STAT 600 - HW 4

Kevin Korsurat

All Rcpp/RcppArmadillo can be found in my GitHub.

Question 1

For this question, I will let t-distribution with 3 df as a density that we can sample from. The plots below shows the distribution of the target distribution $q(x)$, the distribution that we can sample from $g(x)$, and the envelope function based on $g(x)$ with the different value of α.
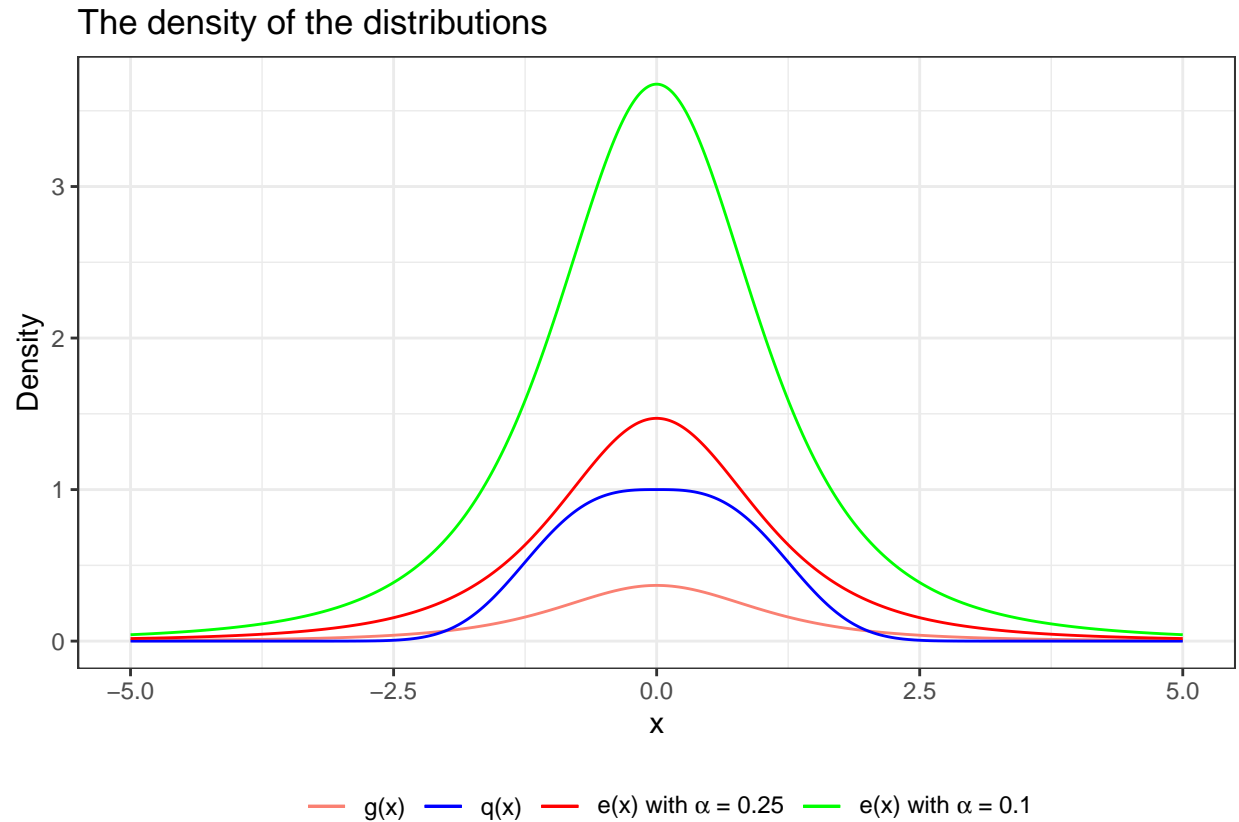

The density of the distributions

Figure 1: The density for all function used in the analysis.

