

Week 4 exercises

This is the second page of *assessed* questions, as described in the background notes. These questions form 70% of your mark for Week 4. The introductory questions in the notes and the Week 4 discussion group task form the remaining 30% of your mark for Week 4.

Unlike the questions in the notes, you'll not immediately see any example answers on this page. However, you can edit and resubmit your answers as many times as you like until the deadline (Friday 16 October 4pm UK time). This is a *hard deadline*: This course does not permit extensions and any work submitted after the deadline will receive a mark of zero. See the late work policy.

Queries: Please don't discuss/query the assessed questions on hypothesis until after the deadline. If you think there is a mistake in a question this week, please email Iain.

Please only answer what's asked. Markers will reward succinct to-the-point answers. You can put any other observations in the "Add any extra notes" button (but this is for your record, or to point out things that seemed strange, not to get extra credit). Some questions ask for discussion, and so are open-ended, and probably have no perfect answer. For these, stay within the stated word limits, and limit the amount of time you spend on them (they are a small part of your final mark).

Feedback: We'll return feedback on your submission via email by Wednesday 21 October.

Good Scholarly Practice: Please remember the University requirements for all assessed work for credit. Furthermore, you are required to take reasonable measures to protect your assessed work from unauthorised access. For example, if you put any such work on a public repository then you must set access permissions appropriately (permitting access only to yourself). You may not publish your solutions after the deadline either.

1 More practice with Gaussians

1. [The website version of this note has a question here.]

2. [The website version of this note has a question here.]

3. Logistic regression linear separability (30 marks)

Maximum likelihood logistic regression maximizes the log probability of the labels,

$$\sum_n \log P(y^{(n)} | \mathbf{x}^{(n)}, \mathbf{w}),$$

with respect to the weights \mathbf{w} . As usual, $y^{(n)}$ is a binary label at input location $\mathbf{x}^{(n)}$.

The training data is said to be *linearly separable* if the two classes can be completely separated by a hyperplane. That means we can find a decision boundary

$$P(y^{(n)} = 1 | \mathbf{x}^{(n)}, \mathbf{w}, b) = \sigma(\mathbf{w}^\top \mathbf{x}^{(n)} + b) = 0.5, \quad \text{where } \sigma(a) = \frac{1}{1 + e^{-a}},$$

such that all the $y = 1$ labels are on one side (with probability greater than 0.5), and all of the $y \neq 1$ labels are on the other side.

a) [The website version of this note has a question here.]

b) [The website version of this note has a question here.]