Problem 2.6

(a). Let $U \sim \mathcal{U}_{[0,1]}$ and $X = h(U) = -\log(U)/\lambda$. The density of X is $|J_{h^{-1}}(x)|$, where $J_{h^{-1}}$ is the Jacobian of the inverse of the transformation h. Then, $X \sim \mathcal{E}xp(\lambda)$.

(b). Let (U_i) be a sequence of iid $\mathcal{U}_{[0,1]}$ random variables and

$$Y = -2\sum_{j=1}^{\nu} \log(U_j).$$

Then Y is a sum of ν $\mathcal{E}xp(/12)$ rv's, that is, of ν χ^2_2 rv's. By definition, Y is thus a chi-squared variable: $Y \sim \chi^2_{2\nu}$.

Let $Y = -\frac{1}{\beta}\sum_{j=1}^{a}\log(U_j)$. Using the previous result, $2\beta Y \sim \chi^2_{2a}$, that is $2\beta Y \sim \mathcal{G}a(a, 1/2)$. Therefore, $Y \sim \mathcal{G}a(a,\ \beta)$.

Let

$$Y = \frac{\sum_{j=1}^{a}\log(U_j)}{\sum_{j=1}^{a+b}\log(U_j)} = \frac{1}{1+Y_1/Y_2}\,.$$

where $Y_1 = -\sum_{j=1}^{a}\log(U_j)$ and $Y_2 = -\sum_{j=a+1}^{a+b}\log(U_j)$. Using the previous result with $\beta = 1$,we have $Y_1 \sim \mathcal{G}a(a,\ 1)$ and $Y_2 \sim \mathcal{G}a(b,\ 1)$. Denote by f, f_2 and f_3 the density functions of Y, Y_1 and Y_2 respectively. Using the transformation $h : (Y_1,\ Y_2) \longrightarrow (Y,\ Y_2)$, we obtain

$$\begin{aligned} f(y) &= \int_0^\infty |J_{h^{-1}}(y, y_2)| f_1(y_2(\frac{1}{y}-1)) f_2(y_2) dy_2 \\ &= \frac{(1-y)^{a-1}}{\Gamma(a)\Gamma(b)y^{a+1}} \int_0^\infty y_2^{a+b-1} e^{-\frac{y_2}{y}}\, dy_2 \\ &= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} y^{a-1}(1-y)^{b-1} = \frac{y^{a-1}(1-y)^{b-1}}{B(a,b)}. \end{aligned}$$

Therefore, Y is a $\mathcal{B}e(a, b)$ random variable.

(c). We use the fact that if $X \sim F_{m.n}$, then $\frac{nX}{m+nX} \sim \mathcal{B}e(m, n)$. Thus, we simulate $Y \sim \mathcal{B}e(m, n)$ using the previous result and take $X = \frac{mY}{n(1-Y)}$.

(d). Let $U \sim \mathcal{U}_{[0,1]}$ and $X = h(U) = \log U/1 - U$. The density of X is then

$$f(x) = |(h^{-1})'(x)| = \frac{e^{-x}}{(1+e^{-x})^2}$$

with $h^{-1}(x) = e^{-x}/1 + e^{-x}$. Then X is a Logistic$(0, 1)$ random variable. Using the fact that $\beta X + \mu$ is a Logistic$(\mu,\ \beta)$ random variable, generating a $\mathcal{U}_{[0,1]}$ variable and taking $X = \beta \log U/1 - U + \mu$ gives a Logistic$(\mu,\ \beta)$ variable.

Problem 2.13

(a). Suppose that we have a sequence of $\mathcal{E}xp(\lambda)$ rv's X_i with cdf F. Then the cdf of the sum $X_1 + \ldots + X_n$ is the convolution

$$F_{X_1+\ldots+X_n}(x) = F^{n^*}(x) = 1 - e^{x\lambda}\sum_{k=0}^{n-1}\frac{(x\lambda)^k}{k!}$$

and the probability that $X_1 + \ldots + X_n \leq 1 \leq X_1 + \ldots + X_{n+1}$ is

$$\begin{aligned} & F^{n^*}(1) - F^{(n-1)^*}(1) = 1 - e^{x\lambda}\sum_{k=0}^{n-1}\frac{(x\lambda)^k}{k!} - \left(1 - e^{\lambda}\sum_{k=0}^{n}\frac{\lambda^k}{k!}\right) \\ & = e^{-\lambda}\frac{(\lambda)^n}{n!} = P_\lambda(N = k)\,. \end{aligned}$$

