MTH210: Lab 6

Importance Sampling

1. **Moments of Gamma Distribution using Importance Sampling:**

   We will estimate moments of Gamma(α, β) distribution, so that we are interested in $h(x) = x^k$ for some $k \geq 1$. Let $X \sim$ Gamma(α, β), where α is the shape and β is the rate parameter.

   The Simple Monte Carlo estimator of $E_F[X^k]$ can be found by sampling from Gamma and taking sample means. Suppose $\alpha = 2$ and $\beta = 5$. I will obtain 1e3 samples from this Gamma and suppose $k = 2$ (second moment). Luckily for this problem, I know the true values of the second moment, and can compare the results with this value

$$\frac{\alpha}{\beta^2} + \left(\frac{\alpha}{\beta}\right)^2$$

```r
k <- 2
N <- 1e3
alpha <- 2
beta <- 5

(truth <- (alpha / beta^2) + (alpha/beta)^2)  # true second moment
#[1] 0.24

samples <- rgamma(N, shape = alpha, rate = beta)
# simple Monte Carlo
mean(samples^k)
```

   When you run the above, you get an estimator of the required expectation. We can also do this using importance sampling. Suppose the choice of importance distribution is Exp(λ), which has density $g(x)$. Then recall that for $Z_1, Z_2, \dots, Z_t \sim G$ the importance sampling estimator for this g is

$$\hat{\theta}_g = \frac{1}{N} \sum_{t=1}^{N} Z_t^k \frac{f(Z_t)}{g(Z_t)}$$

   We can now obtain the importance sampling estimator with $\lambda = 3$ say.

```r
set.seed(1)
lambda <- 3 #proposal

N <- 1e4
samp <- rexp(N, rate = lambda)  # importance samples

#evaluate inside the sum
funcs <- samp^k * dgamma(samp, shape = alpha, rate = beta) / dexp(samp, rate = lambda)

# take the average gives me the Importance Sampling estimator
mean(funcs)

# Estimate of sigma^2_g: Sample variance of h(Z)f(Z)/g(Z)
var(funcs)
```

We obtain the estimate $\hat{\theta}_g$ in `mean(funcs)`, and in `var(funcs)` we obtain an estimate of σ_g^2

$$\sigma_g^2 = \text{Var}_G\left(\underbrace{h(Z)\frac{f(Z)}{g(Z)}}_{s(Z)}\right).$$

In order to estimate σ_g^2, we obtain all the $s(Z_t)$ and take the sample variance of the $s(Z_t)$. This is what `var(funcs)` does.

a. Repeat the above experiment for different values of λ. Which λ value returns the smallest estimate of σ_g^2?

b. Repeat the above experiment for a different proposal Gamma$(3, 3)$.

c. Repeat the above experiment for $\alpha = 4, \beta = 10, k = 3$ and the importance distribution Gamma$(7, 10)$. What are the values in `funcs`? What is `var(funcs)` in this case? Why?

2. **Law of Large Numbers:**

We would like to "verify" the claim of convergence in probability of the importance sampling estimator. We know theoretically that as $N \to \infty$,

$$\hat{\theta}_g \xrightarrow{p} \theta\,,$$

where, recall $\theta = E_F[h(X)]$. In order to visualize this result, we will generate large number of samples from G (N is large) and plot the average $\hat{\theta}_g$ of the samples as N increases.

```r
## Checking convergence
N <- 1e5 # very large N
```

```r
samp <- rexp(N, rate = lambda)  # importance samples

func <- samp^k * dgamma(samp, shape = alpha, rate = beta) / dexp(samp, rate = lambda)


x.axis <- 1:N # sample size on the x-axis
y.axis <- cumsum(func)/(1:N)  # IS estimator for each N

# Plotting the running average
plot(x.axis, y.axis , type = 'l', xlab = "N", ylab = "Running average")
abline(h = truth, col = "red")
```

3. **Implement Problem 3 from Exercises of Section 5 in R.**

   Note, you may choose a grid of values for t:

```r
# choosing 50 values of t between (-5, 5)
t <- seq(-5, 5, length = 50)
```