Variance reduction

Vitaliy Pozdnyakov

January 2020

Preferential sampling

Let

$$g(x) = \max\left(0, K - e^{\beta x}\right)$$

Try to estimate $\mathbb{E}g(X)$ where $X \sim N(0,1)$. Analytical solution

$$\mathbb{E}g(X) = K\Phi\left(\frac{\log K}{\beta}\right) - e^{\frac{1}{2}\beta^2}\Phi\left(\frac{\log K}{\beta} - \beta\right)$$

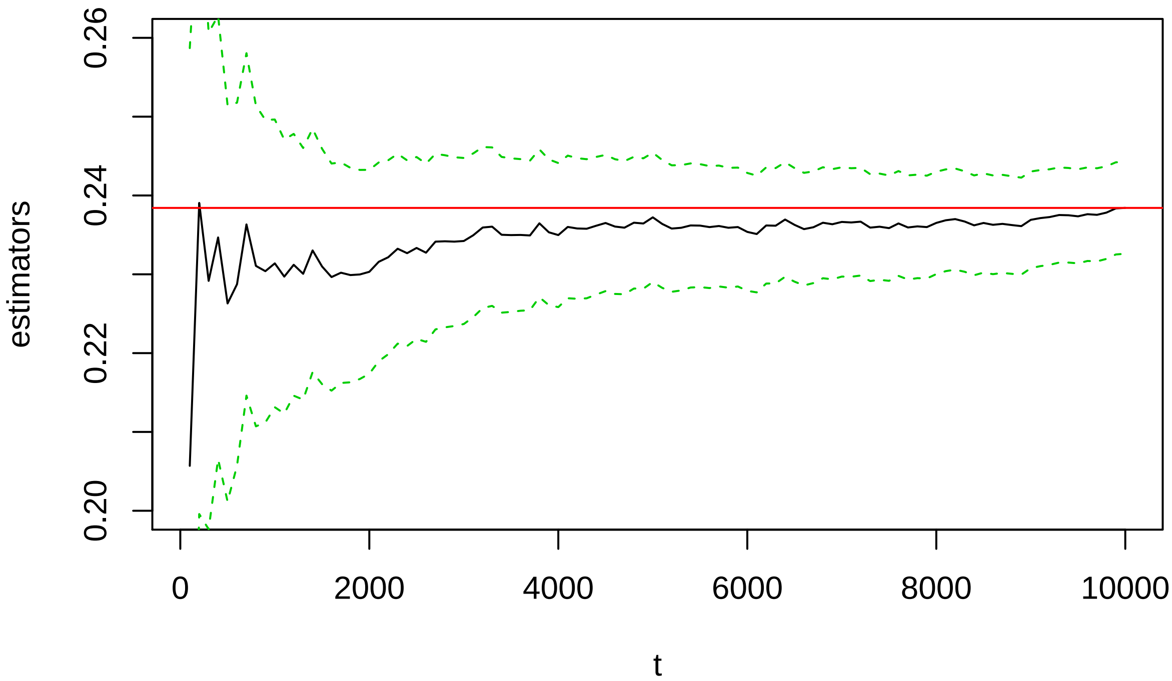
Generate instances of $g(X)$

```
set.seed(123)
n <- 10000
beta <- 1
K <- 1
x <- rnorm(n)
y <- sapply(x, function(x) max(0, K-exp(beta*x)))

# the true value
true_value <- K*pnorm(log(K)/beta)-exp(beta^2/2)*pnorm(log(K)/beta-beta)

t <- seq(100, n, 100)
upper_bound = numeric(n/100)
lower_bound = numeric(n/100)
estimators = numeric(n/100)
for(i in (1:100)){
  res <- t.test(y[1:t[i]])
  lower_bound[i] <- res[["conf.int"]][1]
  upper_bound[i] <- res[["conf.int"]][2]
  estimators[i] <- res[["estimate"]]}

plot(t, estimators, type = "l", ylim = c(0.20, 0.26), col=palette()[1])
lines(t, lower_bound, lty = 2, col=palette()[3])
lines(t, upper_bound, lty = 2, col=palette()[3])
abline(h = true_value, col=palette()[2])
```



