AMS 553 Project

Zhe Zhou

12/1/2022

Set Random Seed

```
library(ggplot2)
set.seed(123)
```

Example Function

```
g <- function(x) {
  exp(-x - log(1+x^2)) * (x > 0) * (x < 1)
}
```

1. Importance Sampling

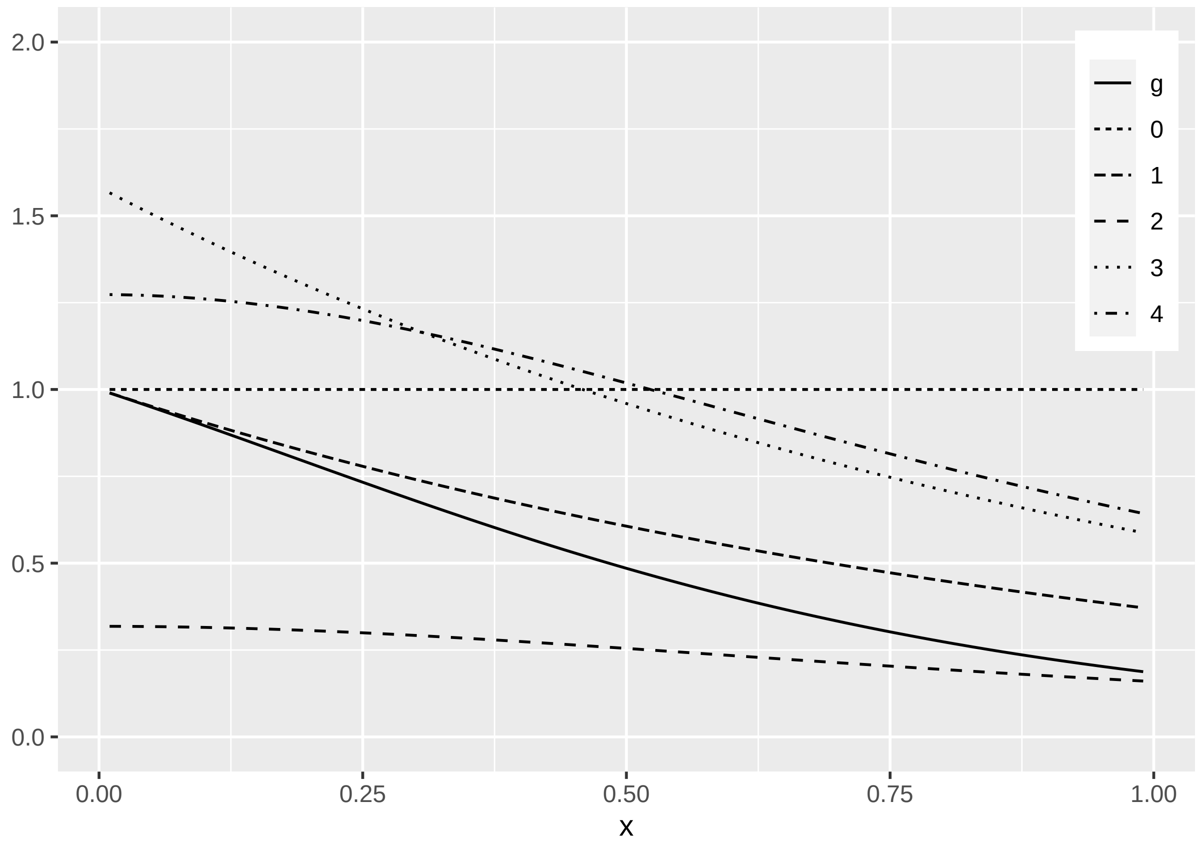
```
# Importance Functions
  # f0 = 1
rf0 <- function(m) {
  return(runif(m))
}
f0 <- function(x) {
  return(rep(1, length(x)))
}
  # f1 = e^(-x)
rf1 <- function(m) {
  return(rexp(m, 1))
}
f1 <- function(x) {
  return(exp(-x))
}
  # f2 = (1 + x^2)^-1 / pi
rf2 <- function(m) {
  x <- rcauchy(m)
  i <- c(which(x > 1), which(x < 0))
  x[i] <- 2
  return(x)
}
f2 <- function(x) {
  return(dcauchy(x))
}
  # f3 = e^(-x) / (1 - e^(-1))
rf3 <- function(m) {
```

```r
  u <- runif(m)
  return(-1 * log(1 - u * (1 - exp(-1))))
}
f3 <- function(x) {
  return(exp(-x) / (1 - exp(-1)))
}
  # f4 = 4(1 + x^2)^(-1) / pi
rf4 <- function(m) {
  u <- runif(m)
  return(tan(pi * u / 4))
}
f4 <- function(x) {
  return(4 * (1 + x^2)^(-1) / pi)
}

# Importance Sampling By Different Functions
m <- 10000
theta.hat <- var <- numeric(5)
theta.est <- function(m, g, rf, f) {
  x <- rf(m)
  fg <- g(x) / f(x)
  return(c(mean(fg), var(fg)))
}
theta.hat[1] <- theta.est(m, g, rf0, f0)[1]
var[1] <- theta.est(m, g, rf0, f0)[2]
theta.hat[2] <- theta.est(m, g, rf1, f1)[1]
var[2] <- theta.est(m, g, rf1, f1)[2]
theta.hat[3] <- theta.est(m, g, rf2, f2)[1]
var[3] <- theta.est(m, g, rf2, f2)[2]
theta.hat[4] <- theta.est(m, g, rf3, f3)[1]
var[4] <- theta.est(m, g, rf3, f3)[2]
theta.hat[5] <- theta.est(m, g, rf4, f4)[1]
var[5] <- theta.est(m, g, rf4, f4)[2]

# Results
results <- rbind(theta.hat, var)
colnames(results) <- c("f0", "f1", "f2", "f3", "f4")
rownames(results) <- c("theta.hat", "var")
results
```

