Computational Methods in Statistics I

Final Project

prepared by

Tamika Royal

Department of Statistics

Florida State University

Goal

The main purpose of the project is to estimate the right tail of an exponential random variable using ideas from importance sampling and tilted distributions. Another aim of the project is to compare the classical Monte Carlo estimator for θ with the estimator using the Importance sampling with tilted distributions for this random variable.

Problem Statement

We are given X which is an exponential random variable with mean $1/\lambda$ or intensity λ. For a constant a > 0, our goal is to estimate the probability:

$\theta = \Pr\{X>a\} = \lambda\int_a^\infty e^{-\lambda x}dx$ Since the event X > a occurs with very small probability, the use of the Classical Monte Carlo approach is not efficient, hence, it is better to use importance sampling with a tilted density serving as the density to sample from.

Introduction

In analyzing the idea behind Importance sampling one has to explore the concept of the Monte Carlo technique. Many inference problems can be written as integrals under some given probability measure. In some situations this probability measure is too difficult to analytically integrate out therefore leading to the use of numerical approximations. One technique for numerical approximation is by sampling, that is, to approximate the integral using samples generated from the given probability measure. [2] This method of approximating the integral is called the Monte Carlo method where the main goal is to estimate θ where:

$$\theta = \int g(x) f(x) dx = E(g(x))$$

For a random variable X which has distribution density function f(x); g(x) is any function on R such that θ and $Eg(X)^2$ are bounded. Suppose we sample $X_1 \ldots X_n$ ~iid f(x) then one can approximate θ by $\hat{\theta}_n = \frac{1}{n}\sum_{i=1}^{n} g(X_i)$; $\hat{\theta}_n$ is an estimator of θ because E($\hat{\theta}_n$) = θ. The variance of $\hat{\theta}_n$ is 1/n*Var(g(X)).

It is possible to reduce the variance of the estimator by various methods which will not be ventured into for purposes of this project. One method of reducing variance in Monte Carlo methods is the importance sampling. The general idea behind importance sampling will be explored in the methodology section as the Monte Carlo Method sets the foundation for it.

Methodology:

(a) **Importance sampling**:

Importance sampling is a general technique for estimating the properties of a particular distribution, while only having samples generated from a different distribution rather than the distribution of interest. Importance sampling is different from the Monte Carlo methods in that instead of sampling from f(x) one samples from another density h(x) and computes the estimate of $\bar{\theta}$ using averages of g(x)f(x)/h(x) instead of g(x) evaluated on those samples. The distribution of θ now becomes:

$$\theta = \int g(x)f(x)dx = \int \frac{g(x)f(x)}{h(x)}h(x)dx$$

h(x) can be any density function as long as the support of h(x) contains the support of f(x).

The idea is to generate samples X_1, X_2…….X_n from the density h(x) and compute the estimate:

$$\hat{\theta}_n = \frac{1}{n}\sum_{i=1}^{n}\frac{g(X_i)f(X_i)}{h(X_i)}$$

The mean of $\hat{\theta}$ is θ and its **variance** is: $\frac{1}{n}\left(E_h\left[\frac{g(X)f(X)}{h(X)}\right]^2 - \theta^2\right)$. The choice of h(x) usually reduces the estimator variance below that of the classical Monte Carlo estimator.

Tilted Sampling: is a specific case of importance sampling where the sampling distribution is a tilted version of the original density function. The idea behind this is that there may be cases where one is interested in estimating the tail probabilities of a distribution. It may be useful to tilt the density while raising the tail probability in cases where the tails are negligible. This is the idea behind the project as the goal is to estimate the tail of the exponential random variable with mean ($1/\lambda$) but because X > a occurs with small probability then instead of using the classical Monte Carlo we use Importance sampling with a tilted density.

(b) In application to the project question, for a scalar t > 0, the tilted density is defined as follows:

$$f_t(x) = \frac{e^{tx} f(x)}{M(t)} \text{ where f(x) = } \lambda e^{-\lambda x}$$

$$m(t) = \int_0^\infty e^{tx} f(x) dx = \lambda \int_0^\infty e^{tx} . e^{-\lambda x} dx = \lambda \int_0^\infty e^{-x(\lambda - t)} dx = -\frac{\lambda}{\lambda - t}[0 - 1] = \frac{\lambda}{\lambda - t}$$

$$f_t(x) = \frac{e^{tx} f(x)}{M(t)} = \frac{e^{tx} \lambda e^{-\lambda x}}{\lambda / \lambda - t} = (\lambda - t) e^{-(\lambda - t)x}$$

From above we see that the tilted density has exponential distribution with mean $1/\lambda - t$ or intensity λ -t .

