Preferential sampling

Vitaliy Pozdnyakov

January 2020

```r
# ex1.02.R
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
K*pnorm(log(K)/beta)-exp(beta^2/2)*pnorm(log(K)/beta-beta)
```

```
## [1] 0.2384217
```

```r
t.test(y[1:100])
```

```
##
##  One Sample t-test
##
## data:  y[1:100]
## t = 7.701, df = 99, p-value = 1.043e-11
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  0.1526982 0.2586975
## sample estimates:
## mean of x
## 0.2056978
```

```r
t.test(y[1:1000])
```

```
##
##  One Sample t-test
##
## data:  y[1:1000]
## t = 24.877, df = 999, p-value < 2.2e-16
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  0.2131347 0.2496388
## sample estimates:
## mean of x
## 0.2313868
```

```r
t.test(y)
```

```
##
##  One Sample t-test
##
## data:  y
## t = 80.356, df = 9999, p-value < 2.2e-16
## alternative hypothesis: true mean is not equal to 0
```

```
## 95 percent confidence interval:
##  0.2326121 0.2442446
## sample estimates:
## mean of x
## 0.2384284
```

```r
# ex1.03.R
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)

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

```
## [1] 0.2384217
```

```r
t.test(y[1:100])
```

```
##
##  One Sample t-test
##
## data:  y[1:100]
## t = 42.467, df = 99, p-value < 2.2e-16
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  0.2284211 0.2508125
## sample estimates:
## mean of x
## 0.2396168
```

```r
t.test(y[1:1000])
```

```
##
##  One Sample t-test
##
## data:  y[1:1000]
## t = 131.76, df = 999, p-value < 2.2e-16
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  0.2315555 0.2385568
## sample estimates:
## mean of x
## 0.2350561
```

```r
t.test(y)
```

```
##
##  One Sample t-test
##
## data:  y
## t = 409.01, df = 9999, p-value < 2.2e-16
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  0.2369690 0.2392513
```

```
## sample estimates:
## mean of x
## 0.2381101
```