(b). We take $X = N-1$ that means $\prod_{i=1}^{N} u_i < c$ and $\prod_{i=1}^{N-1} u_i \geq c$. So, we have $-\frac{1}{\lambda}\sum_{i=1}^{N}\log u_i > 1$ and $-\frac{1}{\lambda}\sum_{i=1}^{N-1}\log u_i \leq 1$ which is equivalent to (a) as $-\frac{1}{\lambda}\log u_i \sim \text{Exp}(\lambda)$

Problem 2.17

(a). To find the distribution of $U_{(i)}$ over all possibilities of ordering $U_1, ..., U_n$ into $U_{(i)} \leq U_{(2)} \leq, ..., \leq U_{(n)}$, let $\mathfrak{S}_n$ be the set of the permutations σ of $\{1, ..., n\}$. (It contains $n!$ elements.) The cdf of $U_{(i)}$ is

$$\begin{aligned} F_{U_{(i)}}(u) &= P(U_{(i)} \leq u) \\ &= \sum_{\sigma\in\mathfrak{S}_n}[\int_0^u(\int_0^{u_{\sigma(2)}} du_{\sigma(1)}(\int_0^{u_{\sigma(3)}} du_{\sigma(2)}(...(\int_0^{u_{\sigma(i)}} du_{\sigma(i-1)})\ldots))) \\ &\quad \times(\int_{u_{\sigma(i)}}^1 du_{\sigma(+1)}(\int_{u_{\sigma(i+1)}}^1 du_{\sigma(i+2)}(\ldots\int_{u_{\sigma(n-1)}}^1 du_{\sigma(n)}))\ldots)du_{\sigma(i)}] \\ &= \sum_{\sigma\in\mathfrak{S}_n}\int_0^u \frac{u_{(i)}^{i-1}}{(i-1)!}\frac{(1-u_{(i)})^{n-i}}{(n-i)!}du_{\sigma(i)} \\ &= \int_0^u \frac{n!}{(i-1)!(n-i)!}x^{(i-1)}(1-x)^{n-i}dx \\ &= \int_0^u \frac{x^{(i-1)}(1-x)^{n-i}}{B(i, n-i+1)}dx. \end{aligned}$$

Therefore, $U_{(i)}$ is distributed as a $\mathcal{B}e(i, n-i+1)$ variable.

(b). The same method used for $(U_{(i_1)}, U_{(i_2)} - U_{(i_1)}, ..., U_{(i_k)} - U_{(i_{k-1})}, 1 - U_{(i_k)})$ gives

$$\begin{aligned} F(u_1, u_2, ..., u_k) &= \sum_{\sigma \in S_n} [\prod_{j=1}^{k} (\int_0^{u_j} (\int_0^{u_{\sigma(i_{j-1}+2)}} du_{\sigma(i_{j-1}+1)} \\ &\quad (... \int_0^{u_{\sigma(i_j)}} du_{\sigma(i_j-1)})...)du_{\sigma(j)})]\mathbb{I}_{\{\sum_{j=1}^k u_{\sigma(i_j)}=1\}} \\ &= \sum_{\sigma \in S_n} \int_{[0,u_1]\times...\times[0,u_k]} \mathbb{I}_{\{\sum_{j=1}^k u_{i_j}=1\}} \prod_{j=1}^{k} \frac{u_{i_j}^{i_j-1}}{(i_j-1)!} du_{i_1}...du_{i_k} \\ &= \frac{n!}{(i_1-i_0-1)!...(i_{k+1}-i_k-1)!} \\ &\times \int_{[0,u_1]\times...\times[0,u_k]} [u_{i_1}^{i_1-i_0-1}...u_{i_k}^{i_{k+1}-i_k-1}\mathbb{I}_{\{\sum_{j=1}^k u_{i_j}=1\}}]du_{i_1}...du_{i_k} \\ &= \frac{\Gamma(n+1)}{\Gamma(i_1)\Gamma(i_2-i_1)...\Gamma(n-i_k)} \\ &\times \int_{[0,u_1]\times...\times[0,u_k]} [u_{i_1}^{i_1-i_0-1}...u_{i_k}^{i_{k+1}-i_k-1}\mathbb{I}_{\{\sum_{j=1}^k u_{i_j}=1\}}]du_{i_1}...du_{i_k} \end{aligned}$$

