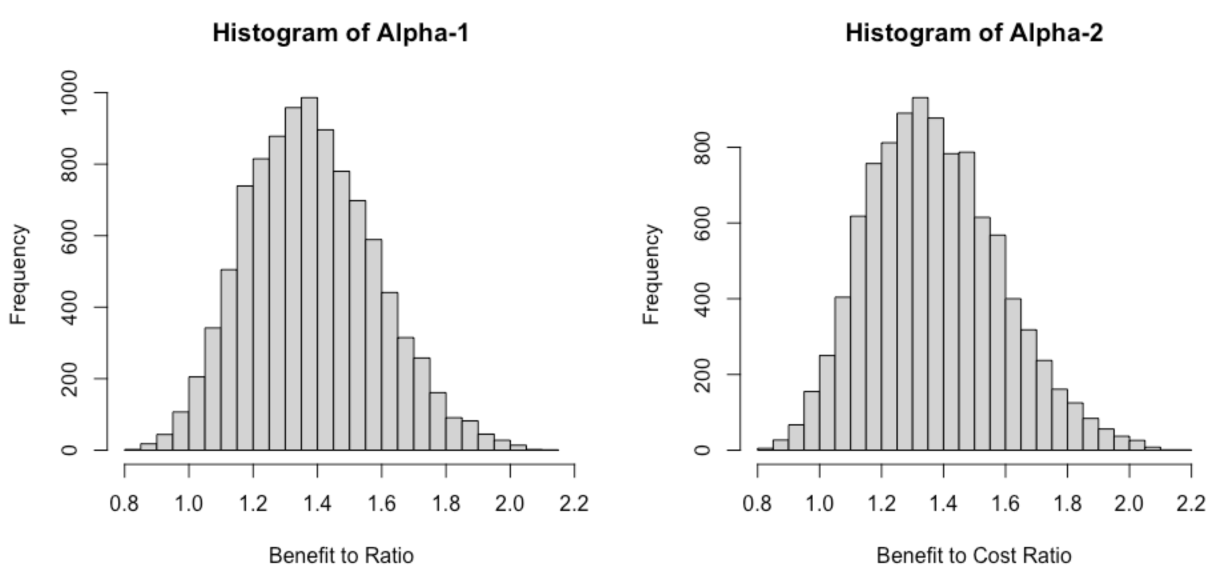


Fig-9: Observed frequencies of Dam-1

```
> dam1_mean_of_benefits <- mean(alpha1[[2]])
> dam1_sd_of_benefits <- sd(alpha1[[2]])
>
> dam1_mean_of_cost <- mean(alpha1[[3]])
> dam1_sd_of_cost<- sd(alpha1[[3]])
>
> dam1_mean_of_cost_to_benefit <- mean(alpha1[[1]])
> dam1_sd_of_cost_to_benefit  <- sd(alpha1[[1]])
```

Fig-10: Theoretical frequencies of Dam-1

```
> theoretical_probability <- function(matrix_m)
+ {
+   mean <- apply(matrix_m, 1, function(x) ((x[1] + x[3])/2))
+   sd <- apply(matrix_m, 1, function(x) ((x[3] - x[1])/(2 * sqrt(3))))
+   return (list(mean, sd))
+ }
>
> benefit_values <- theoretical_probability(benefits_dam1)
> mean_benefit <- sum(benefit_values[[1]])
> sd_benefit <- sum(benefit_values[[2]])
>
> cost_values <-  theoretical_probability(cost_dam1)
> mean_cost <- sum(cost_values[[1]])
> sd_cost <- sum(cost_values[[2]])
```

