Tutorial 6 STQD6214

1. (*Sum of two independent Poisson r.v.s*)
   Suppose X and Y are independent random variables such that X follows the Poisson distribution with mean 3, and Y follows the Poisson distribution with mean 4. The functions `dpois`, `ppois`, `qpois` and `rpois` in R are useful for the Poisson distribution. For these functions, the parameter or argument `lambda` represents the mean of the Poisson distribution. Answer the following questions using R.
   a) Find the probability $P(X = 3)$.
   b) Find the probability $P(X \leq 3)$.
   c) Generate 100 random samples of Y. Calculate its mean and variance. Compare these values with the theoretical mean and variance of Y ($\text{mean}(Y) = \text{var}(Y) = 4$).
   d) In theory the random variable $Z = X + Y$ should follow the Poisson distribution with mean 7.
      - Write a function that generates n samples for the random variable Z. The function should generate n samples of X, and n samples of Y, and calculate $Z = X + Y$.
      - Generate 1000 samples of Z. Calculate its mean and variance.
      - From your generated values of Z, what is the estimated value for the probability $P(Z = 7)$? Compare your answer with the theoretical probability using Poisson distribution with mean 7.

2. (*Monte Carlo integration*)
   Not all integrals are computable analytically. In some cases, approximations can be used to estimate the integrals. One such approximation technique is called the Monte Carlo integration. The simplest Monte Carlo integration approximates an integral by using

$$\int_a^b f(x)\,dx \approx \frac{1}{n}\sum_{i=1}^{n} f(X_i)$$

   where X_i, $i = 1, \dots, n$ are randomly generated using the uniform distribution from a to b. For this question, we will try to approximate the integral

$$\int_1^2 x^x\,dx$$

   a) Write a function `f`, where the function calculates $f(x) = x^x$.
   b) Write a function which takes input n.
      - The function will randomly generate n samples of X from the uniform distribution with min 1 and max 2.
      - The function then estimates the integral by calculating

$$\frac{1}{n}\sum_{i=1}^{n} f(X_i)$$

      - The function will then return the estimate.
   c) Using the function written in (b), estimate the integral

$$\int_1^2 x^x\,dx$$

Note: You can use the command "`integrate(f, 1, 2)`" in R to approximate the integral using another approximation method (not Monte Carlo integration). You can use this to compare your answer.

3. (*Rejection sampling*)

The Beta(3,5) distribution has the probability density function (pdf)

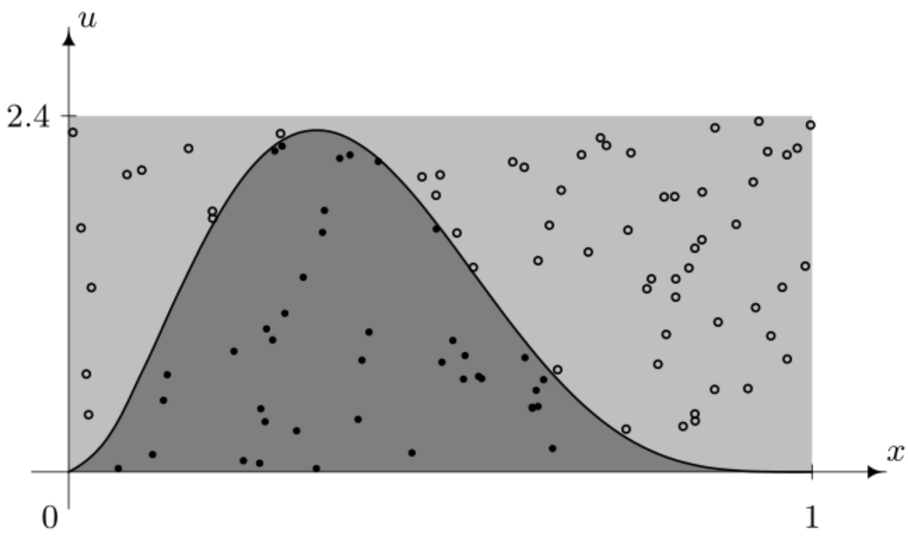
$$f(x) = \frac{\Gamma(8)}{\Gamma(3)\Gamma(5)} x^2(1-x)^4, \qquad \text{for } 0 < x < 1,$$

where $\Gamma(a) = \int_0^\infty t^{a-1} \exp(-t)\, dt$ is the Gamma function. This distribution attains its maximum pdf value at $x = \frac{1}{3}$ with $f(x) \approx 2.305$. The command "`dbeta(x, 3, 5)`" in R can be used to compute the pdf.

Rejection sampling is a method to generate random samples from a given distribution. Using rejection sampling, we can generate random samples from the Beta(3,5) distribution using the following algorithm:

i) Generate a random value of x using uniform distribution from 0 to 1.
ii) Generate a random value of u using uniform distribution from 0 to 2.4.
iii) If $u \le f(x)$ where $f(x)$ is the pdf of Beta(3,5), then accept x as a sample of Beta(3,5) distribution. Otherwise, reject it and repeat step (i) of the algorithm.

The figure below illustrates the sampling process where filled circles denote accepted values, and empty circles denote rejected values.



Write a function that generates n samples from Beta(3,5) distribution using the rejection sampling algorithm. Then using your function, generate 1000 samples, plot a histogram of the samples, and compare it to the pdf of Beta(3,5) distribution.