Normally distributed data can be modeled entirely in terms of their means and variances/covariances (in the uni- or multivariate case). Hence, estimating the mean and the covariance matrix is therefore a problem of great interest in statistics and it is also of great significance to consider the correct statistical model. If a vector $\mathbf{X}$ is $N_p(\mu, \mathbf{\Sigma})$ distributed, then it must be true that any function of $\mathbf{X}$ also follows some distribution, as a function of something random should still be random in its own way. For example, $\mathbf{XAX}^T$ would have some random behaviour (for some suitable symmetric matrix $\mathbf{A}$), or the sample covariance matrix $\mathbf{S} = \frac{\mathbf{X}^T(\mathbf{I}_N - \frac{1}{N}\mathbf{1}_N\mathbf{1}_N^T)\mathbf{X}}{(N-1)}$. In the univariate case, we know that the sample covariance follows a χ^2 distribution (Theorem 8.3.6 in the textbook by Bain and Engelhardt). Our departure point in this assignment is based on this fact.

Consider the following paper on clickUP:

Siriteanu, C., Kuriki, S., Richards, D., & Takemura, A. (2016). Chi-square mixture representations for the distribution of the scalar Schur complement in a noncentral Wishart matrix. Statistics & Probability Letters, 115, 79-87.

Note: you need not study the paper in depth beyond Section 2.

Consider the following questions:

1. Do the following via simulation, using R/Python/SAS.
   (a) Simulate 50 values from a (univariate) normal distribution with mean 44 and standard deviation 15, and calculate the statistic $\frac{(n-1)S^2}{\sigma^2}$. Repeat this 1000 times, and make a histogram of the statistic. Overlay the corresponding density curve onto the statistic, and shortly comment on the fit. (7 marks)
   (b) Repeat the first step, but use 500 simulated values instead. Shortly comment on the fit. (3 marks)
2. Discuss the following (make sure to form succint, meaningful, and informed answers in about five sentences) by consulting, as a start, the paper listed above.
   (a) What is a Wishart matrix/Wishart distribution? How does it, if at all, relate to the χ^2 distribution? (2 marks)
   (b) Describe broadly (in algorithm-style, if you prefer) how one would go about to simulate (central) Wishart random variables. (4 marks)
   (c) Verify the dimensions of equation (1) in the paper above. Draw any parallels between equation (1) i.e. $w_{11.2}$ and something else you have learned before. (4 marks)

Cite any arguments you make. You can include examples if you feel that would strengthen your response. You can use published papers, text books, and internet references. Add all your references and code in an Appendix to the end of the assignment. Additional marks may be awarded for neatness, correct and complete referencing, spelling and grammar, and page outlay.

Instructions:

- You can work in groups of at most 5 to complete this assignment.
- Your assignment must be submitted on Gradescope (the link is on clickUP) before the due date. When you submit, you have to add/list group members to ensure that they are "linked" to your assignment. Only one person per group needs to submit the assignment.
- UP Rules and Regulations regarding this assignment are valid. If any unethical allegations (including but not limited to, adding a classmate's name to an assignment if they have not contributed suitably) come to light, then the matter will be referred to the UP Legal Office and all group members would be scrutinised and considered complicit.
- Submit your assignment in PDF format.
- Your submission must be free from spelling/grammar errors and be presented formally and professionally.
- Your submission may be at most five A4 pages (not including the appendix). Use a sans serif font with usual A4 margins with font size 11. Start each question on a new page. Do not include a cover page or a table of contents!

Submission deadline: Friday 2 May 23h59