Importance sampling.

Exercises. [RC] Chapter 3.

Anatoli Iambartsev
IME-USP

Importance sampling.

Importance sampling is based on a alternative representation of the integral $\mathbb{E}_f(h(X))$. Given an arbitrary density g that is strictly positive when $h \cdot f$ is different from zero

$$\mathbb{E}_f(h(X)) = \int_{supp(g)} h(x)\frac{f(x)}{g(x)}dx = \mathbb{E}_g\Big[\frac{h(X)f(X)}{g(X)}\Big].$$

it justifies the use of the estimator

$$m_n^{IS} = \frac{1}{n}\sum_{i=1}^{n}\frac{f(X_i)}{g(X_i)}h(X_i) \to \mathbb{E}_f(h(X)),$$

where $X_i \sim g$ and the convergence is almost sure if $\mathbb{E}_g\Big|\frac{h(X)f(X)}{g(X)}\Big| < \infty$.

Exercise 3.4 [RC]. For the computation of the expectation $E_f[h(X)]$ when f is the normal pdf and $h(x) = \exp(-(x-3)^2/2) + \exp(-(x-6)^2/2)$.

(a) Show that $E_f[h(X)]$ can be computed in closed form and derive its value.

$$E_f[h(X)] = \frac{1}{\sqrt{2\pi}} \int \left(e^{-\frac{(x-3)^2}{2}} + e^{-\frac{(x-6)^2}{2}} \right) e^{-\frac{x^2}{2}} dx$$

$$= \frac{1}{\sqrt{2\pi}} \int e^{-(x-3/2)^2 - 9/4} dx + \frac{1}{\sqrt{2\pi}} \int e^{-(x-3)^2 - 9} dx$$

$$= \frac{e^{-9/4} + e^{-9}}{\sqrt{2}} \cong 0.0746.$$

Exercise 3.4 [RC]. For the computation of the expectation $E_f[h(X)]$ when f is the normal pdf and $h(x) = \exp(-(x-3)^2/2) + \exp(-(x-6)^2/2)$.

(b) Construct a regular Monte Carlo approximation based on a normal $N(0,1)$ sample of size $n = 10^3$ and produce an error evaluation.

$$m_n = \frac{1}{n}\sum_{i=1}^{n} h(X_i) \to \mathbb{E}_f(h(X)),\ \mathbb{V}ar_f(m_n) = \frac{\mathbb{V}ar_f h(X)}{n}$$

Let us calculate $\mathbb{V}ar_f h(X)$.

Exercise 3.4 [RC]. For the computation of the expectation $E_f[h(X)]$ when f is the normal pdf and $h(x) = \exp(-(x-3)^2/2) + \exp(-(x-6)^2/2)$.

(b) Construct a regular Monte Carlo approximation based on a normal $N(0,1)$ sample of size $n = 10^3$ and produce an error evaluation.

$$\mathbb{E}_f\left(e^{-\frac{(X-3)^2}{2}}\right) = \frac{e^{-9/4}}{\sqrt{2}}, \quad \mathbb{E}_f\left(e^{-\frac{(X-6)^2}{2}}\right) = \frac{e^{-9}}{\sqrt{2}}.$$

$$\mathbb{E}_f\left(e^{-(X-3)^2}\right) = \frac{1}{\sqrt{2\pi}} \int e^{-\frac{3}{2}(x-2)^2-3} dx = \frac{e^{-3}}{\sqrt{3}}$$

$$\mathbb{E}_f\left(e^{-(X-6)^2}\right) = \frac{1}{\sqrt{2\pi}} \int e^{-\frac{3}{2}(x-4)^2-12} dx = \frac{e^{-12}}{\sqrt{3}}$$

$$\mathbb{V}ar_f\left(e^{-\frac{(X-3)^2}{2}}\right) = \frac{e^{-3}}{\sqrt{3}} - \frac{e^{-9/2}}{2}, \quad \mathbb{V}ar_f\left(e^{-\frac{(X-6)^2}{2}}\right) = \frac{e^{-12}}{\sqrt{3}} - \frac{e^{-18}}{2}$$

$$\mathbb{E}_f\left(e^{-\frac{(X-3)^2}{2}} e^{-\frac{(X-6)^2}{2}}\right) = \frac{1}{\sqrt{2\pi}} \int e^{-\frac{3}{2}(x-3)^2-9} dx = \frac{e^{-9}}{\sqrt{3}}$$

$$cov_f\left(e^{-\frac{(X-3)^2}{2}}, e^{-\frac{(X-6)^2}{2}}\right) = \frac{e^{-9}}{\sqrt{3}} - \frac{e^{-(9/4+9)}}{2}.$$

