Spearman–Brown Split-Half Reliability

Armen Bagdasarov

# Note: Analyses With outliers identified and removed

# Load packages  
library(splithalfr)  
library(MASS)

# Turn off scientific notation  
options(scipen = 999)

# 1 Minute

# Read data  
data <- read.csv("60s.csv")

# Determine outliers  
  
# Create subset of data   
data\_subset <- subset(data, select = c(even\_60s\_D\_gev, odd\_60s\_D\_gev))  
  
# Minimum Covariance Determinant  
output75 <- cov.mcd(data\_subset, quantile.used = nrow(data\_subset)\*.75)  
mhmcd75 <- mahalanobis(data\_subset, output75$center, output75$cov)  
alpha <- .001  
cutoff <- (qchisq(p = 1 - alpha, df = ncol(data\_subset)))  
names\_outlier\_MCD75 <- which(mhmcd75 > cutoff)  
data\_reduced <- data[-c(names\_outlier\_MCD75), ] # Remove outliers

# GEV  
# With Outliers Removed  
spearman\_brown(data\_reduced$even\_60s\_D\_gev, data\_reduced$odd\_60s\_D\_gev)

## [1] 0.8618193

48-nrow(data\_reduced)

## [1] 1

# Determine outliers  
  
# Create subset of data   
data\_subset <- subset(data, select = c(even\_60s\_A\_duration, odd\_60s\_A\_duration))  
  
# Minimum Covariance Determinant  
output75 <- cov.mcd(data\_subset, quantile.used = nrow(data\_subset)\*.75)  
mhmcd75 <- mahalanobis(data\_subset, output75$center, output75$cov)  
alpha <- .001  
cutoff <- (qchisq(p = 1 - alpha, df = ncol(data\_subset)))  
names\_outlier\_MCD75 <- which(mhmcd75 > cutoff)  
data\_reduced <- data[-c(names\_outlier\_MCD75), ] # Remove outliers

# Duration  
# With Outliers Removed  
spearman\_brown(data\_reduced$even\_60s\_A\_duration, data\_reduced$odd\_60s\_A\_duration)

## [1] 0.4720425

48-nrow(data\_reduced)

## [1] 1

# Determine outliers  
  
# Create subset of data   
data\_subset <- subset(data, select = c(even\_60s\_B\_duration, odd\_60s\_B\_duration))  
  
# Minimum Covariance Determinant  
output75 <- cov.mcd(data\_subset, quantile.used = nrow(data\_subset)\*.75)  
mhmcd75 <- mahalanobis(data\_subset, output75$center, output75$cov)  
alpha <- .001  
cutoff <- (qchisq(p = 1 - alpha, df = ncol(data\_subset)))  
names\_outlier\_MCD75 <- which(mhmcd75 > cutoff)  
data\_reduced <- data[-c(names\_outlier\_MCD75), ] # Remove outliers

# Duration  
# With Outliers Removed  
spearman\_brown(data\_reduced$even\_60s\_B\_duration, data\_reduced$odd\_60s\_B\_duration)

## [1] 0.4685547

48-nrow(data\_reduced)

## [1] 3

# Determine outliers  
  
# Create subset of data   
data\_subset <- subset(data, select = c(even\_60s\_C\_duration, odd\_60s\_C\_duration))  
  
# Minimum Covariance Determinant  
output75 <- cov.mcd(data\_subset, quantile.used = nrow(data\_subset)\*.75)  
mhmcd75 <- mahalanobis(data\_subset, output75$center, output75$cov)  
alpha <- .001  
cutoff <- (qchisq(p = 1 - alpha, df = ncol(data\_subset)))  
names\_outlier\_MCD75 <- which(mhmcd75 > cutoff)  
data\_reduced <- data[-c(names\_outlier\_MCD75), ] # Remove outliers

# Duration  
# With Outliers Removed  
spearman\_brown(data\_reduced$even\_60s\_C\_duration, data\_reduced$odd\_60s\_C\_duration)

## [1] 0.7242014

48-nrow(data\_reduced)

## [1] 1

# Determine outliers  
  
# Create subset of data   
data\_subset <- subset(data, select = c(even\_60s\_D\_duration, odd\_60s\_D\_duration))  
  
# Minimum Covariance Determinant  
output75 <- cov.mcd(data\_subset, quantile.used = nrow(data\_subset)\*.75)  
mhmcd75 <- mahalanobis(data\_subset, output75$center, output75$cov)  
alpha <- .001  
cutoff <- (qchisq(p = 1 - alpha, df = ncol(data\_subset)))  
names\_outlier\_MCD75 <- which(mhmcd75 > cutoff)  
data\_reduced <- data[-c(names\_outlier\_MCD75), ] # Remove outliers

# Duration  
# With Outliers Removed  
spearman\_brown(data\_reduced$even\_60s\_D\_duration, data\_reduced$odd\_60s\_D\_duration)

## [1] 0.77406

48-nrow(data\_reduced)

## [1] 2

# Determine outliers  
  
# Create subset of data   
data\_subset <- subset(data, select = c(even\_60s\_F\_duration, odd\_60s\_F\_duration))  
  
# Minimum Covariance Determinant  
output75 <- cov.mcd(data\_subset, quantile.used = nrow(data\_subset)\*.75)  
mhmcd75 <- mahalanobis(data\_subset, output75$center, output75$cov)  
alpha <- .001  
cutoff <- (qchisq(p = 1 - alpha, df = ncol(data\_subset)))  
names\_outlier\_MCD75 <- which(mhmcd75 > cutoff)  
data\_reduced <- data[-c(names\_outlier\_MCD75), ] # Remove outliers

# Duration  
# With Outliers Removed  
spearman\_brown(data\_reduced$even\_60s\_F\_duration, data\_reduced$odd\_60s\_F\_duration)

## [1] 0.6187202

48-nrow(data\_reduced)