Variance reduction. Rewrite $\mathbb{E}g(X)$ as

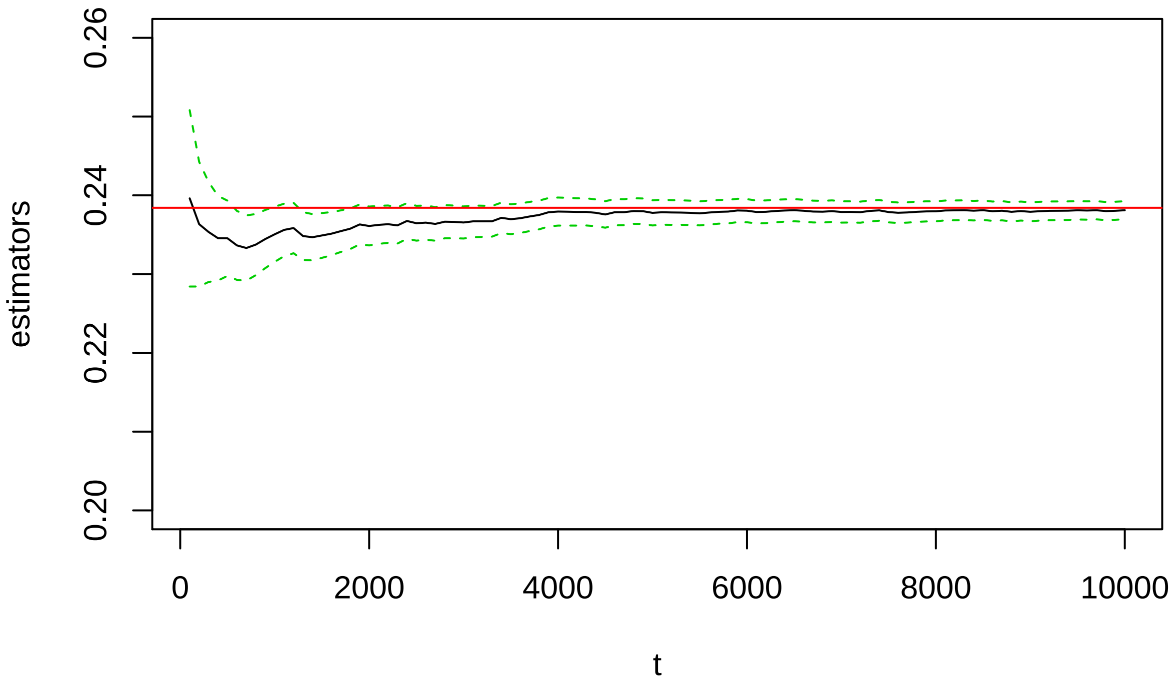
$$\mathbb{E}\left(\frac{\max\left(0, 1 - e^{\beta\sqrt{Y}}\right) + \max\left(0, 1 - e^{-\beta\sqrt{Y}}\right)}{\sqrt{2\pi Y}}\right)$$

where Y is Poisson with $\lambda = 1/2$ and $K = 1$

```
n <- 10000
beta <-1
K <- 1
x <- rexp(n,rate=0.5)
h <- function(x) (max(0,1-exp(beta*sqrt(x)))+max(0,1-exp(-beta*sqrt(x))))/sqrt(2*pi*x)
y <- sapply(x, h)

t <- seq(100, n, 100)
upper_bound = numeric(n/100)
lower_bound = numeric(n/100)
estimators = numeric(n/100)
for(i in (1:100)){
  res <- t.test(y[1:t[i]])
  lower_bound[i] <- res[["conf.int"]][1]
  upper_bound[i] <- res[["conf.int"]][2]
  estimators[i] <- res[["estimate"]]}

plot(t, estimators, type = "l", ylim = c(0.20, 0.26), col=palette()[1])
lines(t, lower_bound, lty = 2, col=palette()[3])
lines(t, upper_bound, lty = 2, col=palette()[3])
abline(h = true_value, col=palette()[2])
```


0.20
0.22
0.24
0.26
estimators
0
2000
4000
6000
8000
10000
t