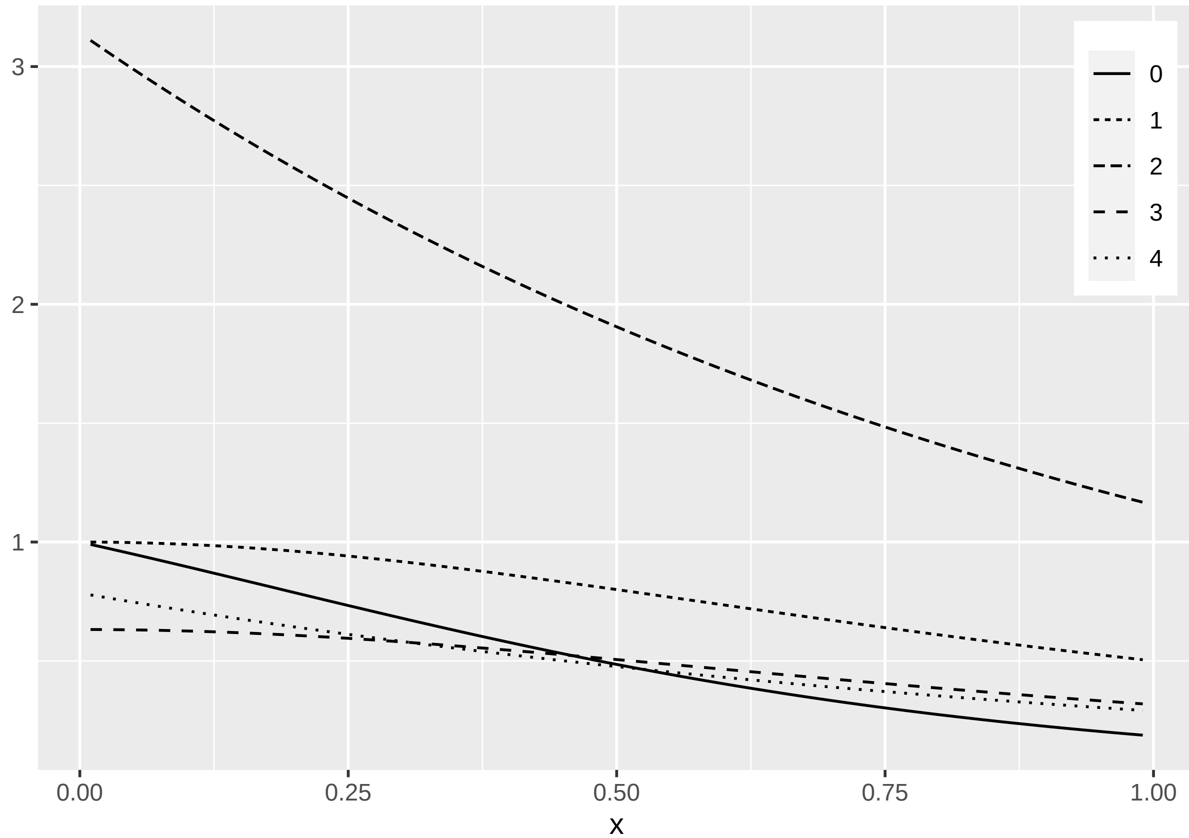
```
##                   f0        f1        f2          f3         f4
## theta.hat 0.52631065 0.5237532 0.5121693 0.523956396 0.52328145
## var       0.06005922 0.1749603 0.9221703 0.009425546 0.02019974
```

```r
# Plots
x <- seq(0.01, 0.99, .01)
w <- 2
# Figure (a)
df1 <- data.frame(x = rep(x, 6), variable = c(rep("g", length(x)), rep("0", length(x)),
                                              rep(1, length(x)), rep(2, length(x)),
                                              rep(3, length(x)), rep(4, length(x))),
                  value = c(g(x), f0(x), f1(x), f2(x), f3(x), f4(x)))
df1$variable <- factor(df1$variable, levels = c("g", "0", "1", "2", "3", "4"))
```

```r
ggplot(df1, aes(x = x, y = value, linetype = variable)) +
  geom_line() +
  scale_y_continuous(limits = c(0, 2)) +
  labs(y = "") +
  theme(legend.title = element_blank(), legend.position = c(0.94, 0.76))
```




```r
# Figure (b)
df2 <- data.frame(x = rep(x, 5), variable = c(rep(0, length(x)), rep(1, length(x)),
                                              rep(2, length(x)), rep(3, length(x)),
                                              rep(4, length(x))),
                  value = c(g(x), g(x)/f1(x), g(x)/f2(x), g(x)/f3(x), g(x)/f4(x)))
df2$variable <- factor(df2$variable, levels = c("0", "1", "2", "3", "4"))

ggplot(df2, aes(x = x, y = value, linetype = variable)) +
  geom_line() +
  labs(y = "") +
  theme(legend.title = element_blank(), legend.position = c(0.94, 0.8))
```