## [1] 3

# Determine outliers  
  
# Create subset of data   
data\_subset <- subset(data, select = c(even\_60s\_A\_coverage, odd\_60s\_A\_coverage))  
  
# Minimum Covariance Determinant  
output75 <- cov.mcd(data\_subset, quantile.used = nrow(data\_subset)\*.75)  
mhmcd75 <- mahalanobis(data\_subset, output75$center, output75$cov)  
alpha <- .001  
cutoff <- (qchisq(p = 1 - alpha, df = ncol(data\_subset)))  
names\_outlier\_MCD75 <- which(mhmcd75 > cutoff)  
data\_reduced <- data[-c(names\_outlier\_MCD75), ] # Remove outliers

# Coverage  
# With Outliers Removed  
spearman\_brown(data\_reduced$even\_60s\_A\_coverage, data\_reduced$odd\_60s\_A\_coverage)

## [1] 0.7657565

48-nrow(data\_reduced)

## [1] 1

# Determine outliers  
  
# Create subset of data   
data\_subset <- subset(data, select = c(even\_60s\_C\_coverage, odd\_60s\_C\_coverage))  
  
# Minimum Covariance Determinant  
output75 <- cov.mcd(data\_subset, quantile.used = nrow(data\_subset)\*.75)  
mhmcd75 <- mahalanobis(data\_subset, output75$center, output75$cov)  
alpha <- .001  
cutoff <- (qchisq(p = 1 - alpha, df = ncol(data\_subset)))  
names\_outlier\_MCD75 <- which(mhmcd75 > cutoff)  
data\_reduced <- data[-c(names\_outlier\_MCD75), ] # Remove outliers

# Coverage  
# With Outliers Removed  
spearman\_brown(data\_reduced$even\_60s\_C\_coverage, data\_reduced$odd\_60s\_C\_coverage)

## [1] 0.8738947

48-nrow(data\_reduced)

## [1] 1

# Determine outliers  
  
# Create subset of data   
data\_subset <- subset(data, select = c(even\_60s\_C\_occurrence, odd\_60s\_C\_occurrence))  
  
# Minimum Covariance Determinant  
output75 <- cov.mcd(data\_subset, quantile.used = nrow(data\_subset)\*.75)  
mhmcd75 <- mahalanobis(data\_subset, output75$center, output75$cov)  
alpha <- .001  
cutoff <- (qchisq(p = 1 - alpha, df = ncol(data\_subset)))  
names\_outlier\_MCD75 <- which(mhmcd75 > cutoff)  
data\_reduced <- data[-c(names\_outlier\_MCD75), ] # Remove outliers

# Occurrence  
# With Outliers Removed  
spearman\_brown(data\_reduced$even\_60s\_C\_occurrence, data\_reduced$odd\_60s\_C\_occurrence)

## [1] 0.8085326

48-nrow(data\_reduced)

## [1] 1

# Determine outliers  
  
# Create subset of data   
data\_subset <- subset(data, select = c(even\_60s\_D\_occurrence, odd\_60s\_D\_occurrence))  
  
# Minimum Covariance Determinant  
output75 <- cov.mcd(data\_subset, quantile.used = nrow(data\_subset)\*.75)  
mhmcd75 <- mahalanobis(data\_subset, output75$center, output75$cov)  
alpha <- .001  
cutoff <- (qchisq(p = 1 - alpha, df = ncol(data\_subset)))  
names\_outlier\_MCD75 <- which(mhmcd75 > cutoff)  
data\_reduced <- data[-c(names\_outlier\_MCD75), ] # Remove outliers

# Occurrence  
# With Outliers Removed  
spearman\_brown(data\_reduced$even\_60s\_D\_occurrence, data\_reduced$odd\_60s\_D\_occurrence)

## [1] 0.6715264

48-nrow(data\_reduced)

## [1] 1

# 2 Minute

# Read data  
data <- read.csv("120s.csv")

# Determine outliers  
  
# Create subset of data   
data\_subset <- subset(data, select = c(even\_120s\_A\_gev, odd\_120s\_A\_gev))  
  
# Minimum Covariance Determinant  
output75 <- cov.mcd(data\_subset, quantile.used = nrow(data\_subset)\*.75)  
mhmcd75 <- mahalanobis(data\_subset, output75$center, output75$cov)  
alpha <- .001  
cutoff <- (qchisq(p = 1 - alpha, df = ncol(data\_subset)))  
names\_outlier\_MCD75 <- which(mhmcd75 > cutoff)  
data\_reduced <- data[-c(names\_outlier\_MCD75), ] # Remove outliers

# GEV  
# With Outliers Removed  
spearman\_brown(data\_reduced$even\_120s\_A\_gev, data\_reduced$odd\_120s\_A\_gev)

## [1] 0.7785856

48-nrow(data\_reduced)

## [1] 2

# Determine outliers  
  
# Create subset of data   
data\_subset <- subset(data, select = c(even\_120s\_B\_gev, odd\_120s\_B\_gev))  
  
# Minimum Covariance Determinant  
output75 <- cov.mcd(data\_subset, quantile.used = nrow(data\_subset)\*.75)  
mhmcd75 <- mahalanobis(data\_subset, output75$center, output75$cov)  
alpha <- .001  
cutoff <- (qchisq(p = 1 - alpha, df = ncol(data\_subset)))  
names\_outlier\_MCD75 <- which(mhmcd75 > cutoff)  
data\_reduced <- data[-c(names\_outlier\_MCD75), ] # Remove outliers

# GEV  
# With Outliers Removed  
spearman\_brown(data\_reduced$even\_120s\_B\_gev, data\_reduced$odd\_120s\_B\_gev)

## [1] 0.9207496

48-nrow(data\_reduced)

## [1] 2

# Determine outliers  
  
# Create subset of data   
data\_subset <- subset(data, select = c(even\_120s\_C\_gev, odd\_120s\_C\_gev))  
  
# Minimum Covariance Determinant  
output75 <- cov.mcd(data\_subset, quantile.used = nrow(data\_subset)\*.75)  
mhmcd75 <- mahalanobis(data\_subset, output75$center, output75$cov)  
alpha <- .001  
cutoff <- (qchisq(p = 1 - alpha, df = ncol(data\_subset)))  
names\_outlier\_MCD75 <- which(mhmcd75 > cutoff)  
data\_reduced <- data[-c(names\_outlier\_MCD75), ] # Remove outliers

