STA 380

Week 6.

4. We want to estimate the following integral using importance sampling:

$$\theta = \int_1^4 \exp\left\{-\frac{(\ln x)^2}{2}\right\} dx$$

(a) Use `integrate()` to compute the "true" value of θ.
Disclaimer: you need to know how to integrate for the tests; we also just want you to learn the `integrate()` function as well. You will not be using the `integrate()` function in `R` for the term tests or exam.

(a)

```
fun <- function(x){exp(-(log(x))^2 / 2)}
true <- integrate(fun, lower = 1, upper = 4)
```

2.032086

(b) Estimate θ using the simple Monte Carlo estimator. That is, the importance function is the pdf of some uniform distribution.

$$\theta = 3 * \int_1^4 \underbrace{\frac{1}{3}}_{\frac{1}{b-a}} \exp\left\{-\frac{(\ln x)^2}{2}\right\} dx$$

$$= 3\, E_{U(1,4)}\left[\exp\left\{-\frac{(\ln x)^2}{2}\right\}\right]$$

Algorithm :

① Generate $u_1 \dots u_n \sim \text{Uniform}(1, 4)$

② Compute $\widehat{E(g(x))} = \frac{1}{n}\sum_{i=1}^{n} \exp\left\{-\frac{(\ln x)^2}{2}\right\}$

③ Deliver $\hat{\theta} = 3 * \widehat{E(g(x))}$.

R:

```
n <- 10^5
u <- runif(n, 1, 4)
est1 <- 3 * mean(exp(-(log(u))^2/2))
```

(c) Consider the log-normal distribution:

$$f_X(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left\{-\frac{(\ln x - \mu)^2}{2\sigma^2}\right\}.$$

$\mu = 0$, $\sigma^2 = 1$

Estimate θ using the log-normal pdf as the importance function. Use $\sigma^2 = 1, \mu = 0$.

$\sigma = 1$

$\theta = b - a \int f_X(x) * \cancel{\text{target}}.\ dx.$

$$f_X(x) = \frac{1}{x\sqrt{2\pi}} \exp\left\{\frac{-(\ln x)^2}{2}\right\}$$

$$\theta = \int_1^4 \exp\left\{-\frac{(\ln x)^2}{2}\right\} dx$$

$$= \int_1^4 \exp\left\{-\frac{(\ln x)^2}{2}\right\} \cdot \frac{f_X(x)}{f_X(x)}\, dx$$

$$= \int_1^4 \exp\left\{-\frac{(\ln x)^2}{2}\right\} \cdot \frac{\frac{1}{x\sqrt{2\pi}} \exp\left\{\frac{-(\ln x)^2}{2}\right\}}{\frac{1}{x\sqrt{2\pi}} \exp\left\{\frac{-(\ln x)^2}{2}\right\}}\, dx$$

$$= \int_1^4 \frac{\left(\frac{1}{x\sqrt{2\pi}}\right) \exp\left\{\frac{-(\ln x)^2}{2}\right\}}{\frac{1}{x\sqrt{2\pi}}}\, dx$$

$$\frac{1}{x\sqrt{2\pi}} \div \frac{1}{x\sqrt{2\pi}} = \frac{x\sqrt{2\pi}}{x\sqrt{2\pi}}$$

$$= \int_1^4 \frac{x\sqrt{2\pi}\, \exp\left\{\frac{-(\ln x)^2}{2}\right\}}{x\sqrt{2\pi}}\, dx$$

$f_X(x)$

$$= E_{f_X(x)}\left[x\sqrt{2\pi} \cdot 1_{\{1 \le x \le 4\}}\right]$$

Algorithm :

① Let $x_1, \dots, x_n$ be random variables generated from LogNormal $(\mu = 0, \sigma^2 = 1)$

② $\hat{\theta} = \frac{1}{n} \sum_{i=1}^{n} x \sqrt{2\pi} \cdot \mathbb{1}_{\{1 \le x \le 4\}}$.

R codes

```
n <- 10^5
x <- rlnorm(n, mean=0, sd=1)
est2 <- mean(x* sqrt(2*pi) * (x>=1 & x<=4))
```

(d) I used `all.equal()` for comparison, although you're free to simply check the absolute distance or any other reasonable metric as well.

```
all.equal(true$value, est1)
## [1] "Mean relative difference: 0.0005045298"
all.equal(true$value, est2)
## [1] "Mean relative difference: 0.001480595"
```

$$\mathbb{V}ar(\hat{\theta}) = \frac{1}{n}\left[\left(\int \frac{g(x)^2}{f(x)}dx\right) - \theta^2\right]$$

This means that the variance of the first estimator is then:

$$\mathbb{V}ar(\hat{\theta_{(b)}})\left[\left(\int_1^4 3\exp\left\{-(\ln x)^2\right\}dx\right) - \theta^2\right]$$

And then the variance of the second must be:

$$\mathbb{V}ar(\hat{\theta}) = \frac{1}{n}\left[\left(\int_1^4 \exp\left\{-\frac{(\ln x)^2}{2}\right\} x\sqrt{2\pi}\,dx\right) - \theta^2\right]$$

Let's make our lives easier and `integrate()` in `R`:

```
fun1 <- function(x){3 * exp(-(log(x))^2)}
fun2 <- function(x){exp(-(log(x))^2/2) * x * sqrt(2*pi)}

int1 <- integrate(fun1, 1, 4)
int2 <- integrate(fun2, 1, 4)

# variance with uniform importance function
(1/n) * (int1$value - true$value^2)
## [1] 3.442339e-06
# variance with lognormal importance function
(1/n) * (int2$value - true$value^2)
## [1] 7.335948e-05
```

5. Recall the beta distribution, which has pdf:

$$f_X(x) = \left[\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\right] x^{\alpha-1}(1-x)^{\beta-1}, \quad 0 < x < 1.$$

$\Gamma(n) = (n-1)!$

Can we use antithetic variables to estimate the expected value of the $X \sim Beta(\underline{\alpha = 2}, \underline{\beta = 4})$ distribution? Why or why not? If the antithetic approach can be used, then use the antithetic approach to estimate the expected value of X.