Exercise 3.4 [RC]. For the computation of the expectation $E_f[h(X)]$ when f is the normal pdf and $h(x) = \exp(-(x-3)^2/2) + \exp(-(x-6)^2/2)$.

(b) Construct a regular Monte Carlo approximation based on a normal $N(0,1)$ sample of size $n = 10^3$ and produce an error evaluation.

$$\begin{aligned}
\mathbb{V}ar_f h(X) &= \mathbb{V}ar_f\left(e^{-\frac{(X-3)^2}{2}}\right) + \mathbb{V}ar_f\left(e^{-\frac{(X-6)^2}{2}}\right) + 2cov_f\left(e^{-\frac{(X-3)^2}{2}}, e^{-\frac{(X-6)^2}{2}}\right) \\
&= \frac{e^{-3}}{\sqrt{3}} - \frac{e^{-9/2}}{2} + \frac{e^{-12}}{\sqrt{3}} - \frac{e^{-18}}{2} + 2\left(\frac{e^{-9}}{\sqrt{3}} - \frac{e^{-(9/4+9)}}{2}\right) \\
&= \frac{e^{-3} + e^{-12} + 2e^{-9}}{\sqrt{3}} - \frac{e^{-9/2} + e^{-18} + 2e^{-(9/4+9)}}{2} \\
&\cong 0.0233 \\
r_n &= 0.6745\sqrt{\frac{0.0233}{n}} \cong 0.0032 \\
r_n^{0.95} &= 1.96\sqrt{\frac{0.0233}{n}} \cong 0.0094
\end{aligned}$$

Exercise 3.4 [RC]. For the computation of the expectation $E_f[h(X)]$ when f is the normal pdf and $h(x) = \exp(-(x-3)^2/2) + \exp(-(x-6)^2/2)$.

(b) Construct a regular Monte Carlo approximation based on a normal $N(0,1)$ sample of size $n = 10^3$ and produce an error evaluation.

$$\mathbb{E}_f\left(e^{-\frac{(X-3)^2}{2}} + e^{-\frac{(X-6)^2}{2}}\right) \cong 0.0746.$$

```
x=rnorm(1000)
y=exp(-(x-3)^2/2) + exp(-(x-6)^2/2)
mean(y)
> 0.07764772
```

$$\begin{aligned} CI_{95\%}\left(\mathbb{E}_f\left(e^{-\frac{(X-3)^2}{2}} + e^{-\frac{(X-6)^2}{2}}\right)\right) &\cong 0.0776 \pm 0.0094 \\ &= (0.0682, 0.087) \end{aligned}$$

Exercise 3.4 [RC]. For the computation of the expectation $E_f[h(X)]$ when f is the normal pdf and $h(x) = \exp(-(x-3)^2/2) + \exp(-(x-6)^2/2)$.

(c) Compare the above with an importance sampling approximation based on an importance function g corresponding to the $U[-8,-1]$ distribution and a sample of size Nsim=10^3. (Warning: This choice of g does not provide a converging approximation of $\mathbb{E}_f[h(X)]$)

$$m_n^{IS} = \frac{1}{n}\sum_{i=1}^{n} \frac{7}{\sqrt{2\pi}} e^{-X_i^2/2}\Big(e^{-(X_i-3)^2/2} + e^{-(X_i-6)^2/2}\Big)$$

where $X_i \sim U[-8,-1]$.

$$\mathbb{E}_g\Big(\frac{7}{\sqrt{2\pi}} e^{-X^2/2} h(X)\Big) = \frac{1}{\sqrt{2\pi}} \int_{-8}^{-1} e^{-x^2/2}\Big(e^{-(x-3)^2/2} + e^{-(x-6)^2/2}\Big)dx$$

$$\neq \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-x^2/2}\Big(e^{-(x-3)^2/2} + e^{-(x-6)^2/2}\Big)dx = \mathbb{E}_f\Big(h(X)\Big)$$

Defensive sampling.