# GEV  
# With Outliers Removed  
spearman\_brown(data\_reduced$even\_120s\_C\_gev, data\_reduced$odd\_120s\_C\_gev)

## [1] 0.8746808

48-nrow(data\_reduced)

## [1] 1

# Determine outliers  
  
# Create subset of data   
data\_subset <- subset(data, select = c(even\_120s\_D\_gev, odd\_120s\_D\_gev))  
  
# Minimum Covariance Determinant  
output75 <- cov.mcd(data\_subset, quantile.used = nrow(data\_subset)\*.75)  
mhmcd75 <- mahalanobis(data\_subset, output75$center, output75$cov)  
alpha <- .001  
cutoff <- (qchisq(p = 1 - alpha, df = ncol(data\_subset)))  
names\_outlier\_MCD75 <- which(mhmcd75 > cutoff)  
data\_reduced <- data[-c(names\_outlier\_MCD75), ] # Remove outliers

# GEV  
# With Outliers Removed  
spearman\_brown(data\_reduced$even\_120s\_D\_gev, data\_reduced$odd\_120s\_D\_gev)

## [1] 0.9323797

48-nrow(data\_reduced)

## [1] 6

# Determine outliers  
  
# Create subset of data   
data\_subset <- subset(data, select = c(even\_120s\_F\_gev, odd\_120s\_F\_gev))  
  
# Minimum Covariance Determinant  
output75 <- cov.mcd(data\_subset, quantile.used = nrow(data\_subset)\*.75)  
mhmcd75 <- mahalanobis(data\_subset, output75$center, output75$cov)  
alpha <- .001  
cutoff <- (qchisq(p = 1 - alpha, df = ncol(data\_subset)))  
names\_outlier\_MCD75 <- which(mhmcd75 > cutoff)  
data\_reduced <- data[-c(names\_outlier\_MCD75), ] # Remove outliers

# GEV  
# With Outliers Removed  
spearman\_brown(data\_reduced$even\_120s\_F\_gev, data\_reduced$odd\_120s\_F\_gev)

## [1] 0.6512246

48-nrow(data\_reduced)

## [1] 1

# Determine outliers  
  
# Create subset of data   
data\_subset <- subset(data, select = c(even\_120s\_B\_duration, odd\_120s\_B\_duration))  
  
# Minimum Covariance Determinant  
output75 <- cov.mcd(data\_subset, quantile.used = nrow(data\_subset)\*.75)  
mhmcd75 <- mahalanobis(data\_subset, output75$center, output75$cov)  
alpha <- .001  
cutoff <- (qchisq(p = 1 - alpha, df = ncol(data\_subset)))  
names\_outlier\_MCD75 <- which(mhmcd75 > cutoff)  
data\_reduced <- data[-c(names\_outlier\_MCD75), ] # Remove outliers

# Duration  
# With Outliers Removed  
spearman\_brown(data\_reduced$even\_120s\_B\_duration, data\_reduced$odd\_120s\_B\_duration)

## [1] 0.5795808

48-nrow(data\_reduced)

## [1] 1

# Determine outliers  
  
# Create subset of data   
data\_subset <- subset(data, select = c(even\_120s\_F\_duration, odd\_120s\_F\_duration))  
  
# Minimum Covariance Determinant  
output75 <- cov.mcd(data\_subset, quantile.used = nrow(data\_subset)\*.75)  
mhmcd75 <- mahalanobis(data\_subset, output75$center, output75$cov)  
alpha <- .001  
cutoff <- (qchisq(p = 1 - alpha, df = ncol(data\_subset)))  
names\_outlier\_MCD75 <- which(mhmcd75 > cutoff)  
data\_reduced <- data[-c(names\_outlier\_MCD75), ] # Remove outliers

# Duration  
# With Outliers Removed  
spearman\_brown(data\_reduced$even\_120s\_F\_duration, data\_reduced$odd\_120s\_F\_duration)

## [1] 0.5226496

48-nrow(data\_reduced)

## [1] 1

# Determine outliers  
  
# Create subset of data   
data\_subset <- subset(data, select = c(even\_120s\_A\_coverage, odd\_120s\_A\_coverage))  
  
# Minimum Covariance Determinant  
output75 <- cov.mcd(data\_subset, quantile.used = nrow(data\_subset)\*.75)  
mhmcd75 <- mahalanobis(data\_subset, output75$center, output75$cov)  
alpha <- .001  
cutoff <- (qchisq(p = 1 - alpha, df = ncol(data\_subset)))  
names\_outlier\_MCD75 <- which(mhmcd75 > cutoff)  
data\_reduced <- data[-c(names\_outlier\_MCD75), ] # Remove outliers

# Coverage  
# With Outliers Removed  
spearman\_brown(data\_reduced$even\_120s\_A\_coverage, data\_reduced$odd\_120s\_A\_coverage)

## [1] 0.7095196

48-nrow(data\_reduced)

## [1] 2

# Determine outliers  
  
# Create subset of data   
data\_subset <- subset(data, select = c(even\_120s\_C\_coverage, odd\_120s\_C\_coverage))  
  
# Minimum Covariance Determinant  
output75 <- cov.mcd(data\_subset, quantile.used = nrow(data\_subset)\*.75)  
mhmcd75 <- mahalanobis(data\_subset, output75$center, output75$cov)  
alpha <- .001  
cutoff <- (qchisq(p = 1 - alpha, df = ncol(data\_subset)))  
names\_outlier\_MCD75 <- which(mhmcd75 > cutoff)  
data\_reduced <- data[-c(names\_outlier\_MCD75), ] # Remove outliers

# Coverage  
# With Outliers Removed  
spearman\_brown(data\_reduced$even\_120s\_C\_coverage, data\_reduced$odd\_120s\_C\_coverage)

## [1] 0.832319

48-nrow(data\_reduced)

## [1] 1

# Determine outliers  
  
# Create subset of data   
data\_subset <- subset(data, select = c(even\_120s\_D\_coverage, odd\_120s\_D\_coverage))  
  
# Minimum Covariance Determinant  
output75 <- cov.mcd(data\_subset, quantile.used = nrow(data\_subset)\*.75)  
mhmcd75 <- mahalanobis(data\_subset, output75$center, output75$cov)  
alpha <- .001  
cutoff <- (qchisq(p = 1 - alpha, df = ncol(data\_subset)))  
names\_outlier\_MCD75 <- which(mhmcd75 > cutoff)  
data\_reduced <- data[-c(names\_outlier\_MCD75), ] # Remove outliers

