Computational statistics
Homework 1

Salomé Do

October 19, 2018

Exercice 2.30

Let f, g be the properly normalized densities of the Accept-Reject algorithm.

(a) Acception probability.

We are interested in the probability p of accepting a random variable in the Accept-Reject algorithm.
A random variable X is accepted when $U < \frac{f(X)}{Mg(X)}$, so we have :

$$\begin{aligned} p &= P\left(U < \frac{f(X)}{Mg(X)}\right) \\ &= \int_x P\left(U < \frac{f(x)}{Mg(x)} \mid X = x\right) P(X = x)dx \\ &= \int_x \frac{f(x)}{Mg(x)} g(x)dx \\ &= \frac{1}{M} \int_x f(x)dx \\ &= \frac{1}{M} \end{aligned}$$

(b) Show $M \geq 1$.

$p = \frac{1}{M}$ is by definition the probability of a uniform to be on a certain interval. Thus, $0 \leq p \leq 1$ and necessarily $M \geq 1$.

(c) Average time to have k acceptations.

Let N be the number of failed trials until the k-th random variable is accepted. Let's compute $\mathbb{P}(N = n)$, for any $n \in \mathbb{N}$.
First of all, having n failed trials until k-th random variable is accepted means having done $l = n + k$ trials. Let's see the serie of trials as a vector, with 0 denoting a fail and 1 an acceptation. Our vector must look like this:

$$(\underbrace{0, 1, 0, 0, \dots, 0, 1}_{l-1, \text{ with k-1 success}}, 1)$$

Where the only fixed part is the last number, fixed to 1 (acceptation), because the process stops with the k-th acceptation.
There are $\binom{l-1}{k-1}$ combinations corresponding to the first part of the vector.

Thus, we naturally have, $\forall n \in \mathbb{N}$:

$$\begin{aligned}\mathbb{P}(N=n) &= \binom{l-1}{k-1} p^{k-1}(1-p)^{l-1-(k-1)} p \\ &= \binom{n+k-1}{k-1} p^{k}(1-p)^{n+k-1-(k-1)} \\ &= \binom{n+k-1}{k-1} p^{k}(1-p)^{n} \\ \text{So } N &\sim \mathcal{N}\text{eg}(k,p)\end{aligned}$$

We now want an estimation of the average number of failures we are going to wait to have k acceptions. Let's compute $\mathbb{E}[N]$.

We start by letting S_l be the following sum :

$$\begin{aligned}S_l &= \sum_{n=1}^{l} n\mathbb{P}(N=n) \\ &= \sum_{n=1}^{l} n\binom{n+k-1}{k-1} p^{k}(1-p)^{n} \\ &= \sum_{n=1}^{l} \frac{n(n+k-1)!}{(k-1)!n!} p^{k}(1-p)^{n} \\ &= \sum_{n=1}^{l} k\frac{(n+k-1)!}{k!(n-1)!} p^{k}(1-p)^{n} \\ &= \sum_{n=1}^{l} k\binom{n+k-1}{n-1} p^{k}(1-p)^{n} \\ &= kp^{k}(1-p)\sum_{n'=0}^{l-1}\binom{n'+k}{n'}(1-p)^{n'} \qquad (n' \leftarrow n+1)\end{aligned}$$

Using a variant of the binomial serie : $\frac{1}{(1-x)^{\beta+1}} = \sum_{k=0}^{\infty}\binom{k+\beta}{k}x^k$, we have the convergence and the limit of S_l serie when $l \to \infty$:

$$S_l \underset{l\to\infty}{\longrightarrow} kp^{k}(1-p)\frac{1}{(1-(1-p))^{k+1}}$$

Thus $\mathbb{E}[N] = k\frac{(1-p)}{p}$. However, we want the average time to wait for k acceptation, so we want :

$$\begin{aligned}\mathbb{E}[N]+k &= \frac{k}{p} \\ &= kM\end{aligned}$$

Hence the asked result.

(d) No need for a low M bound.

A tight M bound descreases the average time required to find acceptations. However, a M close to 1 is sometimes more difficult to compute and does not give efficient algorithm.

(e) Too tight bound.

If the bound is too tight (when $f(x) > Mg(x)$), the algorithm doesn't generate f.
Let's suppose we have a set $\mathcal{A}$ such as :

$$\mathcal{A} = \{x : f(x) > Mg(x)\}$$

We compute :

$$\begin{aligned}
P\left(X = x | U \leq \frac{f(X)}{Mg(X)}\right) &= P(A|B) \\
&= \frac{P(B|A)P(A)}{P(B)} \\
&= \frac{P(B|A)g(x)}{P(B)}
\end{aligned}$$

Now, we want to compute $P(B|A)$:

$$\begin{aligned}
P(B|A) &= P\left(U \leq \frac{f(X)}{Mg(X)} | X = x\right) \\
&= 1 - P\left(U < \frac{f(X)}{Mg(X)} | X = x\right) \\
&= 1 - \left[1 - \frac{f(x)}{Mg(x)}\right] 1_{\{x \notin \mathcal{A}\}} \\
&= \frac{f(x)}{Mg(x)} 1_{\{x \notin \mathcal{A}\}}
\end{aligned}$$

We now just want $P(B)$: as in question (a), we have:

$$\begin{aligned}
P(B) &= P\left(U \leq \frac{f(X)}{Mg(X)}\right) \\
&= \int_x P\left(U \leq \frac{f(x)}{Mg(x)} \mid X = x\right) P(X = x) dx \\
&= 1 - \int_x P\left(U \geq \frac{f(x)}{Mg(x)} \mid X = x\right) P(X = x) dx \\
&= 1 - \int_{x \notin \mathcal{A}} P\left(U \geq \frac{f(x)}{Mg(x)} \mid X = x\right) P(X = x) dx \\
&= \int_{x \notin \mathcal{A}} P\left(U \leq \frac{f(x)}{Mg(x)} \mid X = x\right) P(X = x) dx \\
&= \int_{x \notin \mathcal{A}} \frac{f(x)}{M} dx \\
&= \frac{1}{M}\left(1 - \int_{x \in \mathcal{A}} f(x) dx\right)
\end{aligned}$$