where $i_0 = 0$ and $i_{k+1} = n$ for the sake of simplicity. Therefore, $(U_{(i_1)}, U_{(i_2)} - U_{(i_1)}, ..., U_{(i_k)} - U_{(i_{k-1})}, 1 - U_{(i_k)})$ is distributed as a Dirichlet $\mathcal{D}_{k+1}(i_1, i_2 - i_1, ..., n - i_k + 1)$.

(c). Let U and V be iid random variables from $\mathcal{U}_{[0,1]}$ and let

$$X = \frac{U^{1/\alpha}}{U^{1/\alpha} + V^{1/\beta}}.$$

Consider the transform

$$h : (u, v) \longrightarrow \left(x = \frac{u^{1/\alpha}}{u^{1/\alpha} + v^{1/\beta}}, y = v \right)$$

The distribution of X conditional on $U^{1/\alpha} + V^{1/\beta} \leq 1$ is

$$\begin{aligned} P(X \leq x | U^{1/\alpha} + V^{1/\beta} \leq 1) &= \frac{\int_0^x (\int_0^{(1-z)^\beta} |J_{h^{-1}}(z, y)| dy) dz}{P(U^{1/\alpha} + V^{1/\beta} \leq 1)} \\ &= \frac{\int_0^x (\int_0^{(1-z)^\beta} \alpha y^{\alpha/\beta} z^{\alpha-1} (1-z)^{-\alpha-1} dy) dz}{\int_0^1 (\int_0^{(1-v^{1/\beta})^\alpha} du) dv} \\ &= \frac{\int_0^x (\beta(1-z)^{\alpha+\beta}(\alpha+\beta)^{-1} \alpha z^{\alpha-1} (1-z)^{-\alpha-1} dz)}{\int_0^1 (1-v^{1/\beta})^\alpha dv} \end{aligned}$$

For computing $\int_0^1 (1 - v^{1/\beta})^\alpha dv$, we use the change of variable $v = t^\beta$. Then $dv = \beta t^{\beta-1} dt$ and

$$\int_0^1 (1 - v^{1/\beta})^\alpha dv = \beta \int_0^1 t^{\beta-1}(1-t)^\alpha dt = \beta B(\alpha+1, \beta) = \frac{\alpha\beta\Gamma(\alpha)\Gamma(\beta)}{(\alpha+\beta)\Gamma(\alpha+\beta)}.$$

Therefore,

$$P(X \le x | U^{1/\alpha} + V^{1/\beta} \le 1) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)} \int_0^x z^{\alpha-1}(1-z)^{\beta-1} dz.$$

And X is a $\mathcal{B}e(\alpha, \beta)$ random variable.

(d). The Renyi representation of

$$u_{(i)} = \frac{\sum_{j=1}^{i} \nu_j}{\sum_{j=1}^{n} \nu_j},$$

where the ν_j's are iid $\sim \mathcal{E}xp(1)$ is the same as the representation of

$$Y = \frac{\sum_{j=1}^{a} \log{(U_j)}}{\sum_{j=1}^{a+b} \log{(U_j)}}$$

with $a = i$, $b = n - i$ and $\nu_j = -\log(U_j)$. Problem **??** implies that Y is a $\mathcal{B}e(a, b)$ rv and $U_{(i)}$ is $\mathcal{B}e(i, n-i)$ distributed. Note that the $U_{(i)}$'s are ordered, $U_{(1)} \le U_{(2)} \le ... \le U_{(n)}$. Thus, the Renyi representation gives the order statistics for a uniform sample.

