Return Analysis

Mingze Li

2015-04-16

```r
# Load required libraries
library(traveltimeCLT)
library(data.table)
```

```
## Warning: package 'data.table' was built under R version 4.3.3
```

```r
trips <- fread("data/trips.csv")
id <- sample(unique(trips$trip), 1000)
train <- trips[!trips$trip %in% id, ]
test <- trips[trips$trip %in% id, ]
timeBin_x_edge <- get_timeBin_x_edges(train)
```
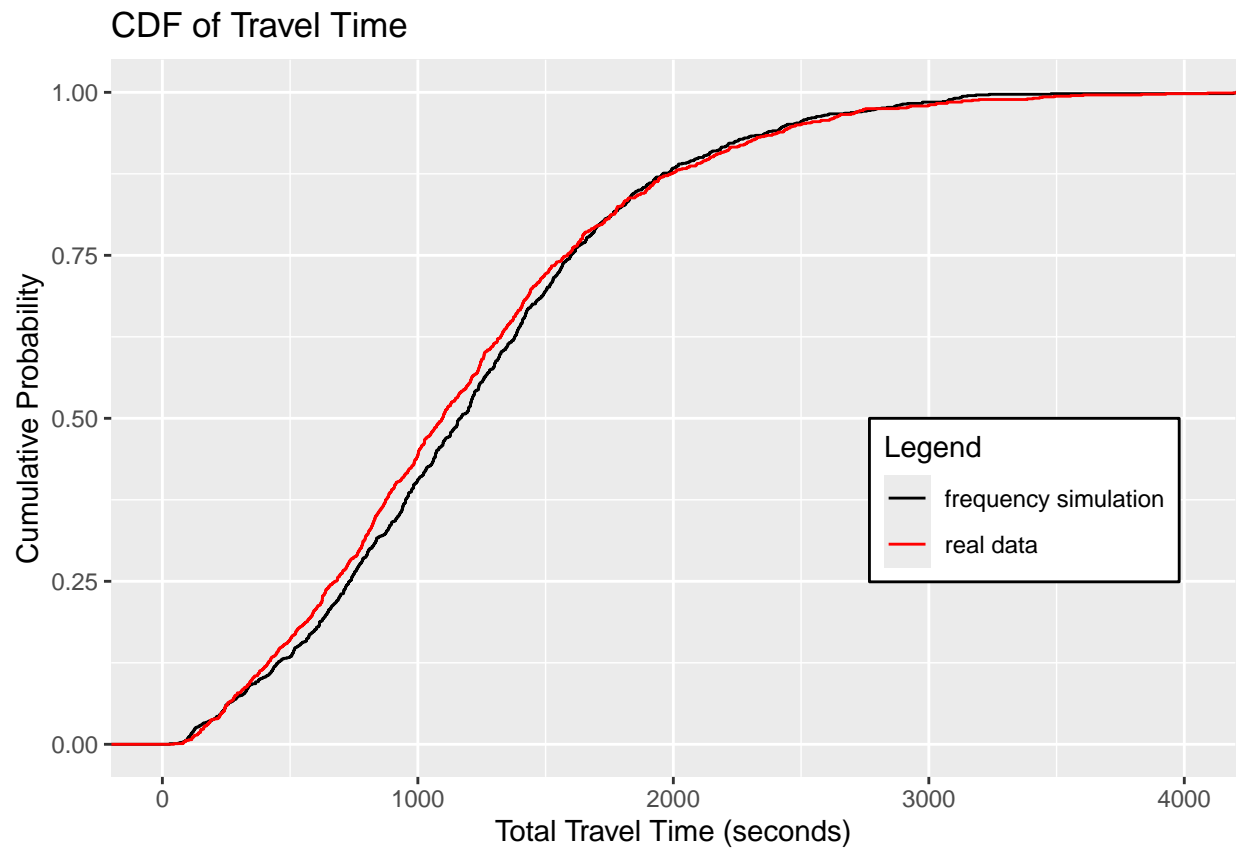
```r
# sample_route(id,trips,1001)
sample <- sample_route(id, trips, 1001, timeBin_x_edge)
pressure_test <- sample_route(id, trips, 1001, timeBin_x_edge, abuse_ratio = 0.3, abuse_multiplier = 1.