# Coverage  
# With Outliers Removed  
spearman\_brown(data\_reduced$even\_120s\_D\_coverage, data\_reduced$odd\_120s\_D\_coverage)

## [1] 0.9392859

48-nrow(data\_reduced)

## [1] 1

# Determine outliers  
  
# Create subset of data   
data\_subset <- subset(data, select = c(even\_120s\_F\_coverage, odd\_120s\_F\_coverage))  
  
# Minimum Covariance Determinant  
output75 <- cov.mcd(data\_subset, quantile.used = nrow(data\_subset)\*.75)  
mhmcd75 <- mahalanobis(data\_subset, output75$center, output75$cov)  
alpha <- .001  
cutoff <- (qchisq(p = 1 - alpha, df = ncol(data\_subset)))  
names\_outlier\_MCD75 <- which(mhmcd75 > cutoff)  
data\_reduced <- data[-c(names\_outlier\_MCD75), ] # Remove outliers

# Coverage  
# With Outliers Removed  
spearman\_brown(data\_reduced$even\_120s\_F\_coverage, data\_reduced$odd\_120s\_F\_coverage)

## [1] 0.6227449

48-nrow(data\_reduced)

## [1] 3

# Determine outliers  
  
# Create subset of data   
data\_subset <- subset(data, select = c(even\_120s\_D\_occurrence, odd\_120s\_D\_occurrence))  
  
# Minimum Covariance Determinant  
output75 <- cov.mcd(data\_subset, quantile.used = nrow(data\_subset)\*.75)  
mhmcd75 <- mahalanobis(data\_subset, output75$center, output75$cov)  
alpha <- .001  
cutoff <- (qchisq(p = 1 - alpha, df = ncol(data\_subset)))  
names\_outlier\_MCD75 <- which(mhmcd75 > cutoff)  
data\_reduced <- data[-c(names\_outlier\_MCD75), ] # Remove outliers

# Occurrence  
# With Outliers Removed  
spearman\_brown(data\_reduced$even\_120s\_D\_occurrence, data\_reduced$odd\_120s\_D\_occurrence)

## [1] 0.9116121

48-nrow(data\_reduced)

## [1] 1

# Determine outliers  
  
# Create subset of data   
data\_subset <- subset(data, select = c(even\_120s\_F\_occurrence, odd\_120s\_F\_occurrence))  
  
# Minimum Covariance Determinant  
output75 <- cov.mcd(data\_subset, quantile.used = nrow(data\_subset)\*.75)  
mhmcd75 <- mahalanobis(data\_subset, output75$center, output75$cov)  
alpha <- .001  
cutoff <- (qchisq(p = 1 - alpha, df = ncol(data\_subset)))  
names\_outlier\_MCD75 <- which(mhmcd75 > cutoff)  
data\_reduced <- data[-c(names\_outlier\_MCD75), ] # Remove outliers

# Occurrence  
# With Outliers Removed  
spearman\_brown(data\_reduced$even\_120s\_F\_occurrence, data\_reduced$odd\_120s\_F\_occurrence)

## [1] 0.8180833

48-nrow(data\_reduced)

## [1] 2

# 3 Minute

# Read data  
data <- read.csv("180s.csv")

# Determine outliers  
  
# Create subset of data   
data\_subset <- subset(data, select = c(even\_180s\_A\_gev, odd\_180s\_A\_gev))  
  
# Minimum Covariance Determinant  
output75 <- cov.mcd(data\_subset, quantile.used = nrow(data\_subset)\*.75)  
mhmcd75 <- mahalanobis(data\_subset, output75$center, output75$cov)  
alpha <- .001  
cutoff <- (qchisq(p = 1 - alpha, df = ncol(data\_subset)))  
names\_outlier\_MCD75 <- which(mhmcd75 > cutoff)  
data\_reduced <- data[-c(names\_outlier\_MCD75), ] # Remove outliers

# GEV  
# With Outliers Removed  
spearman\_brown(data\_reduced$even\_180s\_A\_gev, data\_reduced$odd\_180s\_A\_gev)

## [1] 0.9287306

48-nrow(data\_reduced)

## [1] 1

# Determine outliers  
  
# Create subset of data   
data\_subset <- subset(data, select = c(even\_180s\_B\_gev, odd\_180s\_B\_gev))  
  
# Minimum Covariance Determinant  
output75 <- cov.mcd(data\_subset, quantile.used = nrow(data\_subset)\*.75)  
mhmcd75 <- mahalanobis(data\_subset, output75$center, output75$cov)  
alpha <- .001  
cutoff <- (qchisq(p = 1 - alpha, df = ncol(data\_subset)))  
names\_outlier\_MCD75 <- which(mhmcd75 > cutoff)  
data\_reduced <- data[-c(names\_outlier\_MCD75), ] # Remove outliers

# GEV  
# With Outliers Removed  
spearman\_brown(data\_reduced$even\_180s\_B\_gev, data\_reduced$odd\_180s\_B\_gev)

## [1] 0.8953596

48-nrow(data\_reduced)

## [1] 3

# Determine outliers  
  
# Create subset of data   
data\_subset <- subset(data, select = c(even\_180s\_D\_gev, odd\_180s\_D\_gev))  
  
# Minimum Covariance Determinant  
output75 <- cov.mcd(data\_subset, quantile.used = nrow(data\_subset)\*.75)  
mhmcd75 <- mahalanobis(data\_subset, output75$center, output75$cov)  
alpha <- .001  
cutoff <- (qchisq(p = 1 - alpha, df = ncol(data\_subset)))  
names\_outlier\_MCD75 <- which(mhmcd75 > cutoff)  
data\_reduced <- data[-c(names\_outlier\_MCD75), ] # Remove outliers

# GEV  
# With Outliers Removed  
spearman\_brown(data\_reduced$even\_180s\_D\_gev, data\_reduced$odd\_180s\_D\_gev)

## [1] 0.9493536

48-nrow(data\_reduced)