We can sample from the tilted density by using the Inverse Transform Method where we generate U ~ [0,1] and Xi's ~ - $1/\lambda - t$ log(U).

(c) The optimal amount of tilt t to estimate θ is such that the mean of $f_t(x)$ is equal to a.

Since the mean of $f_t(x)$ = $1/\lambda - t$ then a = $1/\lambda - t$ which implies that t* = $\lambda - (1/a)$

(d) Xi ~ $f_{t*}(x)$ from the optimally tilted density. Therefore X_i ~ exponential distribution with mean $1/\lambda - t^*$ or intensity λ-t*. Therefore substituting t* = $\lambda - (1/a)$ then Xi ~ exponential distribution with mean a or intensity 1/a.

We can sample from the X_i by the Inverse Transform Method by generating U~[0,1] and X_i's ~ a*log(U).

Therefore the expression for $\tilde{\theta}_n = \frac{M(t^*)}{n} \sum_{i=1}^{n} 1_{[a,\infty]}(x_i).e^{-tx}$

Since $M(t^*) = \frac{\lambda}{\lambda - t^*} = \frac{\lambda}{\lambda - (\lambda - \frac{1}{a})} = a\lambda$

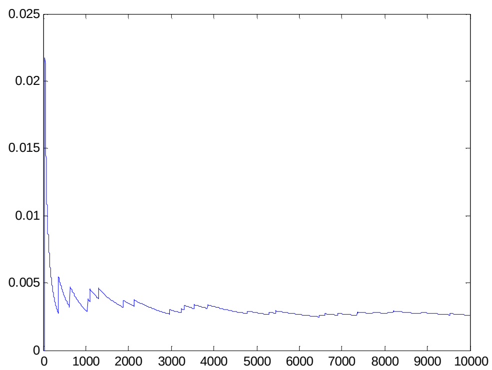
Therefore $\tilde{\theta}_n = \frac{a\lambda}{n} \sum_{i=1}^{n} 1_{[a,\infty]}(x_i).e^{-(\lambda - \frac{1}{a})xi}$

The Classical Monte Carlo estimator is $\hat{\theta}_n = \frac{1}{n} \sum_{i=1}^{n} 1_{[a,\infty]}(x_i)$ where the X_i's ~ exponential with mean $(1/\lambda)$ or intensity λ.

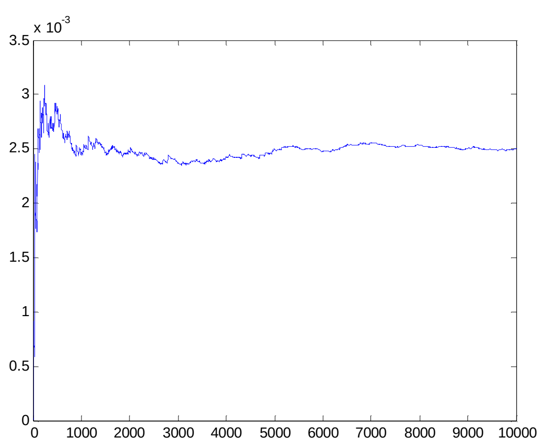
Experimental Results:

Comparing the Classical MC and Important Sampling on Computation time and accuracy.

<u>Convergence of Classical MC for a=6& λ=1</u>



<u>Convergence of Tilted Density for a=6 & λ=1</u>



<u>Accuracy of Estimates</u>

Classical estimator = 0.0018 and the tilted = 0.0025. The true estimate is 0.0025 which indicates that the tilted estimate is more accurate than the classical estimator since it is closer in value to the true value. Also the tilted estimate was more consistent than the tilted estimator.

<u>Computation Time</u>

classical = 3.3017

tilted =3.2006

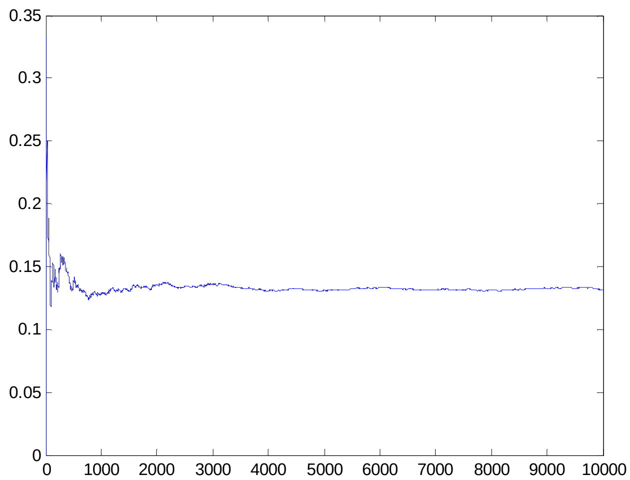
The tilted takes a shorter computation time than the classical estimator and from observance of the convergence graphs one can see that it takes longer for the classical estimator to converge.