Problem 2.19

By the acceptance-rejection algorithm, let X be the outcome random variable of this algorithm. Then,

$$\begin{aligned} P(X \le x | U \le \alpha) &= \frac{P(U \le x, U \le \alpha)}{P(U \le \alpha)} \\ &= x/\alpha \end{aligned}$$

Hence $x \sim \mathcal{U}(0, \alpha)$.

Probability of acceptance is $P(U < \alpha) = \alpha$. For αU method, probability of acceptance is 1. For this method, E(No. of trials)$= 1/\alpha$ (close to 1 for α close to 1 and very big for α close to 0). So αU method is more efficient when α is close to 0.

Problem 2.25a

(a) From the distribution of the noncentral chi-squared distribution,(textbook, p582), if $x_i \sim \mathcal{N}(\theta_i, 1)$, then noncentral chi-squared random variable is $\sum_i^p x_i^2 \sim \chi_p^2(\lambda)$, where $\sum_i^p \theta_i^2 = \lambda$. Then, if we have a chi-squared random variable with degree of freedom $p-1$, $x_1 \sim \chi^2_{p-1}$, and a standard normal random variable, $x_2 \sim \mathcal{N}(0,1)$, let $x = x_1 + (x_2+k)^2$, and k is constant. Then $x = \sum_i^p y_i^2$, where $y^p \sim \mathcal{N}(k,1)$, $y_i \sim \mathcal{N}(0,1)$, $(1 \leq i \leq p-1)$ and $\sum_i^p E^2(y_i) = k^2$. If we set $k^2 = \lambda$, by the definition, we have $x \sim \chi^2_p(\lambda)$.

Problem 2.29

(a) We have Target density: $\pi(\theta|x)$, and Candidate density, $\pi(\theta)$. In A-R algorithm, bound M is,

$$M = \underset{\theta}{Sup}\frac{\pi(\theta|x)}{\pi(\theta)}$$

$$\pi(\theta|x) = \frac{\pi(\theta)f(x|\theta)}{C}$$

Where C is normalizing constant(we can ignore it). Hence,

$$M = \underset{\theta}{Sup}f(x|\theta) = \underset{\theta}{Sup}L(\theta|x)$$

M is determined by $\underset{\theta}{Sup}L(\theta|x)$, which is $L(\hat{\theta}^{MLE}|x)$, by the definition of MLE.$\square$

(b) Let $y \sim g(y)$, $\theta \sim L(\theta|\mathbf{x})$. Then the candidate density is $Cauchy(0,1)$,

$$g(\theta) = \frac{1}{(1-\theta)^2\pi}$$

And the target density is $\mathcal{N}(0,1) \cdot Cauchy(0,1)$,

$$g(\theta)f(\theta|\mathbf{x}) = (2\pi)^{-n/2}exp\left(-\frac{1}{2}\sum_i^n(x_i-\theta)^2\right)\frac{1}{(1-\theta)^2\pi}$$

from a), $M = L(\hat{\theta}^{MLE}|x)$, where $\hat{\theta}^{MLE} = \frac{1}{n}\sum_i^n x_i$,

$$M = \prod_i^n \frac{1}{\sqrt{2\pi}}exp\left(-\frac{(x_i-\bar{x})^2}{2}\right) = (2\pi)^{-n/2}exp\left(-\frac{1}{2}\sum_i^n(x_i-\bar{x})^2\right)$$

Algorithm is following,

1. Generate $y \sim g$, $u \sim \mathcal{U}(0,1)$.
2. Accept $\theta = y$ if $u \leq \frac{L(\theta|\mathbf{x})}{M} = \frac{exp\left(-\frac{1}{2}\sum_i^n(x_i-y)^2\right)}{exp\left(-\frac{1}{2}\sum_i^n(x_i-\bar{x})^2\right)}$.
3. Otherwise return to 1.

