ST425 - Computer session 5

30/10/2020

Q 4.5 - Monte-Carlo Methods

To compute

$$J = \int_{-\infty}^{\infty} g(x)dx = \int_{-\infty}^{\infty} \exp\left(\sin(x) - (x-2)^2\right)dx$$

we can find a density function $f(x)$ of which support is $(-\infty, \infty)$, then rewrite J as

$$J = \int_{-\infty}^{\infty} g(x)dx = \int_{-\infty}^{\infty} h(x)f(x)dx$$

where $h(x) = g(x)/f(x)$. Generally, if $f(x)$ is a valid density function ($f(x) \geq 0 \forall x, \int f(x)dx = 1$), $\mathbb{E}[h(\mathrm{X})] = \int h(x)f(x)dx$. So to approximate J, we can (1) sample n independent $\mathrm{X}_1, \ldots \mathrm{X}_n$ from $f(\cdot)$, (2) compute $h(\mathrm{X}_i)$ for $i = 1, \ldots, n$ and (3) take the average, i.e.,

$$\hat{J} = \frac{1}{n}\sum_{i=1}^{n} h(\mathrm{X}_i)$$

For Q 4.5(c), by looking at the integral, we notice that one of the natural choices for $f(x)$ is $N(2, 1/2)$ (See the document for R workshop week 5 on Moodle). Then

$$J = \int_{-\infty}^{\infty} h(x)f(x)dx = \int_{-\infty}^{\infty} \sqrt{\pi}\exp\left(\sin(x)\right) \times \frac{1}{\sqrt{\pi}}\exp(-(x-2)^2)dx$$

To implement these 3 steps in R,

```
reps<-1000
X <- rnorm(reps, mean=2, sd = sqrt(1/2))
h <- sqrt(pi)*exp(sin(X))
mean(h); sd(h)/sqrt(reps)
```

```
## [1] 3.736936
```

```
## [1] 0.03198649
```

However, there are a number of ways to choose $f(x)$. For example, you can rewrite $g(x)$ as

$$g(x) = \sqrt{\pi}\exp\left(\sin(x) + 4x - 4\right) \times \frac{1}{\sqrt{\pi}}\exp\left(-x^2\right)$$

Then J is approximated by $\frac{1}{n}\sum_{i=1}^{n} \sqrt{\pi}\exp\left(\sin(\mathrm{X}_i) + 4\mathrm{X}_i - 4\right)$ where $X_i \sim N(0, \frac{1}{2})$. We can program this as follows.

```
X <- rnorm(reps, mean=0, sd = sqrt(1/2))
h <- sqrt(pi)*exp(sin(X)+4*X-4)
mean(h); sd(h)/sqrt(reps)
```

```
## [1] 14.92984
```

```
## [1] 12.73765
```

You can also sample from $N(0,1)$, or any $N(\mu,\sigma^2)$ and find appropriate $h(x)$ that corresponds to each $f(x)$. To generalise this in R, noting that $h(x) = g(x)/f(x)$, we can write

```
g <- function(x){exp(sin(x)-(x-2)^2)}
MCresult4.5c <- function(mu=0,sigma=1,reps){
  x <- rnorm(reps,mean=mu, sd = sigma)
  h <- g(x)/(dnorm(x,mean=mu,sd = sigma))
  return(list(mean(h), sd(h)/sqrt(reps)))
}
```

By running a few examples, we see that $\hat{J}$ and its S.E. vary depending on the choice of $f(x)$.

```
MCresult4.5c(mu=2, sigma=sqrt(.5), reps = reps)
```

```
## [[1]]
## [1] 3.754169
##
## [[2]]
## [1] 0.03188613
```

```
MCresult4.5c(mu=0, sigma=sqrt(.5), reps = reps)
```

```
## [[1]]
## [1] 3.23813
##
## [[2]]
## [1] 0.957977
```

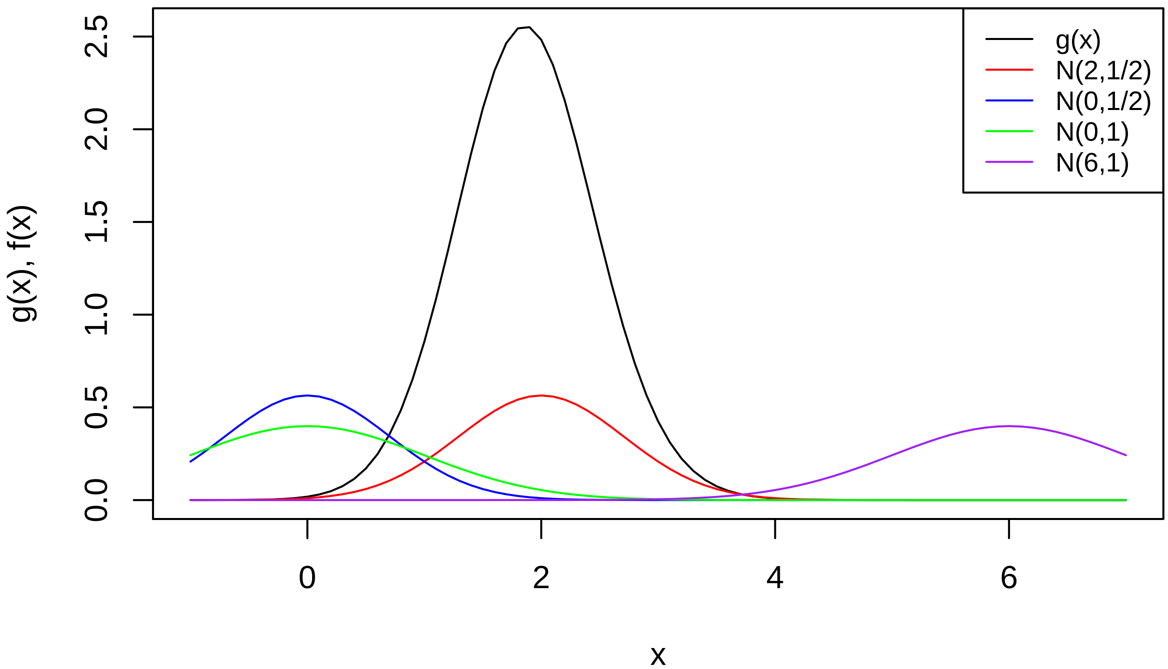
```
MCresult4.5c(mu=0, sigma=1, reps = reps)
```

```
## [[1]]
## [1] 3.84127
##
## [[2]]
## [1] 0.3769763
```

```
MCresult4.5c(mu=6, sigma=1, reps = reps)
```

```
## [[1]]
## [1] 1.474246
##
## [[2]]
## [1] 0.9006234
```

How do we know which one gives the best performance? As discussed in the seminar, we use the one that is naturally given by the problem, i.e. $N(2, \frac{1}{2})$ in this case. We can also see that by plotting $g(x)$ and candidates for $f(x)$.


g(x)
N(2,1/2)
N(0,1/2)
N(0,1)
N(6,1)
g(x), f(x)
x
0.0
0.5
1.0
1.5
2.0
2.5
0
2
4
6