## [1] 1

# Determine outliers  
  
# Create subset of data   
data\_subset <- subset(data, select = c(even\_180s\_F\_gev, odd\_180s\_F\_gev))  
  
# Minimum Covariance Determinant  
output75 <- cov.mcd(data\_subset, quantile.used = nrow(data\_subset)\*.75)  
mhmcd75 <- mahalanobis(data\_subset, output75$center, output75$cov)  
alpha <- .001  
cutoff <- (qchisq(p = 1 - alpha, df = ncol(data\_subset)))  
names\_outlier\_MCD75 <- which(mhmcd75 > cutoff)  
data\_reduced <- data[-c(names\_outlier\_MCD75), ] # Remove outliers

# GEV  
# With Outliers Removed  
spearman\_brown(data\_reduced$even\_180s\_F\_gev, data\_reduced$odd\_180s\_F\_gev)

## [1] 0.8170254

48-nrow(data\_reduced)

## [1] 2

# Determine outliers  
  
# Create subset of data   
data\_subset <- subset(data, select = c(even\_180s\_A\_duration, odd\_180s\_A\_duration))  
  
# Minimum Covariance Determinant  
output75 <- cov.mcd(data\_subset, quantile.used = nrow(data\_subset)\*.75)  
mhmcd75 <- mahalanobis(data\_subset, output75$center, output75$cov)  
alpha <- .001  
cutoff <- (qchisq(p = 1 - alpha, df = ncol(data\_subset)))  
names\_outlier\_MCD75 <- which(mhmcd75 > cutoff)  
data\_reduced <- data[-c(names\_outlier\_MCD75), ] # Remove outliers

# Duration  
# With Outliers Removed  
spearman\_brown(data\_reduced$even\_180s\_A\_duration, data\_reduced$odd\_180s\_A\_duration)

## [1] 0.7273031

48-nrow(data\_reduced)

## [1] 1

# Determine outliers  
  
# Create subset of data   
data\_subset <- subset(data, select = c(even\_180s\_B\_duration, odd\_180s\_B\_duration))  
  
# Minimum Covariance Determinant  
output75 <- cov.mcd(data\_subset, quantile.used = nrow(data\_subset)\*.75)  
mhmcd75 <- mahalanobis(data\_subset, output75$center, output75$cov)  
alpha <- .001  
cutoff <- (qchisq(p = 1 - alpha, df = ncol(data\_subset)))  
names\_outlier\_MCD75 <- which(mhmcd75 > cutoff)  
data\_reduced <- data[-c(names\_outlier\_MCD75), ] # Remove outliers

# Duration  
# With Outliers Removed  
spearman\_brown(data\_reduced$even\_180s\_B\_duration, data\_reduced$odd\_180s\_B\_duration)

## [1] 0.5787027

48-nrow(data\_reduced)

## [1] 2

# Determine outliers  
  
# Create subset of data   
data\_subset <- subset(data, select = c(even\_180s\_B\_coverage, odd\_180s\_B\_coverage))  
  
# Minimum Covariance Determinant  
output75 <- cov.mcd(data\_subset, quantile.used = nrow(data\_subset)\*.75)  
mhmcd75 <- mahalanobis(data\_subset, output75$center, output75$cov)  
alpha <- .001  
cutoff <- (qchisq(p = 1 - alpha, df = ncol(data\_subset)))  
names\_outlier\_MCD75 <- which(mhmcd75 > cutoff)  
data\_reduced <- data[-c(names\_outlier\_MCD75), ] # Remove outliers

# Coverage  
# With Outliers Removed  
spearman\_brown(data\_reduced$even\_180s\_B\_coverage, data\_reduced$odd\_180s\_B\_coverage)

## [1] NA

48-nrow(data\_reduced)

## [1] 48

# Determine outliers  
  
# Create subset of data   
data\_subset <- subset(data, select = c(even\_180s\_F\_coverage, odd\_180s\_F\_coverage))  
  
# Minimum Covariance Determinant  
output75 <- cov.mcd(data\_subset, quantile.used = nrow(data\_subset)\*.75)  
mhmcd75 <- mahalanobis(data\_subset, output75$center, output75$cov)  
alpha <- .001  
cutoff <- (qchisq(p = 1 - alpha, df = ncol(data\_subset)))  
names\_outlier\_MCD75 <- which(mhmcd75 > cutoff)  
data\_reduced <- data[-c(names\_outlier\_MCD75), ] # Remove outliers

# Coverage  
# With Outliers Removed  
spearman\_brown(data\_reduced$even\_180s\_F\_coverage, data\_reduced$odd\_180s\_F\_coverage)

## [1] 0.7578533

48-nrow(data\_reduced)

## [1] 5

# Determine outliers  
  
# Create subset of data   
data\_subset <- subset(data, select = c(even\_180s\_D\_occurrence, odd\_180s\_D\_occurrence))  
  
# Minimum Covariance Determinant  
output75 <- cov.mcd(data\_subset, quantile.used = nrow(data\_subset)\*.75)  
mhmcd75 <- mahalanobis(data\_subset, output75$center, output75$cov)  
alpha <- .001  
cutoff <- (qchisq(p = 1 - alpha, df = ncol(data\_subset)))  
names\_outlier\_MCD75 <- which(mhmcd75 > cutoff)  
data\_reduced <- data[-c(names\_outlier\_MCD75), ] # Remove outliers

# Occurrence  
# With Outliers Removed  
spearman\_brown(data\_reduced$even\_180s\_D\_occurrence, data\_reduced$odd\_180s\_D\_occurrence)

## [1] 0.8690466

48-nrow(data\_reduced)

## [1] 1

# Determine outliers  
  
# Create subset of data   
data\_subset <- subset(data, select = c(even\_180s\_F\_occurrence, odd\_180s\_F\_occurrence))  
  
# Minimum Covariance Determinant  
output75 <- cov.mcd(data\_subset, quantile.used = nrow(data\_subset)\*.75)  
mhmcd75 <- mahalanobis(data\_subset, output75$center, output75$cov)  
alpha <- .001  
cutoff <- (qchisq(p = 1 - alpha, df = ncol(data\_subset)))  
names\_outlier\_MCD75 <- which(mhmcd75 > cutoff)  
data\_reduced <- data[-c(names\_outlier\_MCD75), ] # Remove outliers