$\Gamma(2) = \underline{1!} = 1$

$$f_X(x) = \left[\frac{\Gamma(6)}{\Gamma(2)\Gamma(4)}\right] x^{2-1}(1-x)^{4-1}, \quad 0 < x < 1$$

$$= \frac{5!}{1! \cdot 3!} x(1-x)^3$$

$$f_X(x) = 20 \cdot x(1-x)^3$$

$$E(X) = \int_0^1 x \cdot f_X(x)\,dx = \int_0^1 \underline{20 \cdot x^2(1-x)^3}\,dx.$$

monotone

Review:

- $U \sim$ Uniform $(0,1)$
- Pair $(U, 1-U)$

$$\hat{\mu}_{anti} = \frac{g(U) + g(1-U)}{2}$$

n pairs $\sum_{i=1}^{\frac{n}{2}}$

$$g(x) = 20x^2(1-x)^3$$

$$g'(x) = \ldots$$

☆: Not monotone

can not use antithetic approach.

6. Consider the exponential distribution with a scale parameter θ.

(a) For what values of θ can we compute $\mathbb{E}[X\mathbf{1}_{\{0<X<1\}}]$ using antithetic variables? As a reminder,

$$\mathbb{E}[X\mathbf{1}_{\{0<X<1\}}] = \int_0^1 x f_X(x)dx.$$

(b) Pick any value of θ as long as the antithetic approach is valid. Then, use the antithetic approach to estimate $\mathbb{E}[X\mathbf{1}_{\{0<X<1\}}]$.

(c) Compare the estimator of $\mathbb{E}[X\mathbf{1}_{\{0<X<1\}}]$ to the true value. You will need to remember integration by parts to evaluate this integral.

(a)

Exponential pdf: $f_X(x) = \frac{1}{\theta} e^{-\frac{x}{\theta}}$, $x \in [0, \infty)$.

$$E[X_1 \cdot 1_{\{0<x<1\}}] = 1\int_0^1 \frac{1}{1} x \cdot \frac{1}{\theta} e^{-\frac{x}{\theta}} dx$$

$$= E_{U(0,1)}\Big[\underbrace{\frac{1}{\theta} \cdot x \cdot e^{-\frac{x}{\theta}}}\Big].$$

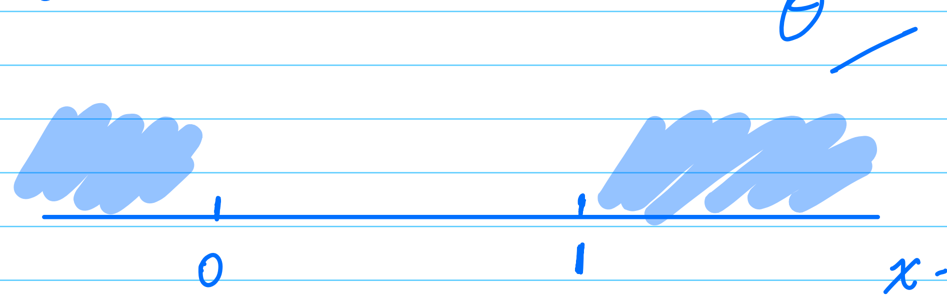
So $g(x) = \boxed{\frac{1}{\theta} \cdot x}\, e^{-\frac{x}{\theta}}$ check monotone.

$$g'(x) = \frac{1}{\theta} \cdot e^{-\frac{x}{\theta}} - \frac{1}{\theta^2} x e^{-\frac{x}{\theta}}$$

$$= e^{-\frac{x}{\theta}} \left(\frac{1}{\theta} - \frac{x}{\theta^2}\right)$$

if $=0$

$\frac{1}{\theta} = \frac{x}{\theta^2} \Rightarrow x = \theta.$ ← critical value.



Parameter of Exponential distri'l

$\theta > 0 \Rightarrow \theta$ should be greater 1

$\theta = 3$

(b) $\theta = 3$.

R codes: set.seed(☆)

n <- 10^4

U <- runif(n/2)

$E_{U(0,1)}\left[\underbrace{\frac{1}{\theta}\cdot x\cdot e^{-\frac{x}{\theta}}}\right]$.

theta <- 3

g(U) g1 <- (1/theta) * u * exp{-u/theta}.

g(1-U) g2 <- (1/theta) * (1-u) * exp{-(1-u)/theta}.

est <- mean((g1+g2)/2). $\underbrace{0.1337041}_{\text{estimator}}$

(c)

$$\mathbb{E}[X\mathbf{1}_{\{0<X<1\}}] = \int_0^1 x f_X(x)dx.$$

f <- function(x) {
(1/3) * x * exp(-x/3)}.

integrate(f, lower = 0, upper = 1) <= $\underbrace{0.1338748}$

Also Integration by parts ----- $-e^{-\frac{1}{3}} * (1+3)+3$ true