Problem 2.31

First, we determine the ratio of the target density to the instrumental density:

$$\frac{f}{g_\lambda} = \frac{1}{\Gamma(n)} x^{n-1} \lambda^{-1} e^{(\lambda-1)x} \,,$$

where $1/\Gamma(n)$ is a normalising constant, f is the density of the $\mathcal{G}a(n,1)$ distribution, and g_λ is the density of the $\mathcal{E}xp(\lambda)$ distribution. Then we find the bound for the ratio, by differentiating w.r.t. x :

$$\frac{\partial}{\partial x}\frac{f}{g_\lambda} = (n-1)x_0^{n-2}\lambda^{-1}e^{(\lambda-1)x_0} + x_0^{n-1}\lambda^{-1}e^{(\lambda-1)x_0}(\lambda-1) = 0\,,$$

that is,

$$x_0 = \frac{n-1}{1-\lambda}$$

for $n > 1$ and $\lambda < 1$. Also, the second order derivative reduces to

$$\frac{\left(\frac{n-1}{1-\lambda}\right)^n (\lambda-1)^3}{(n-1)^2} \leq 0\,.$$

We thus insert the solution for x in the ratio to determine the bound M:

$$M = \frac{1}{\Gamma(n)}\left(\frac{n-1}{(1-\lambda)e}\right)^{n-1}\lambda^{-1}$$

Finally, to minimize the bound, we maximize $\lambda(1-\lambda)^{n-1}$ yielding the solution $\lambda = 1/n$. Inserting the solution for λ in the bound yields

$$M = \frac{\left(\frac{n}{e}\right)^{(n-1)} n}{\Gamma(n)}$$

and so we can plot the bound as a function of n.

Problem 2.35

(a)

$$
\begin{aligned}
P(Z \le z) &= \sum_i^{\infty} P(Z \le z, Z = x_i) \\
&= \sum_i^{\infty} P\left(Z \le z, i^{th} \text{ time}, u_i \le \epsilon_i \frac{f(x_i)}{g(x_i)} \text{ satisfied.}\right) \\
&= \sum_i^{\infty} P\left(Z \le z \middle| i^{th} \text{ time}, u_i \le \epsilon_i \frac{f(x_i)}{g(x_i)} \text{ satisfied.}\right) \\
&\cdot \quad P\left(i^{th} \text{ time}, u_i \le \epsilon_i \frac{f(x_i)}{g(x_i)} \text{ satisfied, and } 1 \le j \le i-1, u_i > \epsilon_i \frac{f(x_i)}{g(x_i)}\right) \\
&= \sum_i^{\infty} P\left(Z \le z, u_i \le \epsilon_i \frac{f(x_i)}{g(x_i)}\right) \prod_j^{i-1} P\left(u_j > \epsilon_j \frac{f(x_j)}{g(x_j)}\right) \\
&= \sum_i^{\infty} \int_{-\infty}^{z} \int_0^{\epsilon_i \frac{f(x_i)}{g(x_i)}} du\, g_i(x_i) dx_i \cdot \prod_j^{i-1} \int_{-\infty}^{\infty} \int_{\epsilon_i \frac{f(x_i)}{g(x_i)}}^{1} du\, g_i(x_i) dx_i \\
&= \sum_i^{\infty} \int_{-\infty}^{z} \epsilon_i f(x_i) dx_i \prod_j^{i-1} \int_{-\infty}^{\infty} (g(x_j) - \epsilon_j f(x_j)) dx_j \\
&= \sum_i^{\infty} \epsilon_i \prod_j^{i-1} (1 - \epsilon_j) \int_{-\infty}^{z} f(x) dx \Box \qquad (x_i = x \text{ is iid.})
\end{aligned}
$$

(b)

Let us change the variable $a_i = \epsilon_i \prod_j^i (1 - \epsilon_j)$,

$$
\begin{aligned}
a_1 &= \epsilon_1 \\
a_2 &= \epsilon_2(1 - \epsilon_1) = \epsilon_2(1 - a_1) \\
a_3 &= \epsilon_3(1 - \epsilon_2)(1 - \epsilon_1) = \epsilon_2(1 - a_1 - a_2) \\
&\cdots \\
\text{Suppose } a_n &= \epsilon_n(1 - a_1 - a_2 \cdots - a_{n-1}) \\
\text{By mathematical Induction, } a_{n+1} &= \epsilon_{n+1} \prod_i^{n-1} (1 - \epsilon_i) - \epsilon_{n+1}\epsilon_n \prod_i^{n} (1 - \epsilon_i) \\
&= \epsilon_{n+1}(1 - a_1 - a_2 - \cdots - a_n)
\end{aligned}
$$