Fig-11: Dam-1 Data Frame

```
> # Create a dataframe.
> df_dam1 <- data.frame(Dam1 = c("Mean of Total Benefits", "SD of Total Benefits",
+                               "Mean of Total Costs","SD of Total Costs",
+                               "Mean of Benefit Cost Ratio", "Standard Deviation of Benefit Cost Ratio"),
+                      Observed = c(dam1_mean_of_benefits, dam1_sd_of_benefits,
+                                   dam1_mean_of_cost , dam1_sd_of_cost ,
+                                   dam1_mean_of_cost_to_benefit, dam1_sd_of_cost_to_benefit),
+                      Theoretical = c(mean_benefit, sd_benefit, mean_cost, sd_cost, '*', '*'))
> df_dam1
                                      Dam1   Observed      Theoretical
1                   Mean of Total Benefits  5.2834372             29.6
2                     SD of Total Benefits  5.2027277 6.29311793416692
3                      Mean of Total Costs 12.4267967             21.6
4                        SD of Total Costs  8.7993809 2.82901631902917
5               Mean of Benefit Cost Ratio  1.3811371                *
6 Standard Deviation of Benefit Cost Ratio  0.2017751                *
```

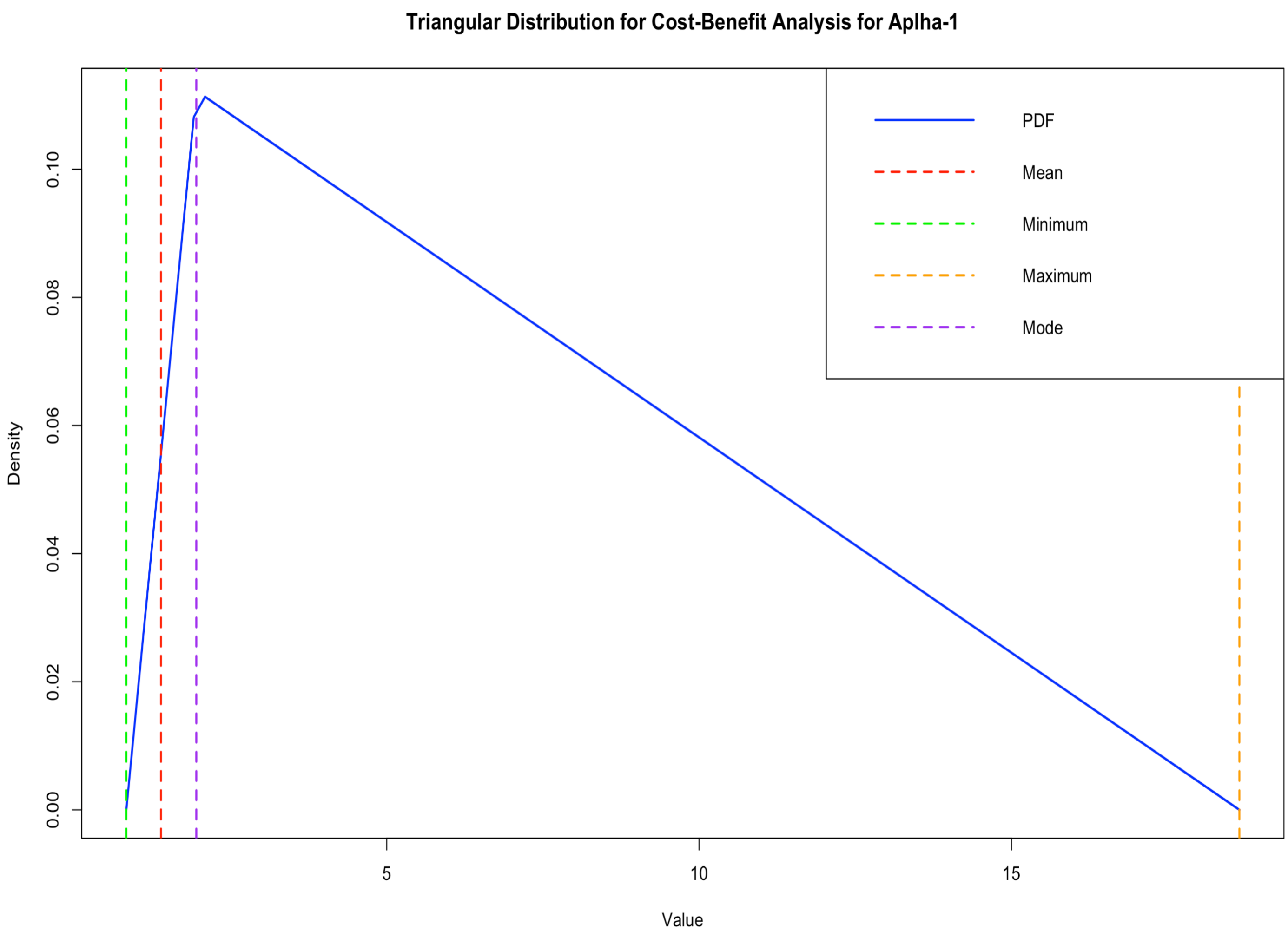
Fig-12: Data Frame for Dam-2

```
> # Observed Values
> dam2_mean_of_benefits  <- mean(alpha2[[2]])
> dam2_sd_of_benefits  <- sd(alpha2[[2]])
>
> dam2_mean_of_cost <- mean(alpha2[[3]])
> dam2_sd_of_cost  <- sd(alpha2[[3]])
>
> dam2_mean_of_cost_to_benefit <- mean(alpha2[[1]])