Plotting Convergence for the Classical MC estimator and the Tilted estimator

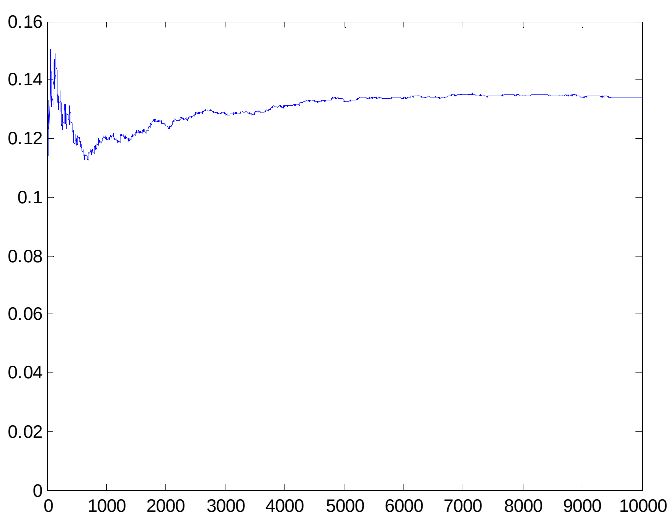
<u>For a=2 and lambda=1:</u>

True value=0.1353

<u>Convergence of Classical MC for a=2& λ=1</u>



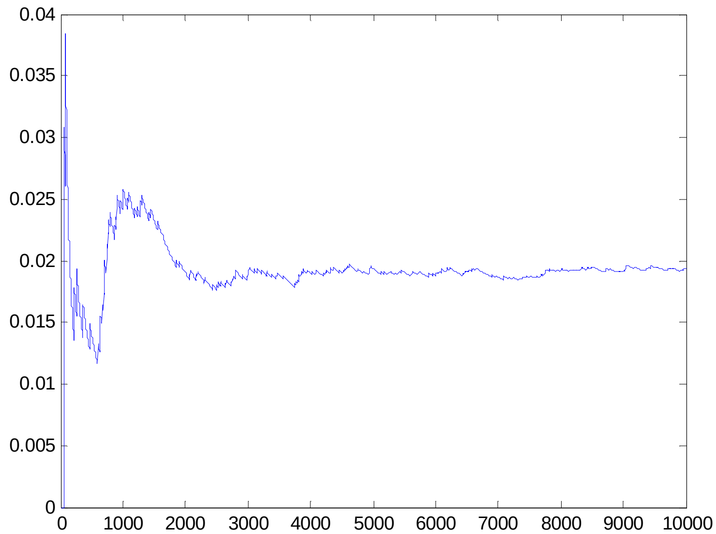
<u>Convergence of Tilted Density for a=2 & λ=1</u>



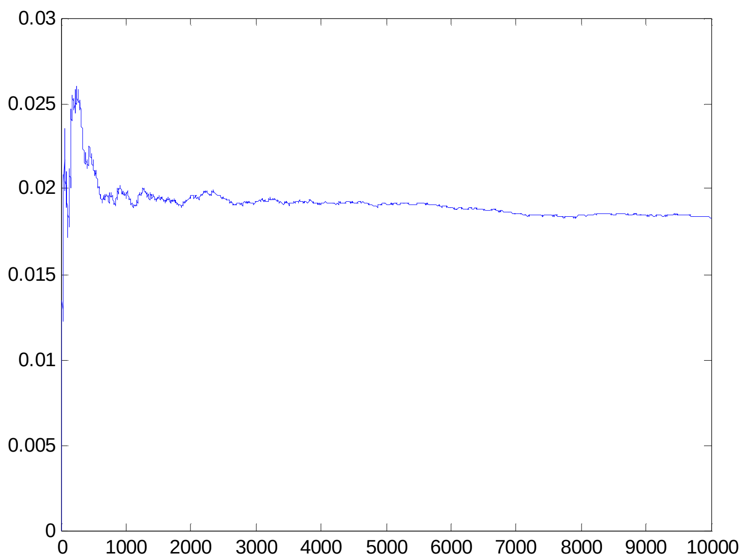
For a=4 and lambda=1:

True value=0.0183

Convergence of Classical MC for a=4& λ=1



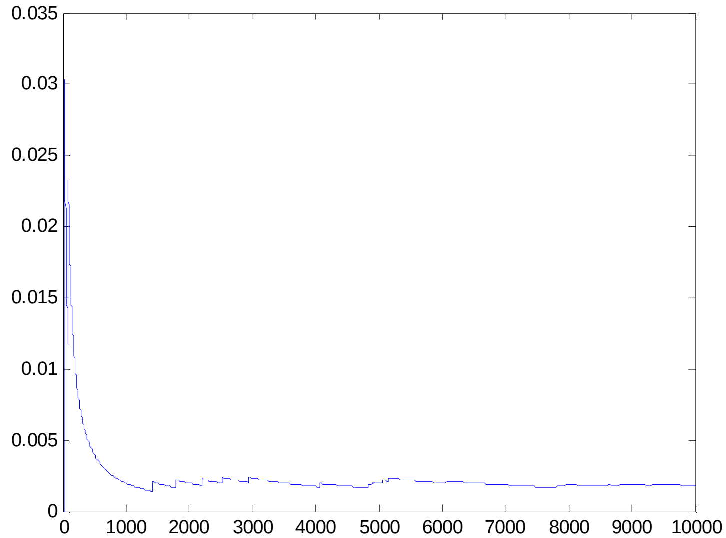
Convergence of Tilted Density for a=4 & λ=1



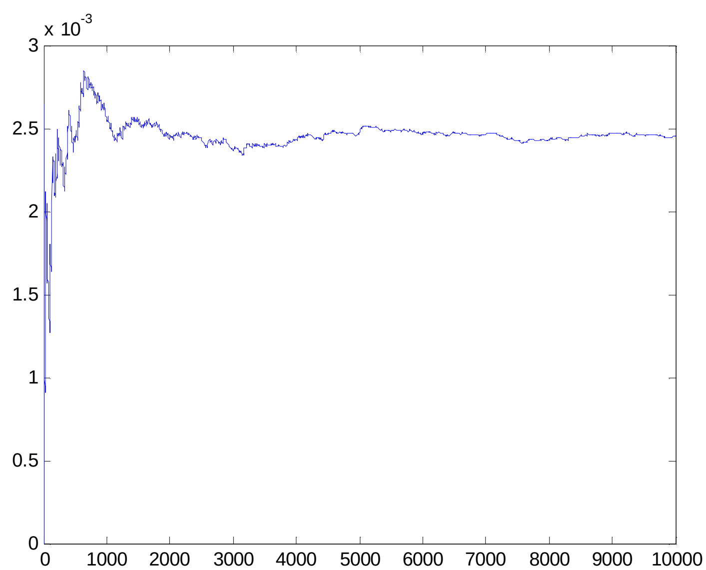
For a=6 and lambda=1:

True value=0.0025

Convergence of Classical MC for a=6& λ=1



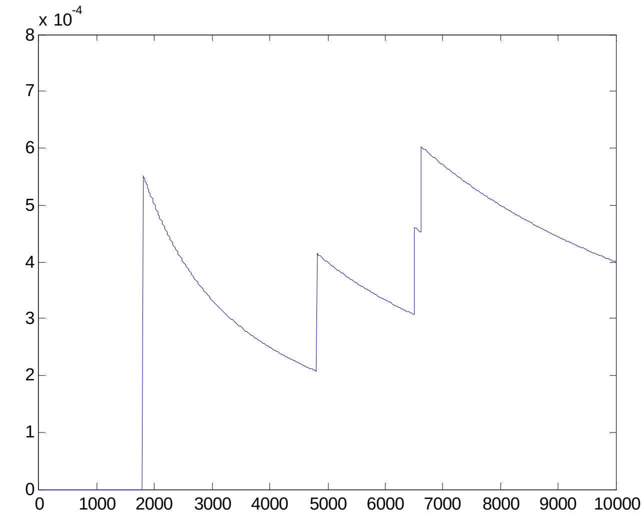
Convergence of Tilted Density for a=6 & λ=1



