Jaesuk Lee

Stat 390

**Hw1:**lect2-1

**Continuous Variables:**

Exact Temperature

Exact Weight

**Discrete Variables:**

Date

Age

**Categorical Variable:**

Color

Race

**1.11b, 11.2 on paper**

**HW2:**1.16

idt <- c(28.1,31.2,13.7,46.0,25.8,16.8,34.8,62.3,28.0,17.9,19.5,21.1,31.9,28.9,60.1,23.7,18.6,21.4,26.6,26.2,32.0,43.5,17.4,38.8,30.6,55.6,25.5,52.1,21.0,22.3,15.5,36.3,19.1,38.4,72.8,48.9,21.4,20.7,57.3,40.9)

idtBins <- seq(from=10, to=80, by=10)

hist(idt, breaks=idtBins)

idtLog <- log10(idt)

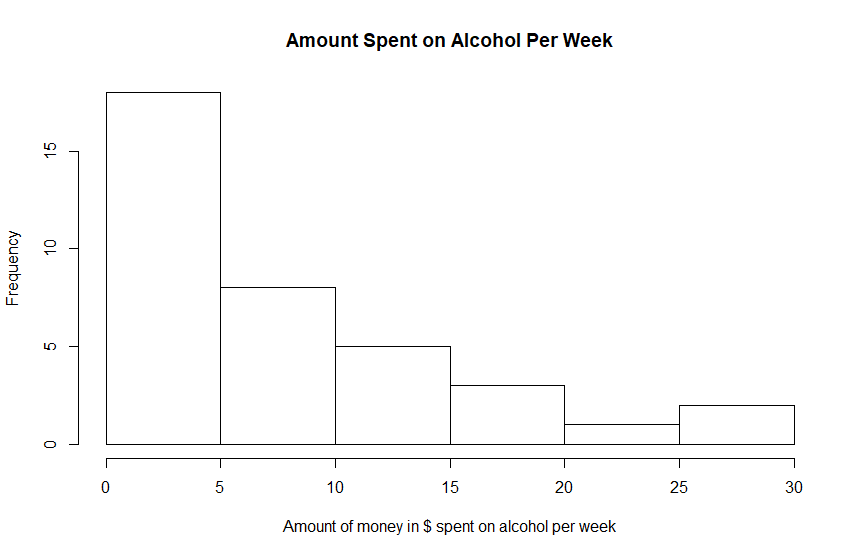
idtLogBin <- seq(from=1.1, to=1.9, by=0.1)

hist(idtLog, breaks=idtLogsBin)

Originally the histogram’s shape looks skewed, but after the transformation the shape looks normal or bell shaped

**Hw2**-lect2-2

See attached spreadsheet for values A screenshot of a cell phone

Description automatically generatedA screenshot of a cell phone

Description automatically generatedA screenshot of a cell phone

Description automatically generated

AlcoholSurvey <- read.csv("C:/Users/jaesu/Downloads/Survey On Alcohol Consumption (Responses) - Form Responses 1.csv")

hist(AlcoholSurvey$Approximately.how.much.do.you.spend.on.alcohol.per.week., main = "Amount Spent on Alcohol Per Week", xlab = "Amount of money in $ spent on alcohol per week")

hist(AlcoholSurvey$How.much.did.you.spend.on.your.last.drink.at.a.bar.or.club., xlab = "Cost of Last Drink at Bar", main= xlab + "Histogram" )

plot(as.factor(AlcoholSurvey$How.many.times.do.you.drink.alcohol.per.week.), xlab="# Of Times Alcohol is Had Per Week", ylab="frequency", main="# Of Times Alcohol is Had Per Week Histogram")

plot(as.factor(AlcoholSurvey$What.is.your.preferred.type.of.alcohol.), xlab="Preferred Alcohol Type", ylab="frequency", main="Prefrred Alcohol Type Histogram")

**HW3**-lect3-1

1. IQ. I’m guessing IQ follows the normal distribution because most people are average, there are very few people with very high or very low IQ. Therefore, the shape of the histogram would resemble a bell curve.
2. Age of death in the US. I think age of death would be skewed because most people in the US die at a later age rather than at an earlier or middle age. And very few people dying at an earlier age.
3. Transistor density of a CPU. I imagine that there are a ton of CPU’s that have low transistor count (embedded machines) that drastically decrease when going higher and higher densities.
4. Age of circumcision. I hope this is okay. I only say this because I’m Korean many Korean men get circumcised when they’re older. I think most people get circumcised when they are very young <1 year old, and many people get circumcised when they’re in their teens.

**HW3**-lect3-3

dataset <- read.table("~/hw\_lect3\_dat.txt", quote="\"", comment.char="")

v1 <- dataset$V1

v2 <- dataset$V2

v1Hist = hist(v1)

v2Hist = hist(v2)

plot(v1Hist$mids, log(v1Hist$density), type="b")

A screenshot of a cell phone

Description automatically generated

plot(log(v1Hist$mids), log(v1Hist$density), type="b")A screenshot of a cell phone

Description automatically generated

plot(v2Hist$mids, log(v2Hist$density), type="b")

A close up of a person

Description automatically generated

plot(log(v2Hist$mids), log(v2Hist$density), type="b")

A close up of a person

Description automatically generated

To determine if a histogram follows power law or exponential shape, the log of frequency, or the log of frequency and log of the variable needs to taken. If the histogram that results from the transformation looks linear, then it follows that law. Here is the proof for the transformation.

Looking at the plot, V1 seems to follow power law. When taken the log(freq) and log(variable), the resulting plot looks linear.

Looking at the plot, V2 seems to follow exponent. When taken the log(freq) the resulting plot looks linear.