[RC, p 81] "Given that importance sampling primarily applies in settings where f is not easy to study, this constraint on the tails of f is often not easy to implement, especially when the dimensionality is high. A generic solution nonetheless exists based on the artificial incorporation of a fat tail component in the importance function g. This solution is called *defensive sampling* by Hesterberg (1995)* and can be achieved by substituting a mixture density for the density g,

$$\rho g(x) + (1 - \rho)\ell(x), \quad 0 < \rho < 1,$$

where ρ is close to 1 and the density ℓ is chosen for its heavy tails (for instance, a Cauchy or a Pareto distribution), not necessarily in conjunction with the problem at hand."

*Hesterberg, T. (1995). Weighted average importance sampling and defensive mixture distributions. Technometrics, 37:185-194.

Example 3.9 [RC]. Consider the computing of the integral

$$\int_1^\infty \sqrt{\frac{x}{x-1}} t_2(x) dx = \frac{\Gamma(3/2)}{\sqrt{2\pi}} \int_1^\infty \sqrt{\frac{x}{x-1}} \frac{dx}{(1+x^2/2)^{3/2}}$$

$$= \mathbb{E}\left(\sqrt{\frac{X}{X-1}} \mathbb{1}(X > 1)\right) \quad \text{where } X \sim t_2.$$

The expectation exists despite of the singularity at $x = 1$, but the second moment is infinite.

This feature means that a mixture of the t_2 density with a well-behaved ℓ is required. To achieve integrability of $h^2(x)f(x)/\ell(x)$ calls for ℓ to be divergent in $x = 1$ and for ℓ to decrease slower than x^{-5} when x goes to infinity. Those boundary conditions suggest that

$$\ell(x) \propto \frac{1}{\sqrt{x-1}} \frac{1}{x^{3/2}} \mathbb{1}(x > 1),$$

(which is defined up to a constant) is an acceptable density.

Example 3.9 [RC].

To characterize this density, you can check that

$$\int_1^y \frac{dx}{\sqrt{x-1}x^{3/2}} = \int_0^{y-1} \frac{dw}{\sqrt{w}(w+1)^{3/2}} = \int_0^{\sqrt{y-1}} \frac{2d\omega}{(\omega^2+1)^{3/2}}$$

$$= \int_0^{\sqrt{2(y-1)}} \frac{\sqrt{2}dt}{(1+t^2/2)^{3/2}}$$

This implies that $\ell(x)$ corresponds to the density of $(1+T^2/2)$ when $T \sim t_2$, indeed, for $y > 1$

$$\mathbb{P}\left(1+\frac{T^2}{2} \le y\right) = \mathbb{P}\left(|T| \le \sqrt{2(y-1)}\right)$$

$$= 2\int_0^{\sqrt{2(y-1)}} \frac{\Gamma(3/2)}{\sqrt{2\pi}} \frac{dt}{(1+t^2/2)^{3/2}} = \int_1^y \frac{\Gamma(3/2)}{\sqrt{\pi}} \frac{dx}{\sqrt{x-1}x^{3/2}},$$

namely the following $\ell(x)$ is density function on $x \in (1, \infty)$

$$\ell(x) = \frac{\Gamma(3/2)}{\sqrt{\pi}} \frac{1}{\sqrt{x-1}x^{3/2}} \mathbb{1}(x > 1).$$

Example 3.9 [RC].

checking numerically:

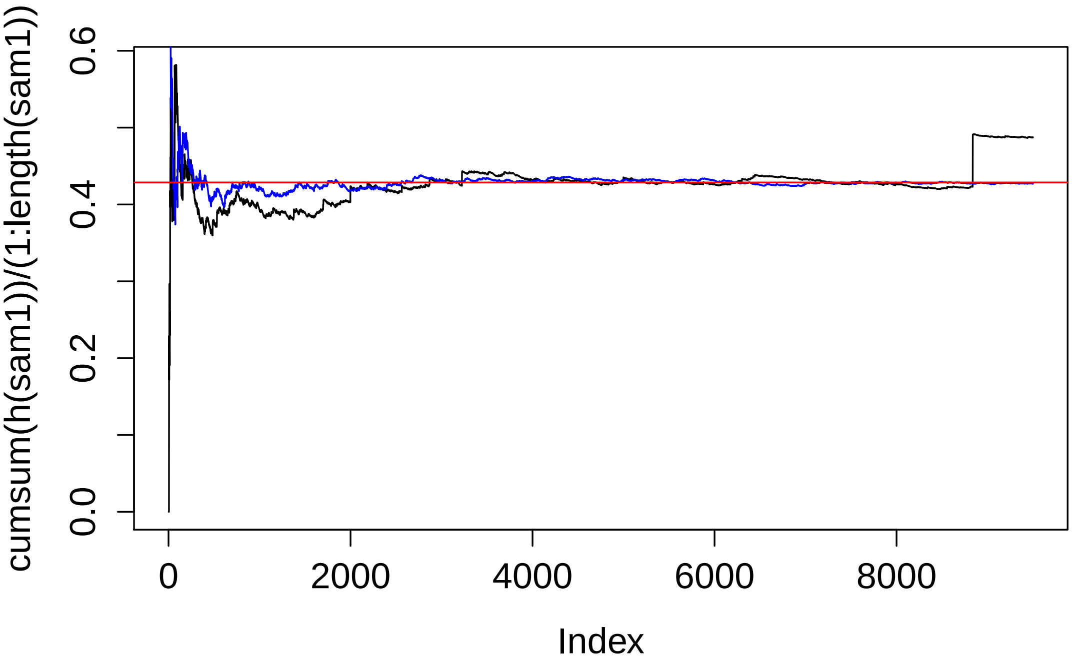
```
integrate(function(x)\{gamma(3/2)/sqrt(pi)/sqrt(x-1)/x^{1.5}\},1,Inf)
> 1 with absolute error < 2.7e-13
```

The comparison of defensive sampling with the original importance sampler thus consists in adding a small sample from ℓ to the original sample from $g = f$:

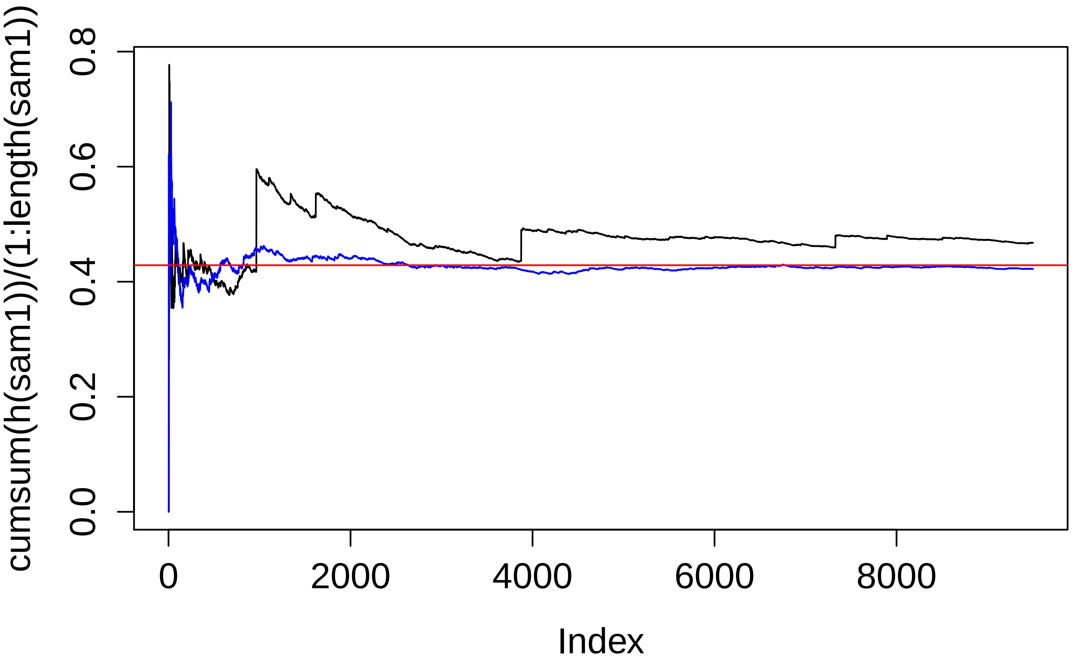
Example 3.9 [RC].

```
> h=function(x)\{z=x; z[z<1]=0; y=sqrt(z/(z-1)); y\}
> int=integrate(function(x){sqrt(x/(x-1))*dt(x,df=2)},1,Inf)\$val
> sam1=rt(.95*10^4,df=2)
> sam2=1+.5*rt(.05*10^4,df=2)^2
> sam=sample(c(sam1,sam2),.95*10^4)
> weit=dt(sam,df=2)/(0.95*dt(sam,df=2)+.05*(sam>0)*
          dt(sqrt(2*abs(sam-1)),df=2)*sqrt(2)/sqrt(abs(sam-1)))
> plot(cumsum(h(sam1))/(1:length(sam1)),ty="l")
> lines(cumsum(weit*h(sam))/1:length(sam1),col="blue")
> abline(a=int, b=0, col="red")
```