Then,

$$\begin{aligned}\sum_i^{\infty} log(1-\epsilon_i) &= \lim_{n\to\infty} log \prod_i^n (1-\epsilon_i)\\ &= \lim_{n\to\infty} log \left((1-a_1)\left(\frac{1-a_1-a_2}{1-a_1}\right)\cdots\left(\frac{1-\sum_i^n a_i}{1-\sum_i^{n-1} a_i}\right)\right)\\ &= log\left(1-\sum_i^{\infty} a_i\right)\end{aligned}$$

Therefore, if $\sum_i^{\infty}\epsilon_i\prod_j^i(1-\epsilon_j)=\sum_i^{\infty}a_i=1$, then $log(0)=-\infty$, hence $\sum_i^{\infty}log(1-\epsilon_i)$ diverges.$\square$
If $\sum_i^{\infty}\epsilon_i\prod_j^i(1-\epsilon_j)=1$, then $P(Z\le z)=\int_{-\infty}^{z}f(x)dx$, thus we have a valid algorithm if condition from (b) satisfied.

(c)

- *Satisfying example*: Let $a_i=P(x=i-1)=\frac{e^{-\lambda}\lambda^{i-1}}{(i-1)!}$ where $x\sim\mathcal{P}(\lambda)$, then $\sum_i^{\infty}\epsilon_i\prod_j^i(1-\epsilon_j)=1$, where $\epsilon_i=\frac{a_i}{1-\sum_j^{i-1}a_j}$.

- *Not satisfying example*: Let $a_i=\frac{1}{2}(\frac{1}{2})^i$ then $\epsilon_i=\frac{\frac{1}{2^i}}{1+\frac{1}{2^i}}$, and $\sum_i^{\infty}\epsilon_i\prod_j^i(1-\epsilon_j)=\frac{1}{2}$, hence not satisfied.

Problem 2.40

(a). The density of a natural exponential family with respect to a measure ν is

$$f_1(x)=\exp\{x\cdot\theta-\psi(\theta)\}$$

The logarithm of the density is

$$\log f_1(x)=x\cdot\theta-\psi(\theta)$$

which is linear in x and hence concave (but not *strictly* concave). Thus natural exponential families have log-concave densities.

(b). The density of the logistic distribution is

$$f_2(x)=\frac{1}{\beta}+\frac{2}{\beta}\frac{\exp\{-(x-a)/\beta\}}{1+\exp\{-(x-\alpha)/\beta\}},$$

hence

$$\frac{d^2}{dx^2}\log f_2(x) = -\frac{2}{\beta^2}\frac{\exp(-(x-\alpha)/\beta)}{(1+\exp(-(x-\alpha)/\beta))^2}$$

which is negative and hence the logistic density is log-concave.

(c). The density of the Gumbel distribution is

$$f_3(x) = \frac{k^k}{(k-1)!}\exp\left\{-k\,x - k\,e^{-x}\right\}$$

where the parameter k is a natural number. So the logarithm of the density is

$$\log f_3(x) = k\log k - \log\{(k-1)!\} - kx - xe^{-x}\,,$$

then

$$\frac{d}{dx}\log f_3(x) = -k + ke^{-x}\,,$$

hence

$$\frac{d^2}{dx^2}\log f_3(x) = -ke^{-x}$$

which is negative and hence the logistic density is log-concave.

(d). The density of the generalized inverse Gaussian distribution is

$$f_4(x) \propto x^{\alpha}e^{-\beta x - \alpha/x}\,,$$

so the logarithm of the density is

$$\log f_4(x) = C + \alpha\log x - \beta x - \alpha/x\,.$$

Then

$$\frac{d}{dx}\log f_3(x) = \alpha/x - \beta + \alpha/x^2\,,$$

hence

$$\frac{d^2}{dx^2}\log f_3(x) = -\alpha/x^2 - 2\alpha/x^2$$

which is negative and the generalized inverse Gaussian density is log-concave.

