1. (a) [5 marks] (B/S) We have
$$\int_0^y x^\alpha \mathrm{d}x = \frac{y^{\alpha+1}}{\alpha+1},$$
so $Z_p = 2^{\alpha+1}/(\alpha+1)$ and the cdf $F(x) = \frac{x^{\alpha+1}}{2^{\alpha+1}}$ for $0 \leqslant x \leqslant 2$. The inverse is $F^{-1}(u) = 2u^{1/(\alpha+1)}$ so
$$X = F^{-1}(U) = 2U^{1/(\alpha+1)}.$$

(b) [15 marks] (i) [5 marks] (S) We have
$$M = \sup_{0\leqslant x\leqslant 2} \frac{p(x)}{q(x)} = \frac{2^\alpha/Z_\alpha}{1/2} = \alpha+1,$$
so the acceptance probability function
$$h(y) = \frac{p(y)}{Mq(y)} = \frac{y^\alpha}{2^\alpha}.$$
The algorithm is: (1) Simulate $Y \sim U(0,2)$ and $U \sim U(0,1)$; (2) if $U \leqslant h(Y)$ return $X = Y$ and otherwise goto (1).

(ii) [6 marks] (S)
```
rpalpha <- function(alpha, n=1) {
  #simulates n samples X~x^alpha
  X <- numeric(n)
  for (i in 1:n) {
    FINISHED=FALSE
    while (!FINISHED) {
      y <- 2*runif(1)
      u <- runif(1)
      FINISHED <- (u <= (y^alpha/2^alpha))
    }
    X[i] <- y
  }
  return(X)
}
```
(iii) [4 marks] (S but a little harder) If $U \sim U(0,1)$ and $Y \sim U(0,2)$, this is
$$M = \mathbb{P}\big(U \leqslant h(Y)\big) = \frac{1}{2}\int_0^2 h(y)\mathrm{d}y = \frac{1}{2^{\alpha+1}}\int_0^2 x^\alpha \mathrm{d}y = \frac{1}{\alpha+1}.$$
Since the number of trials, N say, has a geometric distribution with success probability ξ, we have $\mathbb{E}(N) = 1/\xi = \alpha+1$. [Agrees with the general result $1/M$ which they know.]

(c) [5 marks] (N) Let $x \in [0,1]$, and calculate
$$\mathbb{P}\big(Y \leqslant x, U \leqslant 2Y\big) = \int_0^x \left[\int_0^1 1(u \leqslant 2y)\mathrm{d}u\right]\mathrm{d}y = \int_0^x \min\{1, 2y\}\mathrm{d}y.$$
Also,
$$\mathbb{P}\big(U \leqslant 2Y\big) = \int_0^{1/2} 2y\mathrm{d}y + \int_{1/2}^1 1\mathrm{d}y = \frac{3}{4}.$$

We obtain

$$\begin{aligned}\mathbb{P}(X \leqslant x) &= \mathbb{P}\big(Y \leqslant x \mid U \leqslant 2Y\big) \\ &= \frac{\mathbb{P}\big(Y \leqslant x, U \leqslant 2Y\big)}{\mathbb{P}\big(U \leqslant 2Y\big)} \\ &= \frac{4}{3}\begin{cases} x^2, & 0 \leqslant x \leqslant 1/2 \\ 1/4 + (x - 1/2), & 1/2 \leqslant x \leqslant 1 \end{cases} \\ &= \begin{cases} \frac{4}{3}x^2, & 0 \leqslant x \leqslant 1/2 \\ \frac{1}{3}(4x - 1), & 1/2 \leqslant x \leqslant 1. \end{cases}\end{aligned}$$

This is the conditional probability of X being accepted in each try, and the other tries are independent. [We went through the rejection sampling proof even more formally in the lectures, and more detailed explanation is welcome but not mandatory.]