# Occurrence  
# With Outliers Removed  
spearman\_brown(data\_reduced$even\_180s\_F\_occurrence, data\_reduced$odd\_180s\_F\_occurrence)

## [1] 0.7253558

48-nrow(data\_reduced)

## [1] 5

# 4 Minute

# Read data  
data <- read.csv("240s.csv")

# Determine outliers  
  
# Create subset of data   
data\_subset <- subset(data, select = c(even\_240s\_B\_gev, odd\_240s\_B\_gev))  
  
# Minimum Covariance Determinant  
output75 <- cov.mcd(data\_subset, quantile.used = nrow(data\_subset)\*.75)  
mhmcd75 <- mahalanobis(data\_subset, output75$center, output75$cov)  
alpha <- .001  
cutoff <- (qchisq(p = 1 - alpha, df = ncol(data\_subset)))  
names\_outlier\_MCD75 <- which(mhmcd75 > cutoff)  
data\_reduced <- data[-c(names\_outlier\_MCD75), ] # Remove outliers

# GEV  
# With Outliers Removed  
spearman\_brown(data\_reduced$even\_240s\_B\_gev, data\_reduced$odd\_240s\_B\_gev)

## [1] 0.962708

48-nrow(data\_reduced)

## [1] 7

# Determine outliers  
  
# Create subset of data   
data\_subset <- subset(data, select = c(even\_240s\_D\_gev, odd\_240s\_D\_gev))  
  
# Minimum Covariance Determinant  
output75 <- cov.mcd(data\_subset, quantile.used = nrow(data\_subset)\*.75)  
mhmcd75 <- mahalanobis(data\_subset, output75$center, output75$cov)  
alpha <- .001  
cutoff <- (qchisq(p = 1 - alpha, df = ncol(data\_subset)))  
names\_outlier\_MCD75 <- which(mhmcd75 > cutoff)  
data\_reduced <- data[-c(names\_outlier\_MCD75), ] # Remove outliers

# GEV  
# With Outliers Removed  
spearman\_brown(data\_reduced$even\_240s\_D\_gev, data\_reduced$odd\_240s\_D\_gev)

## [1] 0.9535721

48-nrow(data\_reduced)

## [1] 1

# Determine outliers  
  
# Create subset of data   
data\_subset <- subset(data, select = c(even\_240s\_F\_gev, odd\_240s\_F\_gev))  
  
# Minimum Covariance Determinant  
output75 <- cov.mcd(data\_subset, quantile.used = nrow(data\_subset)\*.75)  
mhmcd75 <- mahalanobis(data\_subset, output75$center, output75$cov)  
alpha <- .001  
cutoff <- (qchisq(p = 1 - alpha, df = ncol(data\_subset)))  
names\_outlier\_MCD75 <- which(mhmcd75 > cutoff)  
data\_reduced <- data[-c(names\_outlier\_MCD75), ] # Remove outliers

# GEV  
# With Outliers Removed  
spearman\_brown(data\_reduced$even\_240s\_F\_gev, data\_reduced$odd\_240s\_F\_gev)

## [1] 0.9354996

48-nrow(data\_reduced)

## [1] 1

# Determine outliers  
  
# Create subset of data   
data\_subset <- subset(data, select = c(even\_240s\_B\_duration, odd\_240s\_B\_duration))  
  
# Minimum Covariance Determinant  
output75 <- cov.mcd(data\_subset, quantile.used = nrow(data\_subset)\*.75)  
mhmcd75 <- mahalanobis(data\_subset, output75$center, output75$cov)  
alpha <- .001  
cutoff <- (qchisq(p = 1 - alpha, df = ncol(data\_subset)))  
names\_outlier\_MCD75 <- which(mhmcd75 > cutoff)  
data\_reduced <- data[-c(names\_outlier\_MCD75), ] # Remove outliers

# Duration  
# With Outliers Removed  
spearman\_brown(data\_reduced$even\_240s\_B\_duration, data\_reduced$odd\_240s\_B\_duration)

## [1] 0.5793405

48-nrow(data\_reduced)

## [1] 3

# Determine outliers  
  
# Create subset of data   
data\_subset <- subset(data, select = c(even\_240s\_B\_coverage, odd\_240s\_B\_coverage))  
  
# Minimum Covariance Determinant  
output75 <- cov.mcd(data\_subset, quantile.used = nrow(data\_subset)\*.75)  
mhmcd75 <- mahalanobis(data\_subset, output75$center, output75$cov)  
alpha <- .001  
cutoff <- (qchisq(p = 1 - alpha, df = ncol(data\_subset)))  
names\_outlier\_MCD75 <- which(mhmcd75 > cutoff)  
data\_reduced <- data[-c(names\_outlier\_MCD75), ] # Remove outliers

# Coverage  
# With Outliers Removed  
spearman\_brown(data\_reduced$even\_240s\_B\_coverage, data\_reduced$odd\_240s\_B\_coverage)

## [1] 0.9384993

48-nrow(data\_reduced)

## [1] 3

# Determine outliers  
  
# Create subset of data   
data\_subset <- subset(data, select = c(even\_240s\_F\_coverage, odd\_240s\_F\_coverage))  
  
# Minimum Covariance Determinant  
output75 <- cov.mcd(data\_subset, quantile.used = nrow(data\_subset)\*.75)  
mhmcd75 <- mahalanobis(data\_subset, output75$center, output75$cov)  
alpha <- .001  
cutoff <- (qchisq(p = 1 - alpha, df = ncol(data\_subset)))  
names\_outlier\_MCD75 <- which(mhmcd75 > cutoff)  
data\_reduced <- data[-c(names\_outlier\_MCD75), ] # Remove outliers

# Coverage  
# With Outliers Removed  
spearman\_brown(data\_reduced$even\_240s\_F\_coverage, data\_reduced$odd\_240s\_F\_coverage)

## [1] 0.9030808

48-nrow(data\_reduced)