Figure 2 shows that all of the proposed distribution have a heavier tail compared to the target distribution. Besides, as shown in the Figure 1, the envelope function for both α are greater than the target distribution for the all range of x.
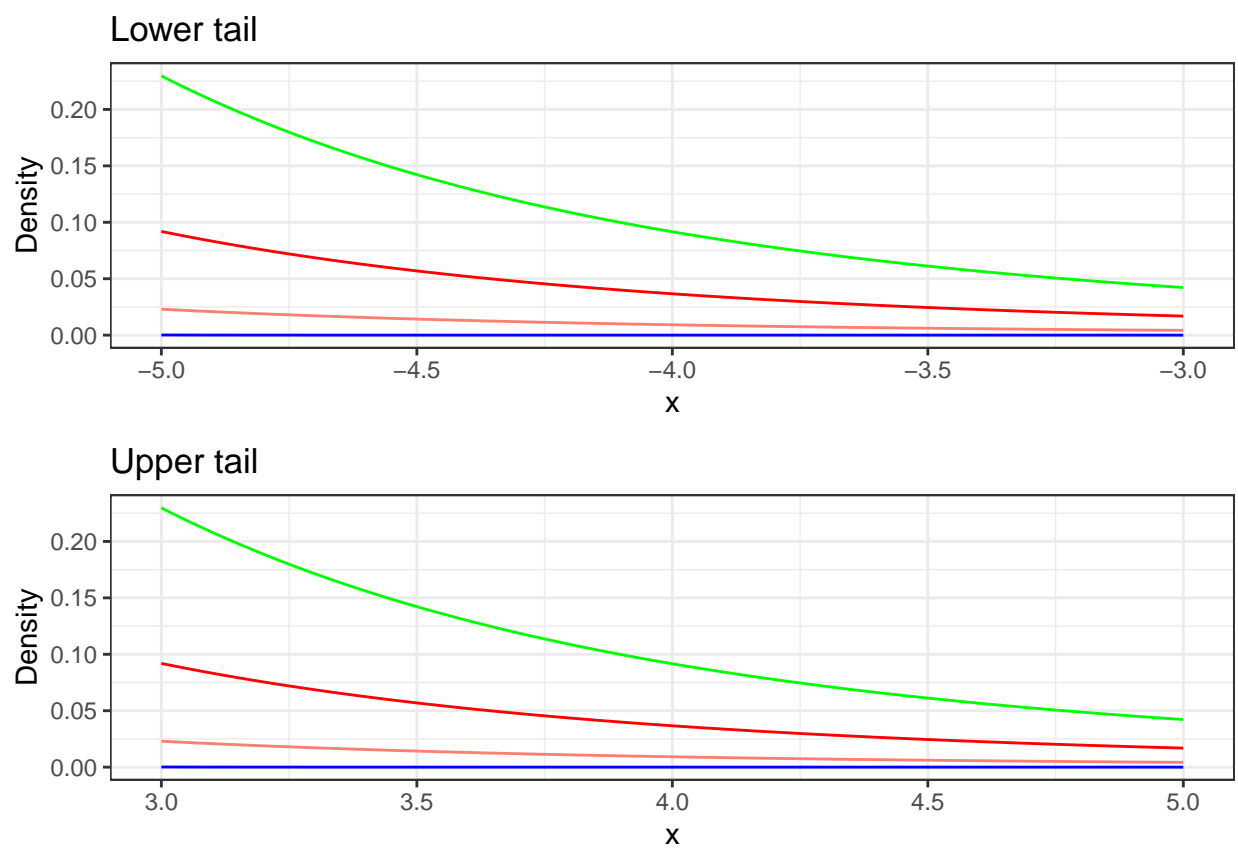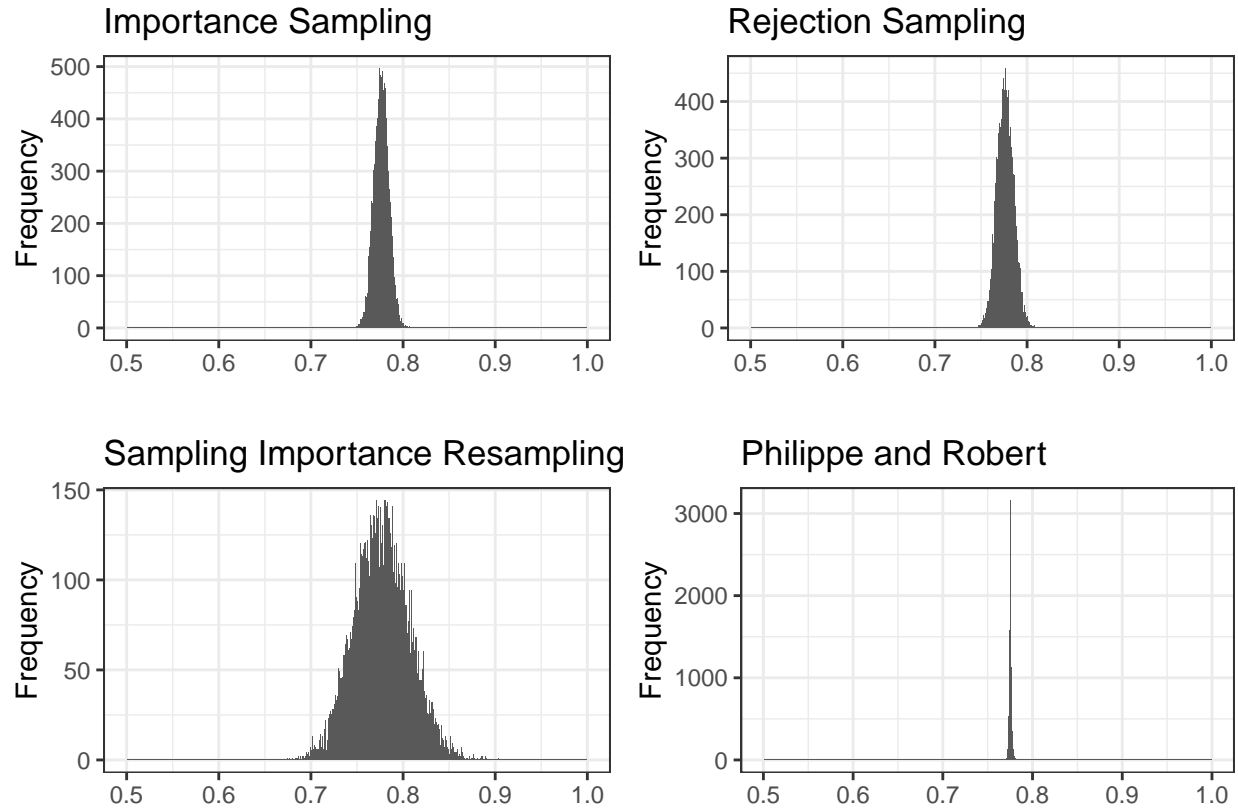
Figure 2: The tail density for all function used in the analysis.