```

```r
plot_CDF_compare(sample[[2]]$real_time, sample[[1]]$dependent_time, "frequency simulation")
```

CDF of Travel Time


```
plot_CDF_compare(pressure_test[[2]]$real_time, pressure_test[[1]]$dependent_time, "frequency simulation"
```

CDF of Travel Time
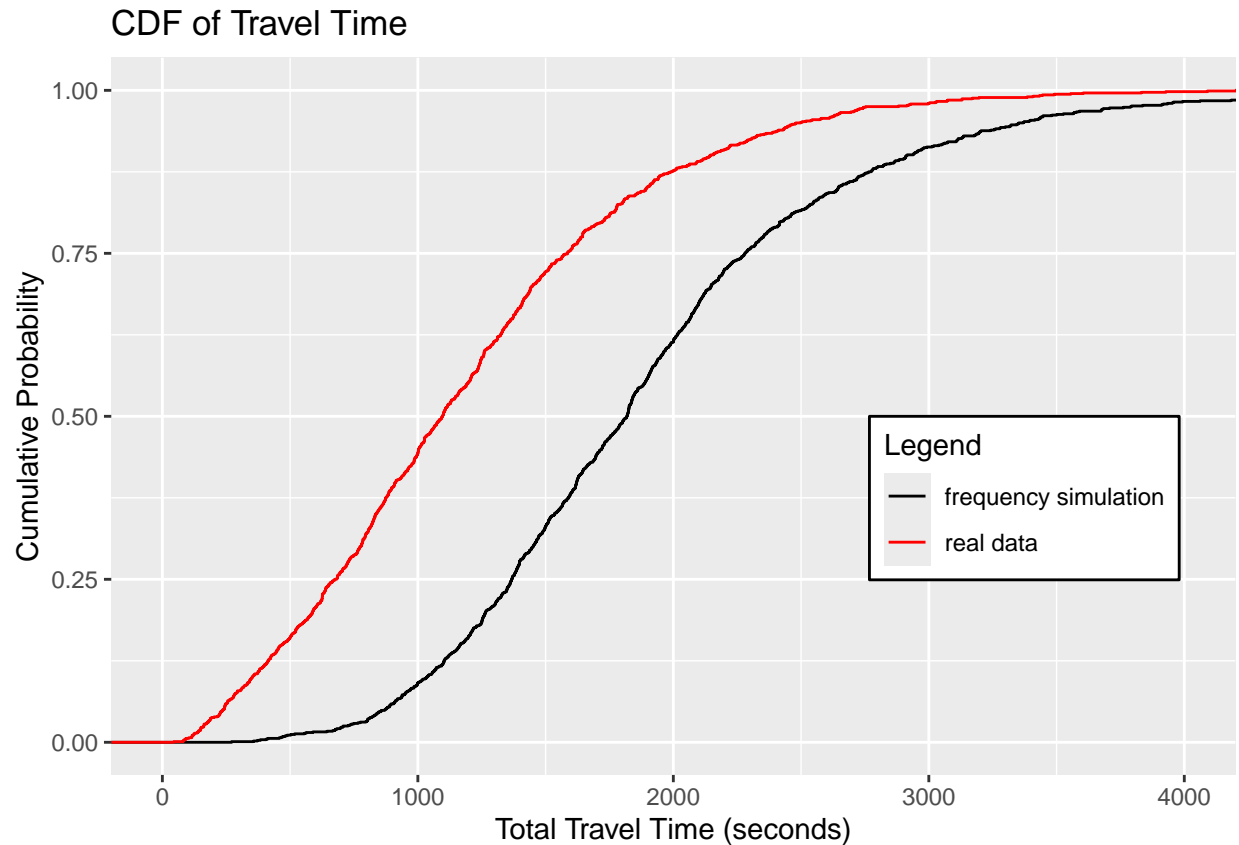

```r
real_price <- price(sample[[2]]$real_time, sample[[2]]$real_length)[, 1]
simulated_price <- data.table(
  simulated_price = price(sample[[1]]$dependent_time, sample[[1]]$simulated_length)[, 1],
  pressured_price = price(pressure_test[[1]]$dependent_time, pressure_test[[1]]$simulated_length)[, 1]
)
```

```r
names(train)[c(2, 3, 5, 7, 8)] <- c("tripID", "entry_time", "duration_secs", "distance_meters", "linkID")
train$speed <- exp(train$logspeed)
train$timeBin <- time_bins_readable(train$entry_time)
fit <- traveltimeCLT(train)
```

```
## Warning in traveltimeCLT(train): 4 trips have less than 1 observation, and will
## not be used to estimate autocorrelations, or residual variance parameters
```

```r
test <- trips[trips$trip %in% id, ]
names(test)[c(2, 3, 5, 7, 8)] <- c("tripID", "entry_time", "1", "distance_meters", "linkID")
names(sample[[3]]) <- c("tripID", "linkID", "entry_time", "distance_meters")
names(pressure_test[[3]]) <- c("tripID", "linkID", "entry_time", "distance_meters")
p <- predict(fit, test)
simulated_p <- predict(fit, sample[[3]])
pressure_p <- predict(fit, pressure_test[[3]])
```

```r
start_times <- test[, .(start_time = entry_time[1]), by = tripID]