$\left(1-\int_{x\in\mathcal{A}} f(x)dx\right)$ is a constant, that we will call $C_{\mathcal{A}}$ Then, we finally have :

$$\begin{aligned} P\left(X = x | U \leq \frac{f(X)}{Mg(X)}\right) &= \frac{f(x)g(x)M}{Mg(x)C_{\mathcal{A}}} \times 1_{x\notin\mathcal{A}} \\ &= \frac{1}{C_{\mathcal{A}}} f(x) \times 1_{x\notin\mathcal{A}} \end{aligned}$$

Thus, the density generated by this algorithm is clearly not f.

Exercice 2.40

(a) Natural exponential family

Let $g_\theta(x) = \log(\exp\{x\theta - \psi(\theta)\})$, for any x.
Then $g_\theta(x) = x\theta - \psi(\theta))$ is affine in x and thus concave.

(b) Logistic distribution

For all $x \in \mathbb{R}$, let :

$$f(x) = \frac{1}{\beta} \frac{\exp(\frac{-(x-\alpha)}{\beta})}{(1+\exp(\frac{-(x-\alpha)}{\beta}))^2}, \beta > 0$$

Let also :

$$\begin{aligned} g(x) = \log f(x) &= \underbrace{-\log\beta - \frac{(x-\alpha)}{\beta}}_{\text{concave}} - 2\underbrace{\log(1+e^{\frac{-(x-\alpha)}{\beta}})}_{\text{h(x)}} \\ \frac{\partial h(x)}{\partial x} &= \frac{-\frac{1}{\beta}e^{-\frac{(x-\alpha)}{\beta}}}{1+e^{-\frac{(x-\alpha)}{\beta}}} \\ &= -\frac{1}{\beta}\left(1 - \frac{1}{1+e^{-\frac{(x-\alpha)}{\beta}}}\right) \\ \frac{\partial^2 h(x)}{\partial x} &= \frac{1}{\beta}\frac{\frac{1}{\beta}e^{-\frac{(x-\alpha)}{\beta}}}{\left(1+e^{-\frac{(x-\alpha)}{\beta}}\right)^2} \\ &\geq 0 \end{aligned}$$

So h is convex, then g concave.

(c) Gumbel distribution

The Gumbel distribution is the following :

$$f(x) = \frac{k^k}{(k-1)!}\exp\{-kx - ke^{-x}\}, k \in \mathbb{N}*$$

Let :

$$\begin{aligned} g(x) &= \log f(x) \\ &= \text{constant } - k(x + e^{-x}) \end{aligned}$$

As $k > 0$, and $x + e^{-x}$ is convex, then g is concave.

(d) Generalized inverse Gaussian distribution

Let $g(x) = \log f(x), \forall x \in \mathbb{R}^+$. Then :

$$g(x) = \underbrace{\text{constant } + \alpha \log x}_{\text{concave, } \alpha>0} \underbrace{-\beta x}_{\text{linear}} \underbrace{-\frac{\alpha}{x}}_{\text{concave}}$$

Exercice 3.4

(a)

Let $X \sim \mathcal{N}(0, \sigma^2)$ be a normal distributed random variable, we then have:

$$\begin{aligned} \mathbb{E}(e^{-X^2}) &= \int_{-\infty}^{+\infty} e^{-x^2} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}} dx \\ &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{+\infty} e^{-(1+\frac{1}{2\sigma^2})x^2} dx \\ &= \frac{1}{\sqrt{2\pi\sigma^2}} \times \sqrt{\frac{\pi}{1+\frac{1}{2\sigma^2}}} \text{ (Gaussian integral)} \\ &= \frac{1}{\sqrt{1+2\sigma^2}} \end{aligned}$$

(b)

Let $X \sim \mathcal{N}(0, \sigma^2)$ be a normal distributed random variable, we then have:

$$\begin{aligned} \mathbb{E}(e^{-X^2}) &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{+\infty} e^{-x^2} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \\ &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{+\infty} e^{-\frac{1+2\sigma^2}{2\sigma^2}\left[(x-\frac{\mu}{1+2\sigma^2})^2 + \frac{\mu^2}{1+2\sigma^2} - \frac{\mu^2}{(1+2\sigma^2)^2}\right]} dx \\ &= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1+2\sigma^2}{2\sigma^2}\left[\frac{\mu^2}{1+2\sigma^2} - \frac{\mu^2}{(1+2\sigma^2)^2}\right]} \int_{-\infty}^{+\infty} e^{-\frac{1+2\sigma^2}{2\sigma^2}(x-\frac{\mu}{1+2\sigma^2})^2} dx \\ &= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\left[\frac{\mu^2}{2\sigma^2} - \frac{\mu^2}{2\sigma^2(1+2\sigma^2)}\right]} \times \sqrt{\frac{\pi}{1+\frac{1}{2\sigma^2}}} \text{ (Gaussian integral)} \\ &= \frac{1}{\sqrt{1+2\sigma^2}} e^{\frac{\mu^2}{2\sigma^2(1+2\sigma^2)} - \frac{\mu^2}{2\sigma^2}} \\ &= \frac{1}{\sqrt{1+2\sigma^2}} e^{-\frac{\mu^2}{1+2\sigma^2}} \end{aligned}$$