2. (a) [6 marks] (B) Draw $Y_1, \ldots, Y_n \sim q$ and

$$I_n^{(q)}(f) = \frac{1}{n}\sum_{k=1}^{n} f(Y_k)w(Y_k), \qquad w(y) = \frac{p(y)}{q(y)} \text{ if } q(y) > 0 \text{ and } 0 \text{ otherwise.}$$

The unbiasedness comes from

$$\mathbb{E}[I_n^{(q)}(f)] = \mathbb{E}[f(Y_1)w(Y_1)] = \sum_{x\in\mathbb{X}:q(y)>0} f(y)\frac{p(y)}{q(y)}q(y) = \sum_{x\in\mathbb{X}:q(y)>0} f(y)p(y) = \mathbb{E}_p[f(X)],$$

because if $q(y) = 0$ then $p(y) = 0$. [This must be explained well in a perfect answer.]

(b) [12 marks] (S)

(i) [6 marks] Set $X_0 = 1$. The state X_k for $k = 2, 3, \ldots$ is determined as follows: (1) simulate $Y_k \sim U\{1, 2, ..., m\}$ and $U_k \sim U(0, 1)$; (2) if $U_k < \alpha(Y_k \mid X_k)$ set $X_k = Y_k$ and otherwise set $X_k = X_{k-1}$. In this algorithm

$$\begin{aligned}\alpha(y \mid x) &= \min\left\{1, \frac{\hat{p}(y)\hat{q}(x)}{\hat{p}(x)\hat{q}(y)}\right\} \\ &= \min\left\{1, \exp(\sqrt{x} - \sqrt{y})\right\},\end{aligned}$$

since $\hat{q}(x) = \hat{q}(y) = 1/m$ and the normalising constants $\exp(c)$ cancel.

(ii) [6 marks] For example

```
mh <- function(n, m) {
#mcmc simulating n steps of MC targeting exp(-sqrt(x)), x=1,2,...,m
  X <- numeric(n)
  x <- 1
  for (k in 1:n) {
    y <- ceiling(runif(1)*m)
    u <- runif(1)
    r
    if (u < exp(-sqrt(y)+sqrt(x))) {
      x <- y
    }
    X[k] <- x
  }
  return(X)
}
```

(c) [7 marks] (N)

(i) The transition probability K can be written down as

$$\begin{aligned}[K]_{(x,z),(y,t)} &= \mathbb{P}\big((X_k, Z_k) = (y, t) \mid (X_{k-1}, Z_{k-1}) = (x, z)\big) \\ &= q(y \mid x)h(t \mid y)\alpha\big((y, t) \mid (x, z)\big) \\ &\quad + 1\big((y, t) = (x, z)\big)\rho(x, z),\end{aligned}$$

with

$$\alpha\big((y, t) \mid (x, z)\big) = \min\left\{1, \frac{t}{z}\frac{q(x \mid y)}{q(y \mid x)}\right\}, \quad \text{(if } zq(y \mid x) > 0 \text{ and 0 otherwise)},$$

$$\rho(x, z) = 1 - \sum_{y\in\mathbb{X},t\in\mathbb{N}} q(y \mid x)h(t \mid y)\alpha\big((y, t) \mid (x, z)\big).$$

Reversibility requires us to show that $\pi(x,z)[K]_{(x,z),(y,t)} = \pi(y,t)[K]_{(y,t),(x,z)}$, which is obvious in case $(y,t) = (x,z)$, and for $(y,t) \neq (x,z)$ with $zq(y \mid x) > 0$ and $tq(x \mid y) > 0$,

$$\begin{aligned}
\pi(x,z)[K]_{(x,z),(y,t)} &= h(z \mid x)z\, q(y \mid x)h(t \mid y)\alpha\big((y,t) \mid (x,z)\big) \\
&= h(z \mid x)h(t \mid y) \min\left\{zq(y \mid x), tq(x \mid y)\right\} \\
&= h(t \mid y)tq(x \mid y)h(z \mid x) \min\left\{\frac{z}{t}\frac{q(y \mid x)}{q(x \mid y)}, 1\right\} \\
&= \pi(y,t)[K]_{(y,t),(x,z)}.
\end{aligned}$$

[The idea of the proof is essentially the same as shown for the MH, but expressions are not exactly the same.]

