ST444: Exercise Sheet 5

Part I: Python

1. Write a function `overlapping()` that takes two lists and returns `True` if they have at least one member in common, or `False` otherwise.

2. Write a (8-digit) password generator in Python. The password should be random. Strong passwords have a mix of lowercase letters, uppercase letters, numbers, and symbols.

   (Hint: use Python's `random` module.)

3. Implement Crout's LU decomposition.

4. Write a function that computes the determinant of a positive-semidefinite square matrix.

   (Hint: use LU decomposition first)

5. Write a function that computes the inverse of a positive-definite square matrix.

   (Hint: could either use Gaussian elimination or LU decomposition)

6. Implement the power method that finds the largest eigenvalue and its corresponding eigenvector of a positive-definite symmetric square matrix.

7. Given n p-dimensional observations, compute its sample covariance matrix and find the first principal component.

8. Implement the Box–Muller algorithm (both the trigonometric version and the rejection polar version). In Python, `sin`, `cos` and `log` can be called as follows:

```
# this line allows us to use functions from the math library
import math
# sin
math.sin(0.1)
# cos
math.cos(0.2)
# log
math.log(0.3)
```

9. Recall the definition of the Student's t distribution (non-examinable). Then write a function that takes a positive integer p and returns a random number from the Student's t distribution of degree p.

   (Hint: make good use your code from the previous question.)

10. Given a univariate function f, our aim is to evaluate $\int_0^1 f(x)dx$. For a fixed N (number of subintervals), implement the rectangle rule, the mid-point rule, the trapezoid rule and Simpson's rule.

Part II: RNG and numerical integration (theory)

1. ($\star$) Prove the validity of Box–Muller algorithm (both the trigonometric version and the rejection polar version)

2. Starting from independent uniform random variables, how would you simulate from each of the following distributions:

   (a) Geometric distribution with probability p.

   (b) Logistic distribution, with density $f(x) = \frac{e^{-x}}{(1+e^{-x})^2}$ for $x \in \mathbb{R}$.

   (c) Cauchy(0,1) distribution, with density $f(x) = \frac{1}{\pi(1+x^2)}$ for $x \in \mathbb{R}$

   Hint: try the inverse transform method.

3. We want to generate $N(0,1)$ random variables using the rejection method described in the lectures, i.e., taking Laplace with parameter λ (its density function is $g_\lambda(x) = \lambda \exp(-\lambda|x|)/2$ for $x \in \mathbb{R}$) as the envelope density.

   (a) Describe the algorithm.

   (b) Find the optimal $\lambda > 0$ that minimises the rejection probability.

   (c) How does this compare with the rejection polar version of Box–Muller?

4. Prove the following result used in Simpson's rule: let $P(x)$ be the quadratic polynomial passing through $(x_1, f(x_1))$, $(x_2, f(x_2))$ and $\left(\frac{x_1+x_2}{2}, f\left(\frac{x_1+x_2}{2}\right)\right)$, then

$$\int_{x_1}^{x_2} P(x)dx = \frac{x_2 - x_1}{6}\left(f(x_1) + 4f\left(\frac{x_1+x_2}{2}\right) + f(x_2)\right).$$

5. Let $I = \int_0^{1/2} \frac{1}{\pi(1+x^2)}dx$. Construct Monte-Carlo estimators for I based on $X_1, \ldots, X_N \overset{i.i.d.}{\sim} f_i$, where

   (a) f_1 is the density of a $U[0,1/2]$.

   (b) f_2 is the density of a Cauchy(0,1).

   and compute their variances.