Importance Sampling

Rejection Sampling

Sampling Importance Resampling

Philippe and Robert

	Estimates	Computational Time
Importance Sampling	0.77644 (SD = 0.00815)	0.00242 (SD = 0.00776)
Rejection Sampling	0.77635 (SD = 0.00919)	0.00381 (SD = 0.00294)
Sampling Importance Resampling	0.77705 (SD = 0.03015)	0.00237 (SD = 0.00086)
Philippe and Robert	0.77541 (SD = 0.00139)	0.00452 (SD = 0.00109)

Question 2

Appendix

```r
knitr::opts_chunk$set(echo = FALSE)

library(tidyverse)
library(knitr)
library(Rcpp)
library(RcppArmadillo)
library(devtools)
library(foreach)
library(doParallel)
library(ggplot2)
library(latex2exp)
library(gridExtra)

path <- "/Users/kevin-imac/Desktop/Github - Repo/"
if(! file.exists(path)){
  path <- "/Users/kevinkvp/Desktop/Github Repo/"
}

compileAttributes()
build()
install()
library(HW4MC)

### User-defined functions -------------------------------------------------
meanSD <- function(x, dplace = 5){
  mm <- round(mean(x), digits = dplace)
  ss <- round(sd(x), digits = dplace)
  paste0(mm, " (SD = ", ss, ")")
}

### Plots
x <- seq(-5, 5, 0.01)
targetD <- sapply(x, tg_dist, a = 3, b = 3)
sampD <- dt(x, 3)

data.frame(x, g = sampD, q = targetD) %>%
  ggplot(aes(x = x)) +
  geom_line(aes(y = sampD, color = "A1")) +
  geom_line(aes(y = targetD, color = "A2")) +
  geom_line(aes(y = sampD/0.25, color = "A3")) +
  geom_line(aes(y = sampD/0.1, color = "A4")) +
  theme_bw() +
  theme(legend.position = "bottom") +
  labs(x = "x", y = "Density", color = " ", title = "The density of the distributions") +
  scale_color_manual(values = c("salmon", "blue", "red", "green"),
                     labels = c("g(x)", "q(x)",
                                TeX("e(x) with $\\alpha$ = 0.25"),
                                TeX("e(x) with $\\alpha$ = 0.1")))

x <- seq(-5, -3, 0.01)
targetD <- sapply(x, tg_dist, a = 3, b = 3)
```

```r
sampD <- dt(x, 3)

p2 <- data.frame(x, g = sampD, q = targetD) %>%
  ggplot(aes(x = x)) +
  geom_line(aes(y = sampD, color = "A1")) +
  geom_line(aes(y = targetD, color = "A2")) +
  geom_line(aes(y = sampD/0.25, color = "A3")) +
  geom_line(aes(y = sampD/0.1, color = "A4")) +
  theme_bw() +
  theme(legend.position = "none") +
  labs(x = "x", y = "Density", color = " ", title = "Lower tail") +
  scale_color_manual(values = c("salmon", "blue", "red", "green"),
                     labels = c("Distribution that we can sample from, g(x)",
                                "Target Distribution, q(x)",
                                TeX("Envelope Function with $\\alpha$ = 0.25"),
                                TeX("Envelope Function with $\\alpha$ = 0.1")))

x <- seq(3, 5, 0.01)
targetD <- sapply(x, tg_dist, a = 3, b = 3)
sampD <- dt(x, 3)

p3 <- data.frame(x, g = sampD, q = targetD) %>%
  ggplot(aes(x = x)) +
  geom_line(aes(y = sampD, color = "A1")) +
  geom_line(aes(y = targetD, color = "A2")) +
  geom_line(aes(y = sampD/0.25, color = "A3")) +
  geom_line(aes(y = sampD/0.1, color = "A4")) +
  theme_bw() +
  theme(legend.position = "none") +
  labs(x = "x", y = "Density", color = " ", title = "Upper tail") +
  scale_color_manual(values = c("salmon", "blue", "red", "green"),