2. Stratified Sampling

```r
M <- 10000
k <- 10
r <- M / k
N <- 50
T2 <- numeric(k)
estimates <- matrix(0, N, 2)

for (i in 1:N) {
  estimates[i, 1] <- mean(g(runif(M)))
  for (j in 1:k)
      T2[j] <- mean(g(runif(r, (j-1)/k, j/k)))
  estimates[i, 2] <- mean(T2)
}

results <- rbind(apply(estimates, 2, mean), apply(estimates, 2, var))
colnames(results) <- c("Standard", "Stratified Sampling")
rownames(results) <- c("theta.hat", "var")
results
```

```
##               Standard Stratified Sampling
## theta.hat 5.256466e-01        5.247863e-01
## var       5.639998e-06        4.619082e-08
```

3. Stratified Importance Sampling

```r
# Using f3 as Importance Function
rf3 <- function(m, a, b) {
  u <- runif(m, a, b)
  return(-1 * log(1 - u * (1 - exp(-1))))
}
f3 <- function(x) {
  return(exp(-x) / (1 - exp(-1)))
}

# Stratified Importance Sampling
M <- 10000
k <- 10
r <- M / k
N <- 50
T2 <- numeric(k)
T3 <- numeric(k)
estimates <- matrix(0, N, 3)

for (i in 1:N) {
  estimates[i, 1] <- mean(g(runif(M)))
  for (j in 1:k) {
    T2[j] <- mean(g(runif(r, (j-1)/k, j/k)))
    x <- rf3(r, (j-1)/k, j/k)
    T3[j] <- mean(g(x) / f3(x))
  }
  estimates[i, 2] <- mean(T2)
  estimates[i, 3] <- mean(T3)
}

results <- rbind(apply(estimates, 2, mean), apply(estimates, 2, var))
colnames(results) <- c("Standard", "Stratified Sampling", "Stratified Importance Sampling")
rownames(results) <- c("theta.hat", "var")
results
```

```
##                 Standard Stratified Sampling Stratified Importance Sampling
## theta.hat 5.244650e-01        5.248126e-01                   5.247795e-01
## var       5.250212e-06        4.775609e-08                   1.107603e-08
```

4. Stratified Importance Sampling

```r
M <- 10000
k <- 1
m <- M/k
si <- numeric(k)
v <- numeric(k)
g <- function(x) exp(-x)/(1 + x^2)
f <- function(x) (k/(1 - exp(-1))) * exp(-x)

for (j in 1:k) {
  u <- runif(m, (j - 1)/k, j/k)
  x <- -log(1 - (1 - exp(-1)) * u)
  fg <- g(x)/f(x)
  si[j] <- mean(fg)
```

```r
  v[j] <- var(fg)
}
sum(si)
```

```
## [1] 0.5225514
```

```r
mean(v)
```

```
## [1] 0.009499111
```