Quiz 0

No marks. This quiz is only to help us understand your background and help you understand desirable bakcground knowledge. The TAs will help you pick up these basics if you need BUT you will need to work as well.

1. Suppose that you have two different types of pumps with different probabilities p's of failure on any particular day. Let these probabilities be p_1 and p_2. What is the probability that the first failure is on the i-th day? Assume that the chance of failure does not depend on the number of days that a pump is in use.

2. Classify the following matrices as positive definite, positive semidefinite, indefinite[1], *etc.*, while providing justification and outlining all the steps:

(a)

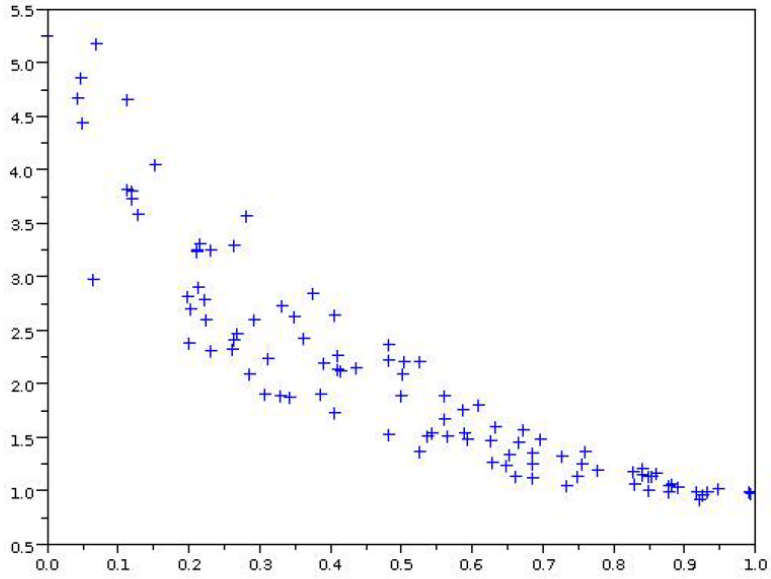
$$A = \begin{bmatrix} 5 & 3 & 1 \\ 3 & 4 & 2 \\ 1 & 2 & 6 \end{bmatrix}$$

(b)

$$A = \begin{bmatrix} 1 & 0 & 4 \\ 0 & 2 & 0 \\ 4 & 0 & 18 \end{bmatrix}$$

3. Look at the figure below which plots the female literacy (X axis) and the average number of children that a woman has (Y axis), for about 100 habitations. Make english language comments about this graph and illustrate what mathematical procedures you would do to validate/justify your comments. Just from the data, is X likely to determine Y or the other way around? Why?

[1] Note that a matrix A is indefinite if neither A nor $-A$ is positive semi-definite.



Problem 4. A lab test is 99% effective in detecting a disease when in fact it is present. However, the test also yields a false positive for 0.5% of the healthy patients tested. If 1% of the population has that disease, then what is the probability that a person has the disease given that his/her test is positive?

Problem 5. Show that $f(x, y) = \frac{1}{x}$, $0 < y < x < 1$, is a joint density function.

Next, assuming that f is the joint density function of X, Y, find

1. the marginal density of Y;
2. the marginal density of X;

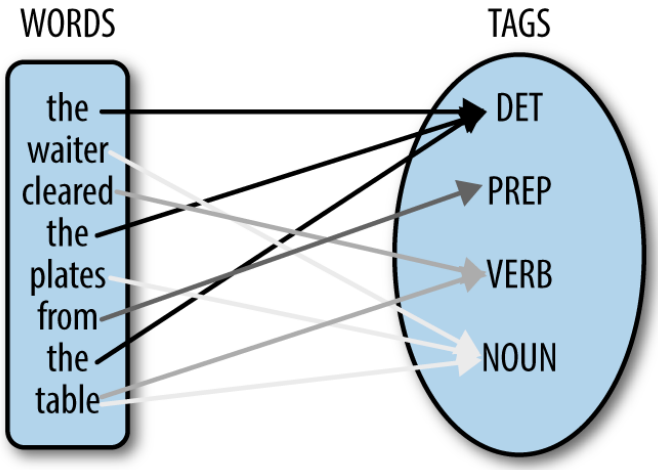
Problem 6. Let X and Y be *independent continuous* random variables with same density functions

$$p(x) = \begin{cases} e^{-x} & \text{if } x > 0; \\ 0 & \text{otherwise.} \end{cases}$$

Find density $\frac{X}{Y}$.

Problem 7. POS tagging is a problem of great importance in the field of Natural Language Processing (**NLP**)
Input: A set of n-words
Output: POS tag for each word

Assuming the picking of words is done independently, find probability that the set contains a 'noun' given that it contains a 'verb'. You can set up notations and define random variables as needed to solve this problem.


WORDS
TAGS
the
waiter
cleared
the
plates
from
the
table
DET
PREP
VERB
NOUN