STA 360: Assignment 5

Michael Lin

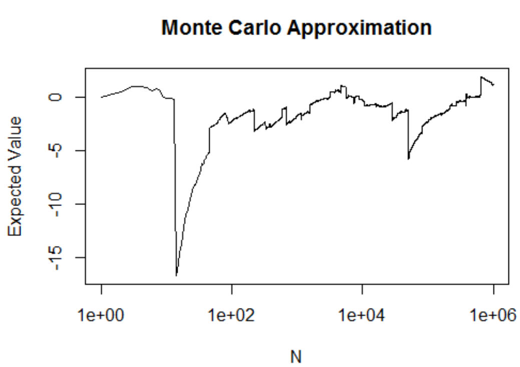
February 10, 2015

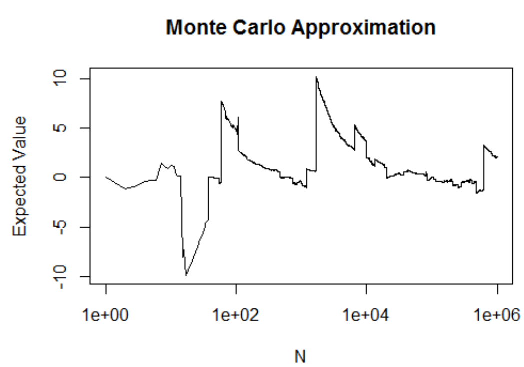
1. Let U be a random variable with Uniform(0,1) distribution. Sample a U, and suppose for this instance $U = u$. Then set u equal to the c.d.f. of the Gumbel(c, β) distribution, i.e.,

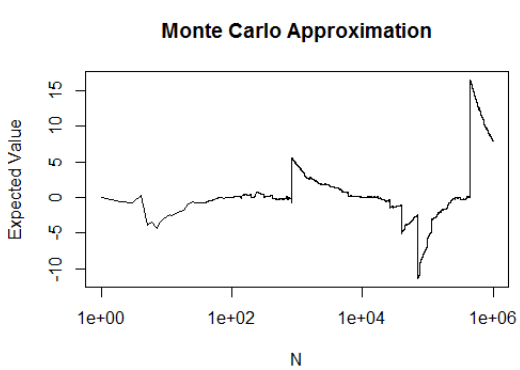
$$\begin{aligned} u &= \exp(-e^{-(x-c)/\beta}) \\ -\log u &= e^{-(x-c)/\beta} \\ \log(\log(1/u)) &= -(x-c)/\beta \\ -\beta \log(\log(1/u)) &= x - c \\ x &= c - \beta \log(\log(1/u)) \end{aligned}$$

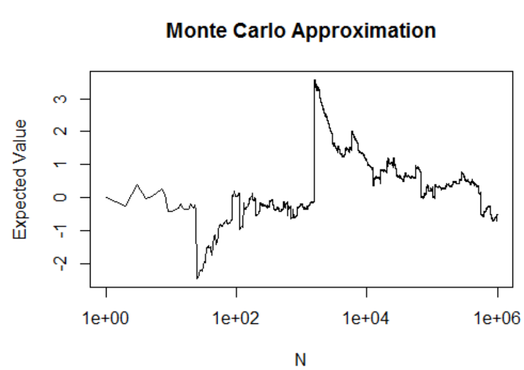
Here, $X = x$, and in general, X would be a random variable with Gumbel(c, β) distribution.

2. Each of the plot shows that the Monte Carlo approximation for the mean of the Cauchy(0,1) distribution does not converge for large N. Even though the mean may appear to converge on an interval, the convergence may be "broken" with a spike despite larger number of samples (it is not clear that the mean will ever converge). This seems to agree with that fact that $\mathbb{E}X$ does not exist for the Cauchy distribution. These plots also support the fact that Monte Carlo approximation may not always be suitable for approximation.









