Name:

|  |  |  |
| --- | --- | --- |
| Problem | Points | Score |
| 1(a) | 20 |  |
| 1(b) | 10 |  |
| 1(c) | 10 |  |
| 2(a) | 10 |  |
| 2(b) | 5 |  |
| 2(c) | 5 |  |
| 3 | 40 |  |
| Total | 100 |  |

Notes:

1. The exam is open books and notes, but no AI tools or Google searching please ☺ You can browse the class web site.
2. Please indicate clearly your answer to the problem.
3. Note that ungrammatical sentences, incoherent statements, or general illegible scratches will get zero credit.
4. If I can’t read or follow your solution, it is wrong, and no partial credit will be awarded.

**Problem No. 1**: You observe the following data: 1, 2, 1, 2. You assume these are drawn from a Poisson distribution: $p\left(θ\right)=\frac{e^{-θ}θ^{x}}{x!}$.

(a) **(20 pts)** Derive an expression for the maximum likelihood estimate of $θ$.

(b) **(10 pts)** Sketch the likelihood function and demonstrate that your answer to (a) is consistent with your plot. Clearly label your axes and critical points on the graph.

(c) **(10 pts)** Suppose you observe two more datapoints: $x = 3$ and $x = 4$. Which datapoint is more consistent with the model trained using the estimate in (a) and the original training set. In other words, as evaluation data, which data point is more likely to be drawn from the distribution. Be as specific as possible and use what you learned in (a) and (b) to justify your answer. Guesses with no supporting evidence will be given a grade of $0$.

**Problem No. 2:** Given the data shown to the right, and assuming each class is equally probable:



 (a) **(10 pts)** Classify the data point “?” using Quadratic Discriminant Analysis (QDA, aka class-dependent PCA). Sketch your solution showing as much detail as possible, including support regions, means, covariances, etc. You do not need to do many calculations – use intuition.

 (b) **(5 pts)** Classify the data using a Euclidean distance measure and a nearest neighbor solution (the class assignment is based on the label of an actual data point closest to the test vector). Use simple sketches or a short description to justify your answer.

(c) **(5 pts)** Explain any similarities or differences between these classifiers. Why are they similar? Why are they different? Don’t simply describe the pictures you have drawn. Analyze the results and justify your conclusions.

**(40 pts) Problem No. 3**: Given the data shown to the right, discuss the differences between classifying the data using class-independent PCA, class-dependent PCA and, most importantly, using EM to estimate the parameters of a two-mixture Gaussian distribution per class.



The black class (upper left, lower right) is class 0, and the blue class (upper right, lower left) is class 1. Within these classes, you know which of the component distributions the data belongs to. For example, class 1 could represent “pets”: “dogs” (upper left) and “cats” (bottom right). Class 2 could represent “farm animals”: “horses” (lower left) and “cows” (upper right). Your classifier should decide between the two classes (“pets” and farm animals”).

For EM, use a two-mixture Gaussian distribution per class. Estimate the means and covariances for each component from the data assuming you know the class labels for each data point, and the sub-class labels as well. Discuss the challenges associated with this and assume a best case scenario where training goes well (this is purposefully a bit vague).

Be very thorough in explaining your answers. Plan your answer before you start writing.