> dam2_sd_of_cost_to_benefit <- sd(alpha2[[1]])
>
> benefit_values2 <- theoretical_probability(benefits_dam2)
> mean_benefit2 <- sum(benefit_values2[[1]])
> sd_benefit2 <- sum(benefit_values2[[2]])
>
> cost_values2 <-  theoretical_probability(cost_dam2)
> mean_cost2 <- sum(cost_values2[[1]])
> sd_cost2 <- sum(cost_values2[[2]])
>
> # Create a dataframe.
> df_dam2 <- data.frame(Dam2 = c("Mean of Total Benefits", "SD of Total Benefits",
+                                  "Mean of Total Costs","SD of Total Costs",
+                                  "Mean of Benefit Cost Ratio", "Standard Deviation of Benefit Cost Ratio"),
+                        Observed = c(dam2_mean_of_benefits, dam2_sd_of_benefits,
+                                     dam2_mean_of_cost , dam2_sd_of_cost ,
+                                     dam2_mean_of_cost_to_benefit, dam2_sd_of_cost_to_benefit),
+                        Theoretical = c(mean_benefit2, sd_benefit2, mean_cost2, sd_cost2, '*', '*'))
> df_dam2
                                      Dam2   Observed      Theoretical
1                   Mean of Total Benefits  5.2271621             30.3
2                     SD of Total Benefits  4.7255709 6.52405804184277
3                      Mean of Total Costs 13.1273030            22.35
4                        SD of Total Costs  9.0715502 3.31976404784035
5               Mean of Benefit Cost Ratio  1.3710580                *
6 Standard Deviation of Benefit Cost Ratio  0.2137134                *
```

