STA380 Practice Problems for Quiz 2

These problems are not to be handed in, but they are for extra practice for students to be prepared for the quiz.

1. Generate the following probability density function using the inverse-transformation method:

$$f_Y(y) = \begin{cases} \frac{1}{8} & 0 < y \le 2, \\ \frac{y}{8} & 2 < y \le 4, \\ 0 & \text{otherwise.} \end{cases}$$

   Additionally, graph the histogram against the true density curve to verify your answer.

2. Let X be a continuous random variable with the following probability density function:

$$f_X(x) = 5x^4, \quad 0 \le x \le 1$$

   Write out the algorithm and code to generate a random variable from the distribution of X using the inverse-transform method.

3. Consider the Weibull distribution:

$$f_X(x) = \begin{cases} \frac{\alpha}{\lambda}\left(\frac{x}{\lambda}\right)^{\alpha-1} e^{-(x/\lambda)^{\alpha}} & x \ge 0, \\ 0 & \text{otherwise.} \end{cases}$$

   Where α represents the shape parameter and λ represents the scale parameter. Use the acceptance-rejection method to generate a sample from the Weibull distribution where $\alpha = 2$ and $\lambda = 3$. Let the trial distribution be $Exponential(\theta = 3)$ where θ is the scale parameter.

   (a) What is $c = \max\left\{\frac{f(x)}{g(x)}\right\}$, where $f(x)$ is the target distribution and $g(x)$ is the trial distribution?

   (b) Write out the acceptance-rejection algorithm for this case.

   (c) Write out the `R` code for the acceptance-rejection algorithm.

   (d) Assume that you would like to generate 10^4 value from the target distribution. Determine how many draws (iterations), in average, are required from the trial distribution. The answer should be a number.

4. Let $X_i \overset{i.i.d.}{\sim} N(\mu = 0, \sigma^2 = 1)$ for $i = 1, 2, \ldots, m$, $m \in \mathbb{N}$. Consider the function $g(x) = \frac{1}{m}\sum_{i=1}^{m} f_{X_i}(x)$.

   (a) Is $g(x)$ a valid density function? Why or why not? Justify your answer.

   (b) Is $g(x)$ a convolution or a mixture?

5. Let $U_1 \sim Uniform(0, 1)$ and $U_2 \sim Uniform\{0, 1, 2\}$. (That is, U_1 is a continuous random variable but U_2 is a discrete random variable.) Consider the convolution $X = 0.5U_1 + 0.5U_2$. How would you simulate this using `R` code? Assume $n = 10^4$.
   Hint: I recommend for the discrete uniform random variable to look at the documentation for `sample()` *before attempting to answer this question.*

Below are the solutions.

1. First, find the cdf:

$$F_Y(y) = \begin{cases} 0 & y \leq 0 \\ y/8 & 0 < y \leq 2, \\ y^2/16 & 2 \leq y \leq 4, \\ 1 & y > 4 \end{cases}$$

   Now we need to find the inverse $F_X^{-1}(x)$:

$$F_X(x) \overset{set}{=} u \;\Rightarrow\; u = y/8 \quad \text{for } 0 < y \leq 2 \;\Rightarrow\; y = 8u \quad \text{for } 0 < 8u \leq 2 \;\Leftrightarrow 0 < u \leq \frac{1}{4}.$$

   Consider the second case:

$$u = y^2/16 \quad \text{for } 2 < y \leq 4 \;\Rightarrow\; y = \pm\sqrt{16u} \quad \text{for either } 2 < \sqrt{16u} \leq 4 \quad \text{or } 2 < -\sqrt{16u} \leq 4$$

   Note that the latter case, $2 < -\sqrt{16u} \leq 4$ is impossible so thus we focus on:

$$2 < \sqrt{16u} \leq 4 \;\Leftrightarrow\; \frac{1}{2} \leq u^{1/2} \leq 1 \;\Leftrightarrow\; \frac{1}{4} \leq u \leq 1$$

   Thus, the algorithm is as follows: for each random variate required,

   - Generate a random variable $u \sim Uniform(0,1)$.
   - Deliver

$$x = \begin{cases} 8u & 0 < u \leq \frac{1}{4}, \\ 4\sqrt{u} & \frac{1}{4} < u \leq 1. \end{cases}$$

   Additionally, see the code below.

```
n <- 10000
u <- runif(n)

y <- ifelse(u <= 1/4, 8*u, 4*sqrt(u))

# Histogram
hist(y, prob = TRUE, breaks = 50,
     main = "Histogram with True Density",
     xlab = "y", border = "white", col = "skyblue")

# Define density
fy <- function(x) {
  ifelse(x <= 0, 0,
         ifelse(x < 2, 1/8,
                ifelse(x < 4, x/8, 0)))
}

# Overlay true density curve
curve(fy, from = 0, to = 4, add = TRUE, col = "red", lwd = 2)
```

2. First, we'll need to derive the cumulative distribution function (cdf) $F_X(x)$. For $0 \leq x \leq 1$:

$$F_X(x) = \int_0^x 5t^4 dt = t^5 \Big|_0^x = x^5$$

   Thus,

$$F_X(x) = \begin{cases} 0 & x < 0, \\ x^5 & 0 \leq x \leq 1, \\ 1 & x > 1. \end{cases}$$

   Now, we need to find the inverse $F_X^{-1}(x)$:

$$F_x(x) \overset{set}{=} u \;\Rightarrow\; x^5 = u \;\Rightarrow\; x = u^{1/5}$$

   Thus $F_X^{-1}(u) = u^{1/5}$. The algorithm is as follows: for each random variate required,

- Generate a random variable $u \sim Uniform(0, 1)$.
- Deliver $x = u^{1/5}$.