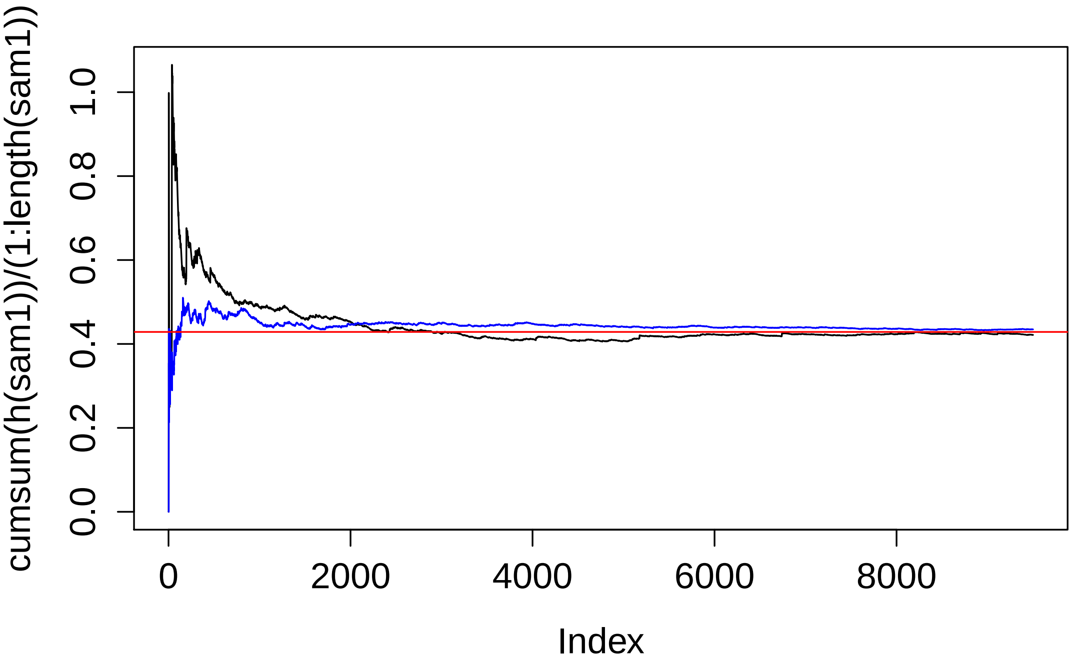
Example 3.9 [RC].



Example 3.9 [RC].



Example 3.9 [RC].



Homework:

- Doubts in Example 3.9.
- Example 3.8.
- Exercise 3.6, 3.10, 3.12

References:

[RC] Cristian P. Robert and George Casella. Introducing Monte Carlo Methods with R. Series "Use R!". Springer