Fig-13: dtriangular Function

```
# Define a custom function to calculate the PDF of the triangular distribution
dtriangular <- function(x, lower, mode, upper) {
  # Calculate PDF of the triangular distribution
  pdf <- ifelse(x < mode,
                2*(x - lower)/((upper - lower)*(mode - lower)),
                2*(upper - x)/((upper - lower)*(upper - mode)))
  pdf[x < lower | x > upper] <- 0   # Set PDF to 0 outside the bounds
  return(pdf)
}
```

Fig-14: Dam-1 Plot:


Triangular Distribution for Cost-Benefit Analysis for Aplha-1
Density
0.00
0.02
0.04
0.06
0.08
0.10
5
10
15
Value
PDF
Mean
Minimum
Maximum
Mode

```
> # Generating x values within the range of the triangular distribution
> x <- seq(lower_bound, upper_bound, length.out = 100)
>
> # PDF of the triangular distribution
> pdf_triangular <- dtriangular(x, lower = lower_bound, mode = mode_value, upper = upper_bound)
>
> # Check if there are any valid PDF values to plot
> #if (all(is.nan(pdf_triangular))) {
>   #stop("No valid PDF values to plot.")
> #}
>
> # Determine y-axis limits for plotting
> y_min <- min(pdf_triangular, na.rm = TRUE)
> y_max <- max(pdf_triangular, na.rm = TRUE)
>
> # Plot PDF of triangular distribution with specified y-axis limits
> plot(x, pdf_triangular, type = "l", col = "blue", lwd = 2,
+      ylim = c(y_min, y_max),
+      main = "Triangular Distribution for Cost-Benefit Analysis for Aplha-1",
+      xlab = "Value", ylab = "Density")
>
> # Add vertical lines for mean, min, max, and mode values
> abline(v = mean_alpha1, col = "red", lty = 2, lwd = 2)  # Mean
> abline(v = min_alpha1, col = "green", lty = 2, lwd = 2)  # Minimum
> abline(v = max_alpha1, col = "orange", lty = 2, lwd = 2)  # Maximum
> abline(v = mode_value, col = "purple", lty = 2, lwd = 2)  # Mode
>
> # Add legend
> legend("topright", legend = c("PDF", "Mean", "Minimum", "Maximum", "Mode"),
+        col = c("blue", "red", "green", "orange", "purple"), lty = c(1, 2, 2, 2, 2), lwd = 2)
```

Fig-15: Dam-1- Chi-Square

```
> # Dam 1.
> # H0: The use of a triangular distribution adequately represents the uncertainty in the costs and bene
fits associated with cost-benefit analysis analysis.
>
> # H1: The use of a triangular distribution does not adequately represents the uncertainty in the costs
and benefits associated with cost-benefit analysis analysis.
>
> alpha1_observed <- hist(alpha1[[1]], breaks = 40, plot = FALSE)$counts
> alpha1_expected <- dnorm(hist(alpha1[[1]], breaks = 20, plot = FALSE)$mids,
+                          mean = mean(alpha1[[1]]), sd = sd(alpha1[[1]])) * length(alpha1[[1]])
>
> chisq_test_alpha1 <- chisq.test(alpha1_observed, p = alpha1_expected /sum(alpha1_expected ))
Warning message:
In chisq.test(alpha1_observed, p = alpha1_expected/sum(alpha1_expected)) :
  Chi-squared approximation may be incorrect
> chisq_test_alpha1

	Chi-squared test for given probabilities

data:  alpha1_observed
X-squared = 209.34, df = 26, p-value < 0.00000000000000022

>
> if(chisq_test_alpha1$p.value<=0.05){
+   cat("p-value is less than 0.05. Therefore, we reject the null hypothesis ")
+ } else {
+   cat("p-value is greater than 0.05. Therefore, we fail to reject the null hypothesis ")
+ }
p-value is less than 0.05. Therefore, we reject the null hypothesis
```

Fig-16: Dam-2 Plot:

