Statistical Data Analysis
Problem sheet #3
Due Monday, 28 October 2024

Exercise 1 [6 marks]: Consider two continuous random variables x and y that follow the joint pdf $f(x, y)$ and define $u = x + y$. Show that the pdf of u can be written

$$g(u) = \int_{-\infty}^{\infty} f(x, u - x)\, dx \ .$$

Use a method analogous to what was shown in the lectures for the product of two random variables (see p. 9 of the week 2 slides).

Exercise 2 [7 marks]: Suppose x and y are independent and exponentially distributed each with mean values θ and define $u = x + y$. By using the result from Ex. 1, find the pdf of u. (In fixing the limits of integration, remember that the pdf is nonzero only for $x \geq 0$ and $y \geq 0$.)

Exercise 3 [7 marks]: Consider a continuous random variable x that follows the pdf $f(x)$ with cumulative distribution $F(x)$, and suppose r follows a uniform distribution on $[0, 1]$. Prove (as was claimed in the lectures) that if we set $F(x) = r$ and solve for x, that $x(r)$ follows the pdf $f(x)$. To do this, use the method discussed in the lectures for finding the pdf of a function, and use the inverse function theorem, which says that

$$\frac{d}{dr} F^{-1}(r) = \frac{1}{\frac{dF}{dx}(x(r))} \ .$$

Exercise 4 (computing warm up [0 marks]): There is nothing to turn in for this exercise – it is just a warm-up exercise to ensure that you have your computing environment set up.

Starting with `simpleMC.py`, `simpleMC.ipynb` or `simpleMC.cc` from the course website, generate 10000 random values uniformly distributed between 0 and 1 and display the result as a histogram with 100 bins. (This is what simpleMC already does; you just need to ensure that you can run it.)