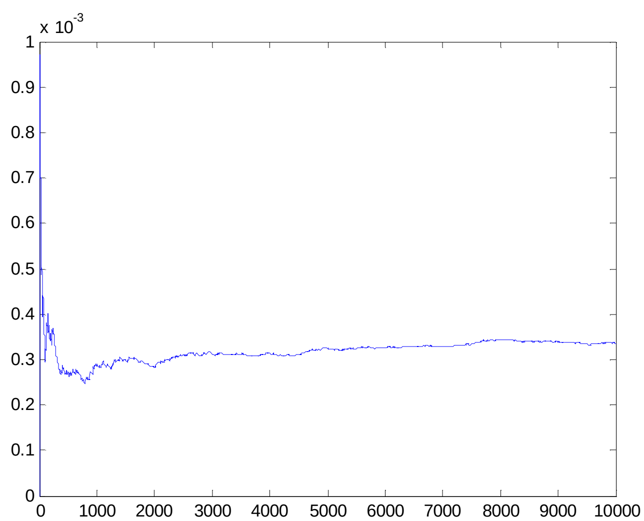
For a=8 and lambda=1:

True value =3.35e-004

Convergence of Classical MC for a=8 & λ=1



Convergence of Tilted Density for a=8 & λ=1



<u>Table of final Estimates</u>

Constant a	$\hat{\theta}_n$ – Classical MC	$\tilde{\theta}_n$ – Tilted Density	True value
2	0.1317	0.1366	0.1353
4	0.0193	0.0183	0.0183
6	0.0018	0.0025	0.0025
8	4.0000e-004	3.3617e-004	3.35e-004

Conclusion

From the results of the final estimates (see Table) one can conclude that the tilted density estimator is a better estimator for θ as the values are closer to the true values of the integral which indicates that it is more accurate that the classical estimator. The computational time is shorter for the tilted estimator and the convergence graphs indicate that the tilted estimator is more stable and converges quickly to a value while the classical graph is less stable and takes a longer time to converge. The tilted density is a better estimator for the tail of the exponential density as seen from the convergence at the different values of a and is it's accuracy and shorter computational time gives it a clear advantage over the classical method.

References

1. http://en.wikipedia.org/wiki/Importance_sampling
2. ClassNotes

Appendix

Matlab Commands

```
clear all; clc;
a=6;
lambda = 1;
n=10000;
%a=1/(lambda-t);
t=lambda -(1/a);
mt= lambda/(lambda-t);
u=rand(1,n);
%classical estimate
tic
for i=1:n
    %u(i) = rand;
    x1(i) = -log(u(i));
    if x1(i)>=a
        gx1(i)=1;
    else
        gx1(i)= 0;
    end
    thetahat(i)=(1/i)*sum(gx1(1:i));
end
classical=toc;
thetahat(i)
plot(thetahat)

%tilted sampling
tic
for i=1:n
    %u(i) = rand;
    x(i) = -a*log(u(i));
    if x(i)>=a
        gx(i)=exp(-t*x(i));
    else
        gx(i)= 0;
    end
    thetatilda(i)=(mt/i)*sum(gx(1:i));
end
tilted=toc;
thetatilda(i)
plot(thetatilda)
classical
tilted
```

Commands – Part 4(b)

```
clear all; clc;
for a=2:2:8;
lambda = 1;
n=10000;
%a=1/(lambda-t);
t=lambda -(1/a);
mt= lambda/(lambda-t);
u=rand(1,n);
%classical estimate
for i=1:n
    x1(i) = -log(u(i));
    if x1(i)>=a
        gx1(i)=1;
    else
        gx1(i)= 0;
    end
    thetahat(i,a)=(1/i)*sum(gx1(1:i));
end

%tilted sampling
for i=1:n
     x(i) = -a*log(u(i));
    if x(i)>=a
        gx(i)=exp(-t*x(i));
    else
        gx(i)= 0;
    end
    thetatilda(i,a)=(mt/i)*sum(gx(1:i));
end
end
plot(thetahat(:,2))
plot(thetahat(:,4))
plot(thetahat(:,6))
plot(thetahat(:,8))


thetahat(i,2)
thetahat(i,4)
thetahat(i,6)
thetahat(i,8)

plot(thetatilda(:,2))
plot(thetatilda(:,4))
plot(thetatilda(:,6))
plot(thetatilda(:,8))
thetatilda(i,2)
thetatilda(i,4)
thetatilda(i,6)
thetatilda(i,8)
thetahat(i,2)
thetahat(i,4)
thetahat(i,6)
thetahat(i,8)
```