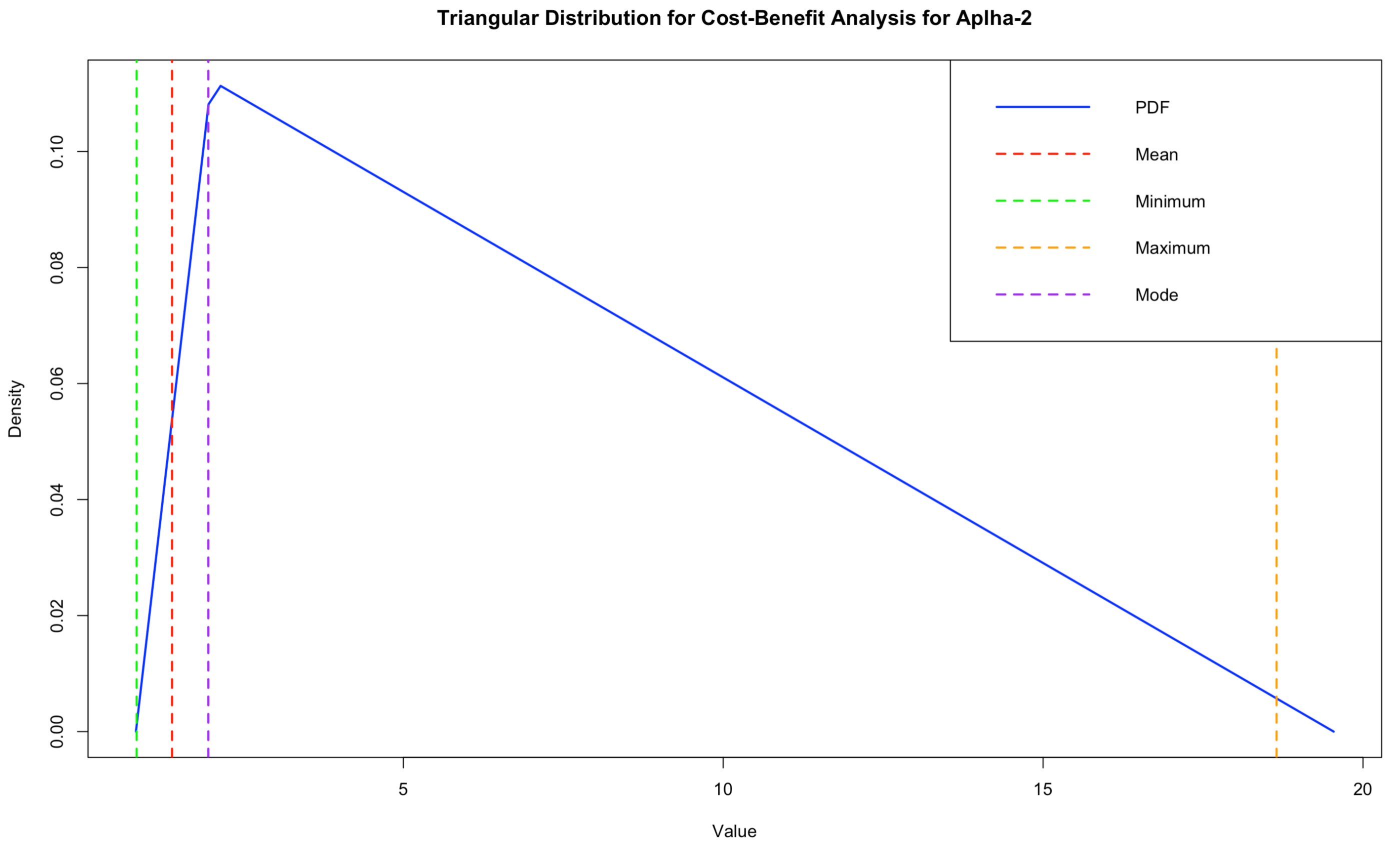


Fig-17: Dam-2- Chi-Square

```
> alpha2_observed <- hist(alpha2[[1]], breaks = 20, plot = FALSE)$counts
> alpha2_expected <- dnorm(hist(alpha2[[1]], breaks = 20, plot = FALSE)$mids,
+                     mean = mean(alpha2[[1]]), sd = sd(alpha2[[1]])) * length(alpha2[[1]])
>
> chisq_test_alpha2  <- chisq.test(alpha2_observed, p = alpha2_expected/sum(alpha2_expected))
Warning message:
In chisq.test(alpha2_observed, p = alpha2_expected/sum(alpha2_expected)) :
  Chi-squared approximation may be incorrect
> chisq_test_alpha2

        Chi-squared test for given probabilities

data:  alpha2_observed
X-squared = 295.72, df = 27, p-value < 0.00000000000000022

>
> if(chisq_test_alpha2$p.value<=0.05){
+   cat("p-value is less than 0.05. Therefore, we reject the null hypothesis ")
+ } else {
+   cat("p-value is greater than 0.05. Therefore, we fail to reject the null hypothesis ")
+ }
p-value is less than 0.05. Therefore, we reject the null hypothesis
```