Problem 3.1

(a) To plot the integrand of the numerator, we can use R command persp(), define

```
f<-function(x,theta){
  theta*exp(-(x-theta)^2/2)/(1+theta^2)
  }
```

and obtain Figure . Now, the simulation of δ can be obtained either by extracting the Cauchy density from both integrals or by extracting the normal density from both integrals, since both these densities appear in the integrands. The first approach leads to the following R code:

```
N=10^4 cau=rt(N,1) delta<-function(x){
  mean(cau*exp(-(x-cau)^2/2))/mean(exp(-(x-cau)^2/2))
  }
x=seq(-10,10,le=100)
plot(x,apply(as.matrix(x),1,delta),type="l",col="sienna",
  xlab="x",ylab=expression(delta(x)),lwd=2)
```

while the second approach uses

```
N=10^4 cau=rnorm(N) delta<-function(x){
  the=cau+x
  mean(the/(1+the^2))/mean(1/(1+the^2))
  }
lines(x,apply(as.matrix(x),1,delta),col="steelblue",
  lwd=2,lty=3)
```

and both give identical approximations to $\delta(x)$.

(b) If we use the representation of the approximation of $\delta(x)$ as $\bar{f}_n/\bar{g}_n$, with $\bar{f}_n = (f(x_1) + \cdots + f(x_n))/n$ [converging to $\mathfrak{f}$] and $\bar{g}_n = (g(x_1) + \cdots + g(x_n))/n$ converging to $\mathfrak{g}$], the normal approximation of the distributions of $\bar{f}_n$ and of $\bar{g}_n$ leads to the decomposition

$$\begin{aligned}\frac{f_n}{g_n} - \delta(x) &= \frac{\bar{f}_n - \mathfrak{f}}{\bar{g}_n} + \frac{f}{\bar{g}_n g}(\mathfrak{g} - \bar{g}_n) \\ &= \frac{\bar{f}_n - \mathfrak{f}}{\sigma_f/\sqrt{n}}\,\frac{\sigma_f}{\sqrt{n}\bar{g}_n} + \frac{\mathfrak{f}\sigma_g}{\sqrt{n}\bar{g}_n\mathfrak{g}}\,\frac{\mathfrak{g} - \bar{g}_n}{\sigma_g/\sqrt{n}} \\ &\equiv \frac{\sigma_f}{\sqrt{n}\mathfrak{g}}\,\mathcal{N}(0,1) + \frac{\mathfrak{f}\sigma_g}{\sqrt{n}\mathfrak{g}^2}\,\mathcal{N}(0,1)\end{aligned}$$

and therefore, if we overlook the dependences between both terms, we can approximate the variance of the average by

$$\frac{\sigma_f^2}{n\mathfrak{g}^2} + \frac{\mathfrak{f}^2\sigma_g^2}{n\mathfrak{g}^4}\,,$$

itself estimated by

$$\frac{\hat{\sigma}_f^2}{n\bar{g_n}^2} + \frac{\bar{f_n}^2\sigma_g^2}{n\bar{g_n}^4}\,.$$

The corresponding R code is given by

```
x=rnorm(1,0,2) N=10^4 #Cauchy sample cau=rt(N,1)
fbar=cumsum(cau*exp(-(x-cau)^2/2))/(1:N)
sigf=(cumsum(cau^2*exp(-(x-cau)^2))/(1:N)) - fbar^2
gbar=cumsum(exp(-(x-cau)^2/2))/(1:N)
sigg=(cumsum(exp(-(x-cau)^2))/(1:N)) - gbar^2
glob=sqrt(((sigf/gbar^2)+(sigg*fbar^2/gbar^4))/(1:N))
plot(fbar/gbar,type="l",xlab="n",ylab=expression(delta[n](x)),
  col="sienna3",lwd=2)
lines(fbar/gbar-2*glob,col="sienna3",lty=2)
lines(fbar/gbar+2*glob,col="sienna3",lty=2) #Normal sample
the=rnorm(N,x) fbar=cumsum(the/(1+the^2))/(1:N)
sigf=(cumsum(the^2/(1+the^2)^2))/(1:N) - fbar^2
gbar=cumsum(1/(1+the^2))/(1:N) sigg=(cumsum(1/(1+the^2)^2))/(1:N) -
gbar^2 glob=sqrt(((sigf/gbar^2)+(sigg*fbar^2/gbar^4))/(1:N))
lines(fbar/gbar,type="l",xlab="n",ylab=expression(delta[n](x)),
  col="steelblue4",lwd=2)
lines(fbar/gbar-2*glob,col="steelblue4",lty=2)
lines(fbar/gbar+2*glob,col="steelblue4",lty=2)
```

we can find that the sample based on the normal density has a smaller variability than the one based on the Cauchy density.