                     labels = c("Distribution that we can sample from, g(x)",
                                "Target Distribution, q(x)",
                                TeX("Envelope Function with $\\alpha$ = 0.25"),
                                TeX("Envelope Function with $\\alpha$ = 0.1")))

grid.arrange(grobs = list(p2, p3))

### Run the models
### Define the function of interests
h_fx <- function(x){x^2}

### (a): Importance Sampling
set.seed(12, kind = "L'Ecuyer-CMRG")
registerDoParallel(5)
impResult <- foreach(t = 1:10000, .combine = "rbind") %dopar% {

  startT <- Sys.time()
  estimate <- impSamp(h = h_fx, n = 10000, a = 3, b = 3, df = 3, showCalc = FALSE)
  totalT <- difftime(Sys.time(), startT)
  c(estimate, totalT)

}
```

```r
stopImplicitCluster()

### (b): Rejection Sampling (alpha = 0.25)
set.seed(12, kind = "L'Ecuyer-CMRG")
registerDoParallel(5)
rejResult <- foreach(t = 1:10000, .combine = "rbind") %dopar% {

  startT <- Sys.time()
  estimate <- rejSamp(h = h_fx, n = 10000, alp = 0.25, a = 3, b = 3, df = 3, showSamp = FALSE)
  totalT <- difftime(Sys.time(), startT)
  c(estimate, totalT)

}
stopImplicitCluster()

### (c): Sampling Importance Resampling
set.seed(12, kind = "L'Ecuyer-CMRG")
registerDoParallel(5)
sirResult <- foreach(t = 1:10000, .combine = "rbind") %dopar% {

  startT <- Sys.time()
  estimate <- SIRSamp(h = h_fx, n = 10000, M = 1000, a = 3, b = 3, df = 3)
  totalT <- difftime(Sys.time(), startT)
  c(estimate, totalT)

}
stopImplicitCluster()

### (d): Philippe and Robert
set.seed(12, kind = "L'Ecuyer-CMRG")
registerDoParallel(5)
phroResult <- foreach(t = 1:10000, .combine = "rbind") %dopar% {

  startT <- Sys.time()
  estimate <- PhRoMethod(h = h_fx, n = 10000, alp = 0.25, a = 3, b = 3, df = 3)
  totalT <- difftime(Sys.time(), startT)
  c(estimate, totalT)

}
stopImplicitCluster()

### Histogram
p1 <- ggplot(data.frame(x = impResult[, 1]), aes(x = x)) +
  geom_histogram(binwidth = 0.001) +
  xlim(0.5, 1) +
  labs(title = "Importance Sampling", x = " ", y = "Frequency") +
  theme_bw()

p2 <- ggplot(data.frame(x = rejResult[, 1]), aes(x = x)) +
  geom_histogram(binwidth = 0.001) +
  xlim(0.5, 1) +
  labs(title = "Rejection Sampling", x = " ", y = "Frequency") +
  theme_bw()
```

```r
p3 <- ggplot(data.frame(x = sirResult[, 1]), aes(x = x)) +
  geom_histogram(binwidth = 0.001) +
  labs(title = "Sampling Importance Resampling", x = " ", y = "Frequency") +
  xlim(0.5, 1) +
  theme_bw()

p4 <- ggplot(data.frame(x = phroResult[, 1]), aes(x = x)) +
  geom_histogram(binwidth = 0.001) +
  labs(title = "Philippe and Robert", x = " ", y = "Frequency") +
  xlim(0.5, 1) +
  theme_bw()

grid.arrange(grobs = list(p1, p2, p3, p4))

### Summary table
estResult <- c(meanSD(impResult[, 1]), meanSD(rejResult[, 1]), meanSD(sirResult[, 1]), meanSD(phroResul
timeResult <- c(meanSD(impResult[, 2]), meanSD(rejResult[, 2]), meanSD(sirResult[, 2]), meanSD(phroResul
data.frame(estResult, timeResult) %>%
  `rownames<-`(c("Importance Sampling", "Rejection Sampling",
              "Sampling Importance Resampling", "Philippe and Robert")) %>%
  kable(col.names = c("Estimates", "Computational Time"))
```