HOMEWORK 2 471

HW2 is due by Tuesday September 29th at 11:59 pm CT pm. Please upload your solutions to Canvas.

Attach the codes/command windows that you use to answer the questions.

(1) Consider the function:

$$f(x) = \frac{1}{2\sqrt{x}} \mathbb{1}_{0<x<1}.$$

- Is f a pdf?
- Suppose that $U \sim \text{Uniform}([0,1])$. How would you use U to construct a random variable X with pdf f?
- Generate a sample of size 10^3 from f and create a histogram using your samples.

(2) Consider the function

$$g(x) = (\sin(x))^2 |x^3 + 2x - 3| \mathbb{1}_{x \in (-1,0) \cup (1/2,1)},$$

and the pdf $f(x) = cg(x)$, where c is chosen so that f is a true density function.

- Use rejection sampling to generate a sample of size 10^3 from f and compare a plot of g to a histogram of your samples. The idea is that you present the histogram and the graph of the function g in the same image. To do so, you have to rescale your histogram. Let us give an intuition on how to rescale the histogram. Suppose that the k-th bin (denoted by B_k) for the histogram, has length w_k. Let n_k be the number of samples in B_k. Then,

$$\frac{n_k}{n} \approx \int_{B_k} f(x)dx = w_k f(\overline{x_k}).$$

In the above, the approximation holds because of the law of large numbers, and the equality holds from the mean value theorem ($\overline{x}_k$ is a point in the bin B_k). We conclude that $\frac{n_k}{n w_k}$ has the same scale as f.
- Use essentially the same rejection sampling algorithm to approximate c.

(3) Consider the function

$$g(x) = \frac{1}{2\sqrt{x}} \sqrt{|\cos(x^2 + x)|} \mathbb{1}_{0<x<1}$$

and let f be the pdf $f(x) := cg(x)$ for appropriate constant c.

- Use rejection sampling to generate a sample of size 10^3 from f (Note that the graph of f can not be enclosed by a box!).
- Use essentially the same rejection sampling algorithm to approximate c.

(4) Consider the function $g : \mathbb{R} \to \mathbb{R}$, given by

$$g(x) := e^{-|x|}.$$

- Is g a probability density function? If not, find a normalization constant Z so that $f(x) = \frac{1}{Z} g(x)$ is a true probability density function.
- Given that you know how to sample from the uniform distribution $Uniform([0,1])$, how would you sample from the distribution that has density f?.
- Use your answer to the previous question and the rejection sampling algorithm to generate 100 samples from a standard Gaussian. In your algorithm, on average how many samples from f were you rejecting for every accepted sample? Is this what you were expecting?

- Create a QQ-plot (quantile -quantile plot) between the samples produced in the previous question and 100 samples of a standard Gaussian generated using the Box-Muller method. What does the QQ plot tell you?

(5) Consider the function $g : [1,2] \times [0,1] \to \mathbb{R}$ (notice the domain of the function) given by

$$g(x,y) := xe^{xy}.$$

- Is g a joint density? If not, find a constant c such that $f(x) = cg(x)$ is a true density.
- The idea is to sample $(X_1, Y_1), \ldots, (X_n, Y_n)$ i.i.d. random vectors with joint density f. How would you obtain these samples? Write a code and obtain 1000 samples from f. Create a scatter plot of the samples.
  Note: You can follow at least two strategies to answer this question: 1) Use rejection sampling with a three dimensional box enclosing the graph of the function, and 2) sample iteratively, first a marginal then a conditional.