Yuan Jingtian

Data Analysis

**Task 2**

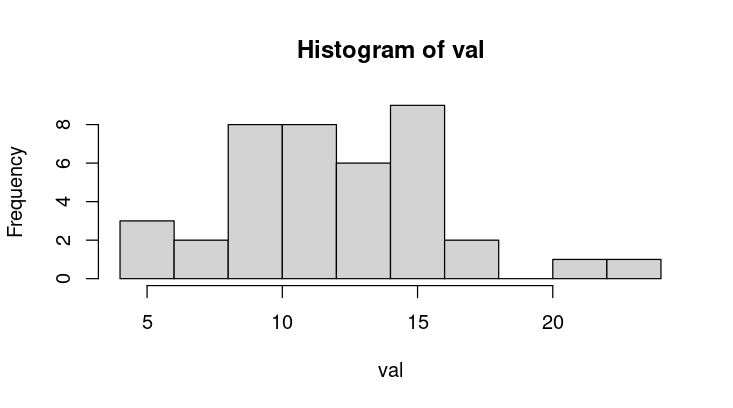
**Statistical Hypothesis**

**Dataset “data\_for\_analysis”**

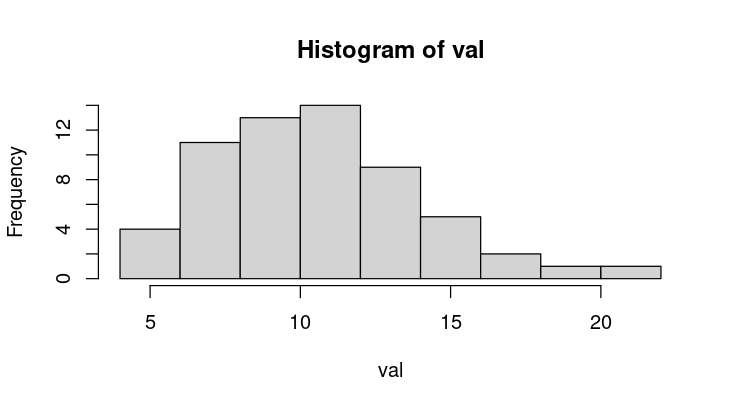
1. **estimate the distribution of continuous variables in data by group (outcome)**



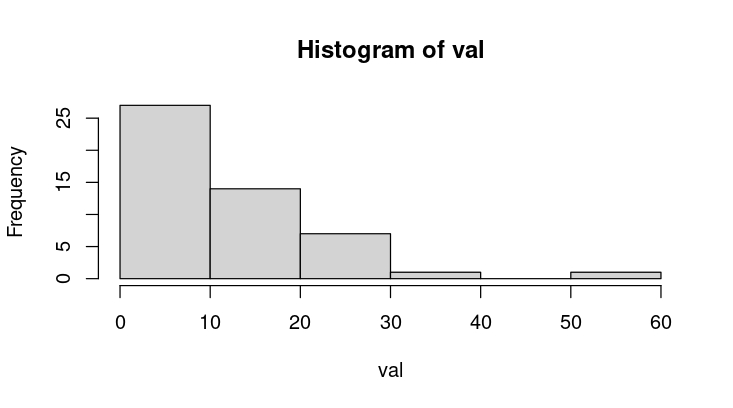
normal distribution



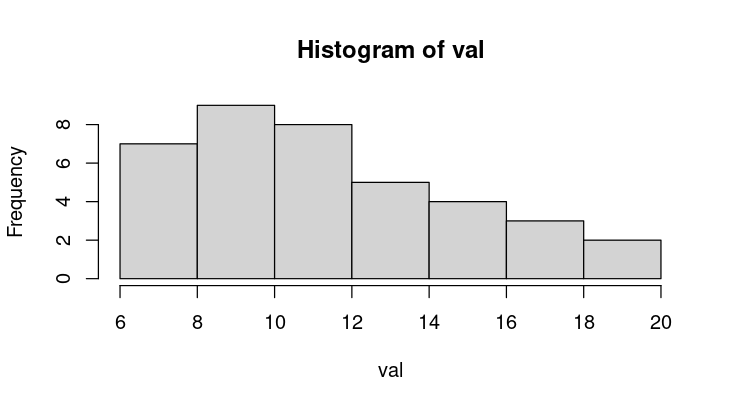
lognormal distribution



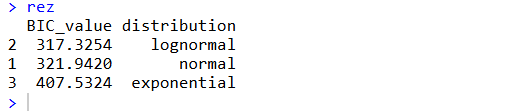
exponential distribution

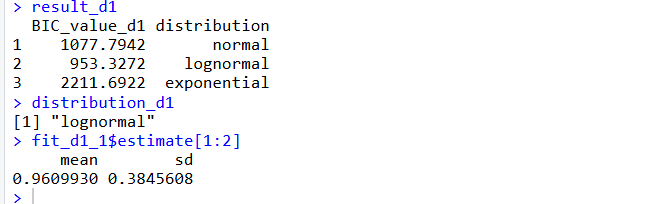


Poisson distribution



1. **create a table with descriptive statistics and specify the parameters according to the selected distribution (by group)**





#---------------Creating a custom table--------------

# Homework: Creating a custom table with descriptive statistics results

calculate\_stats <- function(x) {

if (is.numeric(x)) {

data.frame(

N = length(x),

Mean = mean(x, na.rm = TRUE),

SD = sd(x, na.rm = TRUE),

Median = median(x, na.rm = TRUE),

IQR = IQR(x, na.rm = TRUE),

Min = min(x, na.rm = TRUE),

Max = max(x, na.rm = TRUE),

Missing = sum(is.na(x))

)

} else {

data.frame(

N = length(x),

Categories = length(unique(x)),

Most\_Frequent = names(sort(table(x), decreasing = TRUE))[1],

Freq = max(table(x)),

Missing = sum(is.na(x))

)

}

}

# Select the variables we want to include in our table

vars\_to\_analyze <- c("lipids1", "lipids2", "lipids3", "lipids4")

# Apply the function to each variable and combine results

stats\_list <- lapply(data\_for\_analysis[vars\_to\_analyze], calculate\_stats)

# Combine all results into one table

custom\_stats\_table <- do.call(rbind, stats\_list)

# Add variable names as a column

custom\_stats\_table <- cbind(Variable = rownames(custom\_stats\_table), custom\_stats\_table)

rownames(custom\_stats\_table) <- NULL

# Format the table nicely

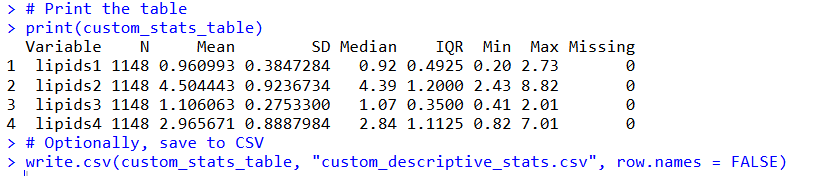
custom\_stats\_table <- as.data.frame(custom\_stats\_table)

# Print the table

print(custom\_stats\_table)

# Optionally, save to CSV

write.csv(custom\_stats\_table, "custom\_descriptive\_stats.csv", row.names = FALSE)



1. **perform the Brunner-Munzel test for 2 independent groups (enter the p-value into the data description table)**