3. Harmonic Mean Approximation

(a) We want to show that

$$p(x_{1:n}) \approx \frac{1}{\frac{1}{N}\sum_{i=1}^{N} 1/p(x_{1:n}|\theta_i)}$$

Observe the following:

$$\begin{aligned}
\frac{1}{\frac{1}{N}\sum_{i=1}^{N} 1/p(x_{1:n}|\theta_i)} &\approx \frac{1}{\mathbb{E}[\frac{1}{p(x_{1:n}|\theta)}]} \\
&= \Big[\mathbb{E}\big[\frac{1}{p(x_{1:n}|\theta)}\big]\Big]^{-1} \\
&= \Big[\int \frac{1}{p(x_{1:n}|\theta)} p(\theta|x_{1:n})d\theta\Big]^{-1} \\
&= \Big[\int \frac{1}{p(x_{1:n}|\theta)} \frac{p(x_{1:n}|\theta)p(\theta)}{p(x_{1:n})} d\theta\Big]^{-1} \\
&= \Big[\frac{1}{p(x_{1:n})} \int p(\theta)d\theta\Big]^{-1} \\
&= \Big[\frac{1}{p(x_{1:n})}\Big]^{-1} \\
&= p(x_{1:n})
\end{aligned}$$

Thus in principle, the harmonic mean approximation converges to the marginal likelihood $p(x_{1:n})$.

(b) The harmonic mean approximation for five independent sets returned the following values:

$$\{0.11058598, 0.09973201, 0.10599124, 0.09182709, 0.10342682\}$$

The true value of the marginal likelihood is 0.03891791, and the approximations do not seem to be converging to the true value.

(c) The results are similar with a different λ_0. The true value of the marginal likelihood is 0.003988426 while the harmonic mean approximation returns the following values:

$$\{0.09475517, 0.10211462, 0.02539194, 0.09336356, 0.07787503\}$$

5. We want to show that

$$P(Z \in S) = P(X \in S | X \in A)$$

for all $S \subset A$. For $X \in \mathbb{R}^d$, let $a = P(X \notin A)$. Therefore,

$$\begin{aligned}
P(Z \in S) &= P(X_1 \in S) + P(X_1 \notin A, X_2 \in S) + P(X_1 \notin A, X_2 \notin A, X_3 \in S) + \cdots \\
&= P(X_1 \in S) + P(X_1 \notin A)P(X_2 \in S) + P(X_1 \notin A)P(X_2 \notin A)P(X_3 \in S) + \cdots \\
&= P(X \in S) + P(X \notin A)P(X \in S) + P(X \notin A)P(X \notin A)P(X \in S) + \cdots \\
&= P(X \in S) + aP(X \in S) + a^2 P(X \in S) + \cdots \\
&= P(X \in S)(1 + a + a^2 + \cdots) \\
&= P(X \in S)\sum_{k=0}^{\infty} a^k \\
&= P(X \in S)\frac{1}{1-a} \\
&= P(X \in S)\frac{1}{1 - P(X \notin A)} \\
&= P(X \in S)\frac{1}{P(X \in A)} \\
&= P(X \in S | X \in A)
\end{aligned}$$

R code for number 2:

```r
N = 10^6

samp = rcauchy(N,0,1)

samp.sum = rep(0,N)
samp.mean = rep(0,N)

samp.sum[1] = samp[1]
for(i in 2:N){
  samp.sum[i] = samp.sum[i-1] + samp[i]
  samp.mean[i] = samp.sum[i]/i
}
samp.mean[1] = samp.sum[1]

x = 1:N
plot(x,samp.mean,log = "x", type="l", lty=1, xlab="N", ylab="Expected Value",
     main="Monte Carlo Approximation")
```

R code for number 3:

```

## Harmonic Mean Function ##
# Courtesy of Radford Neal's Blog #
harmonic.mean.marg.lik <- function (x, s0, s1, n)
{ post.prec <- 1/s0^2 + 1/s1^2
  t <- rnorm(n,(x/s1^2)/post.prec,sqrt(1/post.prec))
  lik <- dnorm(x,t,s1)
  1/mean(1/lik)
}

## Define Variable ##
mu.0 = 0
lambda.0 = 1/100^2
lambda = 1

val = rep(0,5)
for(i in 1:5){
val[i]=harmonic.mean.marg.lik(2,sqrt(lambda.0^-1),sqrt(lambda),10^6)
}

print(val)
```