MLPR background self-test

Please answer these questions before the end of the first week of teaching. If more than one or two of these questions are a struggle, please seriously consider taking a different course.

Make sure to have a proper attempt before looking at the answers. If necessary, work through tutorials on probability, linear algebra, and calculus first. If you look at the answers before you've had an honest attempt, it will be too easy to trick yourself into thinking you have no problems.

1. Two binary outcomes have joint probability table:

$P(x,y)$	$x=1$	$x=2$
$y=1$	0.1	0.3
$y=2$	0.2	0.4

   i) What is the conditional probability distribution $P(x\,|\,y\!=\!1)$?

   ii) Are the two variables independent? Justify your answer.

2. Random variable X has mean μ_x and standard deviation σ_x. Similarly, random variable Y has mean μ_y and standard deviation σ_y. The variables X and Y are independent.

   An outcome from random variable Z is generated by combining the two outcomes from the random variables above as follows: $z = x + 3y$.

   What is the mean and standard deviation of Z?

3.   i) Show that for any discrete random variables X and Y, and any outcomes x and y:

   $$P(X\!=\!x, Y\!=\!y) \leq P(Y\!=\!y).$$

   Also state when the inequality is an equality.

   ii) Does the inequality $p(x,y) \leq p(y)$ hold for the joint and marginal *probability density functions* (PDFs) of real-valued random variables x and y? Justify your answer.

4. Two matrices are:

   $$A = \begin{bmatrix} 1 & 3 & 5 \\ 4 & 6 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 5 & 2 \\ 7 & 9 & 1 \\ 8 & 2 & 3 \end{bmatrix}.$$

   What shape is $C\!=\!AB$, and what is $C_{2,3}$?

5. A is a 2×3 matrix, a matrix with 2 rows and 3 columns.

   X is a $N\times3$ matrix, where each row gives one of a set of N points scattered around a 3-dimensional space.

   i) What is the size of the matrix $Y = XA^\top A$? What can you say about the geometrical arrangement of the points represented by rows of this matrix? (Hint: consider $XA^\top$ first.)

   ii) Could the matrix $A^\top A$ be invertible? Justify your answer.

   iii) Could the matrix $AA^\top$ be invertible? Why? If it can be inverted, is the inverse $A^{-\top}A^{-1}$? (Notation: $A^{-\top} = (A^\top)^{-1} = (A^{-1})^\top$.)

6. Given the function $f(x, y) = (2x^2 + 3xy)^2$, find the vector of partial derivatives:

$$\nabla f = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix}.$$

    i) For a small vector $\delta = [\delta_x \; \delta_y]^\top$, give an interpretation of the inner-product $(\nabla f)^\top \delta$.

    ii) Hence show that for small moves from a position (x, y), the direction that will increase the function f the most points along ∇f.

    (Hint: recall that $\mathbf{a}^\top \mathbf{b} = |\mathbf{a}||\mathbf{b}| \cos \theta$, where θ is the angle between the two vectors $\mathbf{a}$ and $\mathbf{b}$, and $|\cdot|$ gives the length of a vector.)

7. Derive the variance of a random variable drawn from a Uniform$[0, 1]$ distribution.

If you aren't sure where to start, remind yourself of the definition of the variance of a probability distribution. You'll need to do some integration involving the PDF of the Uniform distribution.

8. Programming question for those not from CS backgrounds:

In the programming language of your choice, write a function to return the minimum value in a list of numbers. If your favourite programming language has a `min()` function, please ignore that fact, and write the algorithm yourself in the lower-level tools provided by the language.

This course assumes you have done *some* programming before!