Fig-18: Results Table

	Dam	Alpha1	Alpha2
1	Minimum	0.8310000	0.821000
2	Maximum	2.1100000	2.165000
3	Mean	1.3810000	1.371000
4	Median	1.3700000	1.355000
5	Variance	0.0410000	0.046000
6	Standard Deviation	0.2020000	0.214000
7	skewness	0.3000446	0.373557
8	P(alpha>2)	0.0017000	0.003600
9	P(alpha>1.8)	0.0263000	0.033800
10	P(alpha>1.5)	0.2725000	0.263700
11	P(alpha>1.2)	0.8038000	0.771700
12	P(alpha>1)	0.9829000	0.974600

```
> #install.packages("moments")
> library(moments)
> skew1 <- skewness(alpha1[[1]])
> p1_alpha1_2 <- mean(alpha1[[1]]>2)
> p1_alpha1_1.8 <- mean(alpha1[[1]]>1.8)
> p1_alpha1_1.5 <- mean(alpha1[[1]]>1.5)
> p1_alpha1_1.2 <- mean(alpha1[[1]]>1.2)
> p1_alpha1_1 <- mean(alpha1[[1]]>1)
> #For Dam 2
> skew2 <- skewness(alpha2[[1]])
> p2_alpha2_2 <- mean(alpha2[[1]]>2)
> p2_alpha2_1.8 <- mean(alpha2[[1]]>1.8)
> p2_alpha2_1.5 <- mean(alpha2[[1]]>1.5)
> p2_alpha2_1.2 <- mean(alpha2[[1]]>1.2)
> p2_alpha2_1 <- mean(alpha2[[1]]>1)
> # Dataframe for Displaying results.
> project_results <- data.frame(Dam = c("Minimum", "Maximum","Mean","Median",
+                                       "Variance","Standard Deviation","skewness",
+                                       "P(alpha>2)","P(alpha>1.8)","P(alpha>1.5)",
+                                       "P(alpha>1.2)","P(alpha>1)"),
+                          Alpha1 = c(round(min(alpha1[[1]]),3), round(max(alpha1[[1]]),3),
+                                     round(mean(alpha1[[1]]),3),round(median(alpha1[[1]]),3),
+                                     round(var(alpha1[[1]]),3), round(sd(alpha1[[1]]),3),
+                                     skew1,p1_alpha1_2,p1_alpha1_1.8,
+                                     p1_alpha1_1.5,p1_alpha1_1.2,p1_alpha1_1),
+                          Alpha2 = c(round(min(alpha2[[1]]),3), round(max(alpha2[[1]]),3),
+                                     round(mean(alpha2[[1]]),3),round(median(alpha2[[1]]),3),
+                                     round(var(alpha2[[1]]),3), round(sd(alpha2[[1]]),3),
+                                     skew2,p2_alpha2_2,p2_alpha2_1.8,
+                                     p2_alpha2_1.5,p2_alpha2_1.2,p2_alpha2_1))
```

Fig-19: Probability that the C-B ratio of Dam 1 will be higher

```
> # Probability that the C-B ratio of Dam 1 will be higher
>
> x <- sum(alpha1[[1]] > alpha2[[1]])
> prob <- x / 10000
> cat("Estimate that Alpha 1 > Alpha 2 ", prob)
Estimate that Alpha 1 > Alpha 2  0.6002>
> cat("Alpha1 consistently outperforms Alpha2 across various metrics.")
Alpha1 consistently outperforms Alpha2 across various metrics.
```

References:

ALY 6050: Module 2 Lab 2 — The Triangular Probability Distributions & Random Number Generation

JIM (22 September 2022): Triangular Distribution in R; [online]: Triangular Distribution in R | R-bloggers

Steven P. P. Sanderson: Understanding the Triangular Distribution and Its Application in R; [online]: Understanding the Triangular Distribution and Its Application in R | R-bloggers

R Document: The Triangular Distribution; [online]: R: The Triangular Distribution (r-project.org)

Kausthub Balaji: Triangular Distribution; [online]: RPubs - Triangular Distribution

Zach (3 December 2021): How to Use the Triangular Distribution in R; [online]: How to Use the Triangular Distribution in R (With Examples) - Statology