## [1] 1

# Determine outliers  
  
# Create subset of data   
data\_subset <- subset(data, select = c(even\_240s\_B\_occurrence, odd\_240s\_B\_occurrence))  
  
# Minimum Covariance Determinant  
output75 <- cov.mcd(data\_subset, quantile.used = nrow(data\_subset)\*.75)  
mhmcd75 <- mahalanobis(data\_subset, output75$center, output75$cov)  
alpha <- .001  
cutoff <- (qchisq(p = 1 - alpha, df = ncol(data\_subset)))  
names\_outlier\_MCD75 <- which(mhmcd75 > cutoff)  
data\_reduced <- data[-c(names\_outlier\_MCD75), ] # Remove outliers

# Occurrence  
# With Outliers Removed  
spearman\_brown(data\_reduced$even\_240s\_B\_occurrence, data\_reduced$odd\_240s\_B\_occurrence)

## [1] 0.8980409

48-nrow(data\_reduced)

## [1] 1

# Determine outliers  
  
# Create subset of data   
data\_subset <- subset(data, select = c(even\_240s\_C\_occurrence, odd\_240s\_C\_occurrence))  
  
# Minimum Covariance Determinant  
output75 <- cov.mcd(data\_subset, quantile.used = nrow(data\_subset)\*.75)  
mhmcd75 <- mahalanobis(data\_subset, output75$center, output75$cov)  
alpha <- .001  
cutoff <- (qchisq(p = 1 - alpha, df = ncol(data\_subset)))  
names\_outlier\_MCD75 <- which(mhmcd75 > cutoff)  
data\_reduced <- data[-c(names\_outlier\_MCD75), ] # Remove outliers

# Occurrence  
# With Outliers Removed  
spearman\_brown(data\_reduced$even\_240s\_C\_occurrence, data\_reduced$odd\_240s\_C\_occurrence)

## [1] 0.9230371

48-nrow(data\_reduced)

## [1] 3

# Determine outliers  
  
# Create subset of data   
data\_subset <- subset(data, select = c(even\_240s\_D\_occurrence, odd\_240s\_D\_occurrence))  
  
# Minimum Covariance Determinant  
output75 <- cov.mcd(data\_subset, quantile.used = nrow(data\_subset)\*.75)  
mhmcd75 <- mahalanobis(data\_subset, output75$center, output75$cov)  
alpha <- .001  
cutoff <- (qchisq(p = 1 - alpha, df = ncol(data\_subset)))  
names\_outlier\_MCD75 <- which(mhmcd75 > cutoff)  
data\_reduced <- data[-c(names\_outlier\_MCD75), ] # Remove outliers

# Occurrence  
# With Outliers Removed  
spearman\_brown(data\_reduced$even\_240s\_D\_occurrence, data\_reduced$odd\_240s\_D\_occurrence)

## [1] 0.9031107

48-nrow(data\_reduced)

## [1] 1

# Determine outliers  
  
# Create subset of data   
data\_subset <- subset(data, select = c(even\_240s\_F\_occurrence, odd\_240s\_F\_occurrence))  
  
# Minimum Covariance Determinant  
output75 <- cov.mcd(data\_subset, quantile.used = nrow(data\_subset)\*.75)  
mhmcd75 <- mahalanobis(data\_subset, output75$center, output75$cov)  
alpha <- .001  
cutoff <- (qchisq(p = 1 - alpha, df = ncol(data\_subset)))  
names\_outlier\_MCD75 <- which(mhmcd75 > cutoff)  
data\_reduced <- data[-c(names\_outlier\_MCD75), ] # Remove outliers

# Occurrence  
# With Outliers Removed  
spearman\_brown(data\_reduced$even\_240s\_F\_occurrence, data\_reduced$odd\_240s\_F\_occurrence)

## [1] 0.801909

48-nrow(data\_reduced)

## [1] 3

# 5 Minute

# Read data\_reduced  
data <- read.csv("300s.csv")

# Determine outliers  
  
# Create subset of data   
data\_subset <- subset(data, select = c(even\_300s\_B\_gev, odd\_300s\_B\_gev))  
  
# Minimum Covariance Determinant  
output75 <- cov.mcd(data\_subset, quantile.used = nrow(data\_subset)\*.75)  
mhmcd75 <- mahalanobis(data\_subset, output75$center, output75$cov)  
alpha <- .001  
cutoff <- (qchisq(p = 1 - alpha, df = ncol(data\_subset)))  
names\_outlier\_MCD75 <- which(mhmcd75 > cutoff)  
data\_reduced <- data[-c(names\_outlier\_MCD75), ] # Remove outliers

# GEV  
# With Outliers Removed  
spearman\_brown(data\_reduced$even\_300s\_B\_gev, data\_reduced$odd\_300s\_B\_gev)

## [1] 0.9250248

48-nrow(data\_reduced)

## [1] 2

# Determine outliers  
  
# Create subset of data   
data\_subset <- subset(data, select = c(even\_300s\_F\_gev, odd\_300s\_F\_gev))  
  
# Minimum Covariance Determinant  
output75 <- cov.mcd(data\_subset, quantile.used = nrow(data\_subset)\*.75)  
mhmcd75 <- mahalanobis(data\_subset, output75$center, output75$cov)  
alpha <- .001  
cutoff <- (qchisq(p = 1 - alpha, df = ncol(data\_subset)))  
names\_outlier\_MCD75 <- which(mhmcd75 > cutoff)  
data\_reduced <- data[-c(names\_outlier\_MCD75), ] # Remove outliers

# GEV  
# With Outliers Removed  
spearman\_brown(data\_reduced$even\_300s\_F\_gev, data\_reduced$odd\_300s\_F\_gev)

## [1] 0.943714

48-nrow(data\_reduced)

## [1] 2

# Determine outliers  
  
# Create subset of data   
data\_subset <- subset(data, select = c(even\_300s\_B\_duration, odd\_300s\_B\_duration))  
  
# Minimum Covariance Determinant  
output75 <- cov.mcd(data\_subset, quantile.used = nrow(data\_subset)\*.75)  
mhmcd75 <- mahalanobis(data\_subset, output75$center, output75$cov)  
alpha <- .001  
cutoff <- (qchisq(p = 1 - alpha, df = ncol(data\_subset)))  
names\_outlier\_MCD75 <- which(mhmcd75 > cutoff)  
data\_reduced <- data[-c(names\_outlier\_MCD75), ] # Remove outliers