(ii) If the Markov chain is irreducible and has the invariant probability $\pi(x,z) = h(z \mid x)z$, then

$$\begin{aligned}
\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} f(X_k) &= \mathbb{E}_\pi[f(X,Z)], \qquad \text{with } f(x,z) = f(x) \\
&= \sum_{x\in \mathsf{X}} \sum_{z\in\mathbb{N}} \pi(x,z)f(x) \\
&= \sum_{x\in \mathsf{X}} f(x) \sum_{z\in\mathbb{N}} h(z \mid x)z \\
&= \sum_{x\in \mathsf{X}} f(x)p(x) = \mathbb{E}_p[f(X)].
\end{aligned}$$

3. Let $X = [X_1, .., X_p]$ be an $n \times p$ real matrix of rank p with $p < n$ column vectors $X_i,\ i = 1, ..., p$, and let $\beta = (\beta_1, \beta_2, ..., \beta_p)^T$ and $y = (y_1, y_2, ..., y_n)^T$ be real vectors. The normal equations are $X^T X\beta = X^T y$. Consider solving the normal equations for β.

   (a) [7 marks] (i) [4 marks] (B/S close to collection question) $X^T(X\beta)$ is $2np$ operations. $(X^T X)\beta$ is $np^2 + p^2$. R uses the latter to evaluate `t(X)%*%X%*%beta`.

   (ii) [3 marks] (B) $X^T X\beta = X^T y \Leftarrow R^T Q^T QR\beta = R^T Q^T y \Leftarrow R^T R\beta = R^T Q^T y$ (since Q is orthogonal) and hence if β satisfies $R\beta = Q^T y$ then it is a solution of the normal equations.

   (b) [6 marks] Let $X_{2:p} = [X_2, X_3, ...X_p]$. We have

$$\begin{aligned} X &= [X_1, X_{2:p}] \\ Q &= [Q_1, Q'] \\ X &= QR \\ [X_1, X_{2:p}] &= [Q_1, Q'] \begin{pmatrix} a & r \\ 0_{(p-1)\times 1} & R' \end{pmatrix} \\ [X_1, X_{2:p}] &= [aQ_1, Q_1 r + Q'R'] \end{aligned}$$

   We can read off $X_1 = aQ_1$ and $X_{2:p} = Q_1 r + Q'R'$ from the last line. Now $|Q_1| = 1$ $a = |X_1|$ and so $Q_1 = X_1/a$. Multiplying $X_{2:p} = Q_1 r + Q'R'$ by Q_1^T gives $r = Q_1^T X_{2:p}$. Finally, if $X' = X_{2:p} - Q_1 r$ then $X' = Q'R'$.

   Now R' is upper triangular with positive diagonal entries (since R was) and Q' is clearly orthogonal, hence $Q'R'$ is the QR-factorisation of X'.

   (c) [4 marks] Each time we apply this we reduce the number of columns of X by one. Our Algorithm is
   Step 1. If $p = 1$ then $Q = X/|X|$ and $R = |X|$ and we are done.
   Step 2. Otherwise (if $p > 1$) then
   Step 2.1 set $Q_1 = X_1/|X_1|$, $r = Q_1^T X_{2:p}$ and $X' = X_{2:p} - Q_1 r$.
   Step 2.2 call this algorithm to compute the QR factorisation $X' = Q'R'$.
   Step 2.3 Assemble Q and R from Q_1, Q', r and R'.

   (d) [8 marks] S/N

```
my.qr<-function(X) {

norm<-function(v) {sqrt(sum(v^2))}

n=dim(X)[1]; p=dim(X)[2] #assume p<n

R=matrix(NA,p,p)
Q=matrix(NA,n,p)

R[1,1]=norm(X[,1])
Q[,1]=X[,1]/R[1,1]

if (p==1) return(list(Q=Q,R=R))

R[1,2:p]=t(Q[,1])%*%X[,2:p]
R[2:p,1]=0
Xp=X[,2:p,drop=F]-Q[,1,drop=F]%*%R[1,2:p,drop=F]
```

```
QR=my.qr(Xp)

Q[,2:p]=QR$Q
R[2:p,2:p]=QR$R

return(list(Q=Q,R=R))
}
```