The code is as follows.

```
u <- runif(10000)
x <- u^(1/5)
hist(x, prob = TRUE)
curve(5 * x^4, from = 0, to = 1, add = TRUE, col = "red", lwd = 2)
```

3. (a) Note that in this case,

$$\frac{f(x)}{g(x)} = \frac{\frac{2}{9}xe^{-(x/3)^2}}{\frac{1}{3}e^{-x/3}} = \frac{2}{3}xe^{-(x/3)^2+x/3}$$

Taking the derivative,

$$\begin{aligned}
\frac{d\left(\frac{f(x)}{g(x)}\right)}{dx} &= \frac{2}{3}e^{-(x/3)^2+x/3} + \frac{2}{3}xe^{-(x/3)^2+x/3}\left[-2\left(\frac{x}{3}\right)\left(\frac{1}{3}\right) + \frac{1}{3}\right] \\
&= \frac{2}{3}e^{-(x/3)^2+x/3} + \frac{2}{3}xe^{-(x/3)^2+x/3}\left[-\frac{2}{9}x + \frac{1}{3}\right] \\
0 &\overset{set}{=} \frac{2}{3} + \frac{2}{3}x\left[-\frac{2}{9}x + \frac{1}{3}\right] \\
&= -\frac{4}{27}x^2 + \frac{2}{9}x + \frac{2}{3} \\
&= 2x^2 - 3x - 9 \\
&= (2x+3)(x-3)
\end{aligned}$$

This gives us two solutions: $x = -3/2$ and $x = 3$. We cannot accept the first solution since the support for the Weibull and Exponential distribution are strictly positive. Hence, the maximum occurs when $x = 3$ and therefore:

$$c = \max\left\{\frac{f(x)}{g(x)}\right\} = \frac{2}{3}(3)\exp\left\{-\left(\frac{3}{3}\right)^2 + \frac{3}{3}\right\} = 2$$

(b) For each random variate required,

(i) Generate $Y \sim Exponential(\theta = 3)$.

(ii) Generate $U \sim Uniform(0, 1)$.

(iii) If

$$u < \frac{1}{3}xe^{-(x/3)^2+x/3}$$

accept y and set $x = y$. Otherwise, reject y and generate a random variate again.

(c)

```
n <- 10^4
accepted <- numeric(n)
u_accepted <- numeric(n)
i <- 0
iteration <- 0
while(i < n){
  y <- rexp(n = 1, rate = 1/3) # candidate from g
  u <- runif(1) # u ~ uniform(0, 1)
  ftgt <- (1/3) * y * exp(-(y/3)^2 + y/3) # f(x)/cg(x)

  if(u < ftgt){
    i <- i+1
    accepted[i] <- y
    u_accepted[i] <- u
  }
  iteration <- iteration + 1
}

hist(accepted, prob = TRUE, ylim = c(0, 1), col = "skyblue", border = "white")
curve((2/3) * (x/3) * exp(-(x/3)^2), 0, 8, add = TRUE, col = "red", lwd = 2)
```

(d) $nc = 10^4(2) = 20,000$

4. (a) To check if this is a valid density function, we need to check that it is strictly non-negative and the integral over the support evaluates to 1. Note that $f_{X_i}(x) \geq 0$ because we are given that X_i is Gaussian distributed (which is a well known distribution). Furthermore, $m \in \mathbb{N} \;\Rightarrow\; m \geq 0 \;\Rightarrow\; \frac{1}{m} \geq 0$. Thus,

$$f_{X_i}(x) \geq 0 \;\; \forall i \in \{1, 2, \ldots, m\} \;\Rightarrow\; \sum_{i=1}^{m} f_{X_i}(x) \geq 0 \;\Rightarrow\; \frac{1}{m}\sum_{i=1}^{m} f_{X_i}(x) \geq 0.$$

Thus $g(x) \geq 0$. Now,

$$\int_{-\infty}^{\infty} \frac{1}{m}\sum_{i=1}^{m} f_{X_i}(x)\,dx = \frac{1}{m}\sum_{i=1}^{m}\int_{-\infty}^{\infty} f_{X_i}(x)\,dx = \frac{1}{m}\sum_{i=1}^{m} 1 = \frac{1}{m}m = 1.$$

Thus this is a valid density function.

(b) This is a mixture.

5.
```
n = 10^4
u1 <- runif(n)
u2 <- sample(c(0, 1, 2), size = n, replace = TRUE, prob = c(1/3, 1/3, 1/3))
x <- 0.5*u1 + 0.5*u2
```