Problem 3.3

.1 (a)

- **Algorithm:** Find $P(z > 2.5)$ based on indicator function:

1. Generate $z \sim \mathcal{N}(0,1)$
2. Assign
$$I_i = \begin{cases} 1, & \text{if } z > 2.5 \\ 0, & \text{otherwise} \end{cases}$$
3. Find $\hat{\Phi}(2.5) = \frac{1}{n}\sum_i^n I_i$, and $Var(\hat{\Phi}(2.5)) = \hat{\Phi}(2.5)(1-\hat{\Phi}(2.5))/n$.
4. Find accuracy, $z_{0.975}\sqrt{Var(\hat{\Phi}(2.5))} - z_{0.025}\sqrt{Var(\hat{\Phi}(2.5))}$ for 95% probability.
5. Repeat until the accuracy becomes less than 0.001.

- **Simulation:** Up to the 189000 sample size, we have accuracy less than 0.001, and approximating probability is 0.006188172.

.2 (b)

- **Algorithm:** Show $P(x > 5.3) \approx 0.005$ based on indicator function:

1. Generate $x \sim \mathcal{G}(1,1)$
2. Assign
$$I_i = \begin{cases} 1, & \text{if } x > 5.3 \\ 0, & \text{otherwise} \end{cases}$$
3. Find $\hat{\gamma}(5.3) = \frac{1}{n}\sum_i^n I_i$, and $Var(\hat{\gamma}(5.3)) = \hat{\gamma}(5.3)(1-\hat{\gamma}(5.3))/n$.
4. Find accuracy, $\gamma_{0.9975}\sqrt{Var(\hat{\gamma}(5.3))} - \gamma_{0.0025}\sqrt{Var(\hat{\gamma}(5.3))}$ for 99.5% probability.
5. Repeat until the accuracy becomes less than 0.001.

- **Simulation:** When we simulate 177000 sample sizes, we have three-digit accuracy with 99.5%cut-off. and the mean value is 0.004909605, almost the same as 0.005, hence proved.

Problem 3.14

From the four instrumental densities where $0 < x < \infty$,

- Simulating $x \sim g_1$
  $g_1(x) \sim Double\ Exponential(0,1)$,we can change $g_1(x) \sim \mathcal{E}(1)$.

- Simulating $x \sim g_2$
  $g_2(x)$ is not any known density, so we generate random variables from A-R algorithm. With candidate density $\mathcal{E}(\sqrt{2})$, we find $C = \sup_x \frac{f(x)}{g_2(x)} = \frac{\sqrt{2}e^{x^*}}{e^{\sqrt{2}x^*}+e^{-\sqrt{2}x^*}+2}$, and $x^* = \frac{log(3+2\sqrt{2})}{\sqrt{2}}$. As we see, this distribution is asymmetric around zero, so the normalizing constant for truncating only positive part is just $1/2$.

- Simulating $x \sim g_3,\ g_4$
  We find truncate negative of $g_3(x) \sim \mathcal{C}(0,2)$, and $g_4(x) \sim \mathcal{N}(0,1)$ with naive algorithm, which just takes positive values from original density. Because we can expect that algorithm retrieves 50% from whole generation from original, it is not computationally bad. As the same as $g_2(x)$, the normalizing constant for truncation is $1/2$.

Then we can find sample standard deviations from four instrumental densities by,

$$\mathbb{E}_f(x) \approx \frac{1}{m}\sum_i^m x_i \frac{f(x_i)}{g(x_i)}$$

and

$$\mathbb{V}_f(x) \approx \frac{1}{m^2}\sum_i^m \left[x_i \frac{f(x_i)}{g(x_i)} - \frac{1}{m}\sum_i^m x_i \frac{f(x_i)}{g(x_i)}\right]^2$$

Where $x \sim g$.

From above formulas, we can estimate the size of M needed to obtain three digits of accuracy in estimating $E_f X$.