simulated_start_time <- (sample[[3]][, .(start_time = entry_time[1]), by = tripID][, 2])
```

```r
simulated__start_time <- simulated_start_time[["start_time"]]
pressure_start_time <- (pressure_test[[3]][, .(start_time = entry_time[1]), by = tripID][, 2])
pressure_start_time <- pressure_start_time[["start_time"]]
R <- request_R(p, start_times$start_time, start_times$start_time, sample[[2]]$real_length, 1, risk_free
simulated_R <- request_R(simulated_p, simulated__start_time, simulated__start_time, sample[[1]]$simulate
pressure_R <- request_R(pressure_p, pressure_start_time, pressure_start_time, pressure_test[[1]]$simula

K <- request_K(p, start_times$start_time, start_times$start_time, sample[[2]]$real_length, risk_free = 
simulated_K <- request_K(simulated_p, simulated__start_time, simulated__start_time, sample[[1]]$simulate
pressure_K <- request_K(pressure_p, pressure_start_time, pressure_start_time, pressure_test[[1]]$simula
income <- simulated_R + simulated_K
normal_expand <- real_price
pressure_expand <- simulated_price$pressured_price
mean(income[1:1000] - normal_expand)
```
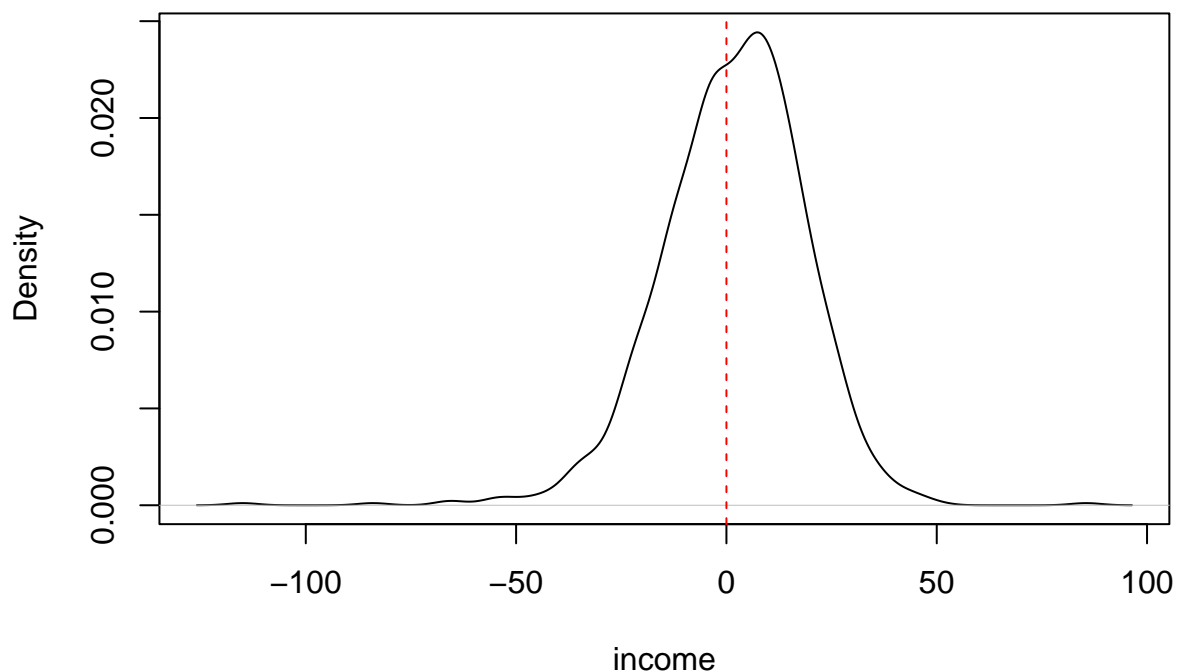
```
## [1] 1.670225
```

```r
mean(income - pressure_expand)
```

```
## [1] -11.10089
```

```r
plot(density(income[1:1000] - normal_expand), xlab = "income", main = "")
abline(v = 0, col = "red", lty = 2)
```

```r
plot(density(income - pressure_expand), xlab = "income", main = "")
abline(v = 0, col = "red", lty = 2)
```