# Duration  
# With Outliers Removed  
spearman\_brown(data\_reduced$even\_300s\_B\_duration, data\_reduced$odd\_300s\_B\_duration)

## [1] 0.8192045

48-nrow(data\_reduced)

## [1] 3

# Determine outliers  
  
# Create subset of data   
data\_subset <- subset(data, select = c(even\_300s\_F\_duration, odd\_300s\_F\_duration))  
  
# Minimum Covariance Determinant  
output75 <- cov.mcd(data\_subset, quantile.used = nrow(data\_subset)\*.75)  
mhmcd75 <- mahalanobis(data\_subset, output75$center, output75$cov)  
alpha <- .001  
cutoff <- (qchisq(p = 1 - alpha, df = ncol(data\_subset)))  
names\_outlier\_MCD75 <- which(mhmcd75 > cutoff)  
data\_reduced <- data[-c(names\_outlier\_MCD75), ] # Remove outliers

# Duration  
# With Outliers Removed  
spearman\_brown(data\_reduced$even\_300s\_F\_duration, data\_reduced$odd\_300s\_F\_duration)

## [1] 0.8756041

48-nrow(data\_reduced)

## [1] 3

# Determine outliers  
  
# Create subset of data   
data\_subset <- subset(data, select = c(even\_300s\_B\_coverage, odd\_300s\_B\_coverage))  
  
# Minimum Covariance Determinant  
output75 <- cov.mcd(data\_subset, quantile.used = nrow(data\_subset)\*.75)  
mhmcd75 <- mahalanobis(data\_subset, output75$center, output75$cov)  
alpha <- .001  
cutoff <- (qchisq(p = 1 - alpha, df = ncol(data\_subset)))  
names\_outlier\_MCD75 <- which(mhmcd75 > cutoff)  
data\_reduced <- data[-c(names\_outlier\_MCD75), ] # Remove outliers

# Coverage  
# With Outliers Removed  
spearman\_brown(data\_reduced$even\_300s\_B\_coverage, data\_reduced$odd\_300s\_B\_coverage)

## [1] 0.9158674

48-nrow(data\_reduced)

## [1] 2

# Determine outliers  
  
# Create subset of data   
data\_subset <- subset(data, select = c(even\_300s\_F\_coverage, odd\_300s\_F\_coverage))  
  
# Minimum Covariance Determinant  
output75 <- cov.mcd(data\_subset, quantile.used = nrow(data\_subset)\*.75)  
mhmcd75 <- mahalanobis(data\_subset, output75$center, output75$cov)  
alpha <- .001  
cutoff <- (qchisq(p = 1 - alpha, df = ncol(data\_subset)))  
names\_outlier\_MCD75 <- which(mhmcd75 > cutoff)  
data\_reduced <- data[-c(names\_outlier\_MCD75), ] # Remove outliers

# Coverage  
# With Outliers Removed  
spearman\_brown(data\_reduced$even\_300s\_F\_coverage, data\_reduced$odd\_300s\_F\_coverage)

## [1] 0.9022011

48-nrow(data\_reduced)

## [1] 1

# Determine outliers  
  
# Create subset of data   
data\_subset <- subset(data, select = c(even\_300s\_B\_occurrence, odd\_300s\_B\_occurrence))  
  
# Minimum Covariance Determinant  
output75 <- cov.mcd(data\_subset, quantile.used = nrow(data\_subset)\*.75)  
mhmcd75 <- mahalanobis(data\_subset, output75$center, output75$cov)  
alpha <- .001  
cutoff <- (qchisq(p = 1 - alpha, df = ncol(data\_subset)))  
names\_outlier\_MCD75 <- which(mhmcd75 > cutoff)  
data\_reduced <- data[-c(names\_outlier\_MCD75), ] # Remove outliers

# Occurrence  
# With Outliers Removed  
spearman\_brown(data\_reduced$even\_300s\_B\_occurrence, data\_reduced$odd\_300s\_B\_occurrence)

## [1] 0.9130604

48-nrow(data\_reduced)

## [1] 1

# Determine outliers  
  
# Create subset of data   
data\_subset <- subset(data, select = c(even\_300s\_C\_occurrence, odd\_300s\_C\_occurrence))  
  
# Minimum Covariance Determinant  
output75 <- cov.mcd(data\_subset, quantile.used = nrow(data\_subset)\*.75)  
mhmcd75 <- mahalanobis(data\_subset, output75$center, output75$cov)  
alpha <- .001  
cutoff <- (qchisq(p = 1 - alpha, df = ncol(data\_subset)))  
names\_outlier\_MCD75 <- which(mhmcd75 > cutoff)  
data\_reduced <- data[-c(names\_outlier\_MCD75), ] # Remove outliers

# Occurrence  
# With Outliers Removed  
spearman\_brown(data\_reduced$even\_300s\_C\_occurrence, data\_reduced$odd\_300s\_C\_occurrence)

## [1] 0.8924716

48-nrow(data\_reduced)

## [1] 1

# Determine outliers  
  
# Create subset of data   
data\_subset <- subset(data, select = c(even\_300s\_D\_occurrence, odd\_300s\_D\_occurrence))  
  
# Minimum Covariance Determinant  
output75 <- cov.mcd(data\_subset, quantile.used = nrow(data\_subset)\*.75)  
mhmcd75 <- mahalanobis(data\_subset, output75$center, output75$cov)  
alpha <- .001  
cutoff <- (qchisq(p = 1 - alpha, df = ncol(data\_subset)))  
names\_outlier\_MCD75 <- which(mhmcd75 > cutoff)  
data\_reduced <- data[-c(names\_outlier\_MCD75), ] # Remove outliers

# Occurrence  
# With Outliers Removed  
spearman\_brown(data\_reduced$even\_300s\_D\_occurrence, data\_reduced$odd\_300s\_D\_occurrence)

## [1] 0.8744441

48-nrow(data\_reduced)

## [1] 3