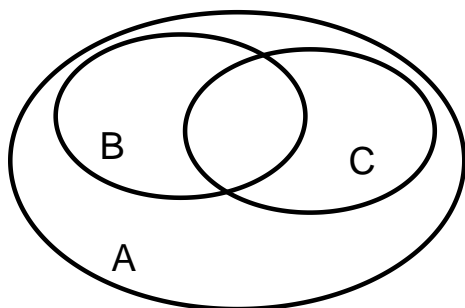


Exercise 1:

In Chapters 3.5 and 3.6 of the lecture, we discussed three key generative classification methods: Linear Discriminant Analysis (LDA), Quadratic Discriminant Analysis (QDA), and Naive Bayes (NB). The Venn diagram below represents these methods as anonymized regions labeled A, B, and C, illustrating the relationships between these methods under a specific assumption.

Your task:

- (a) Summarize the underlying assumptions of LDA, QDA, and NB.
- (b) Assign each method (LDA, QDA, and NB) to the appropriate region (A, B, or C) in the Venn diagram based on these assumptions.
Hint: Firstly, think about how B and C could be seen as special case of A, respectively. Then, consider the shared assumptions that might overlap between B and C to specify their intersection.
- (c) As mentioned, the Venn diagram is true under a "specific assumption". What exactly is this assumption?



Exercise 2:

Consider a classification problem with n observations and p features. The response variable y is categorical with K classes ($y \in \{1, 2, \dots, K\}$). Assume all features are continuous and there are no interaction terms or higher-order terms. For each of the following classification methods, determine the number of parameters that need to be estimated:

- (a) Linear Discriminant Analysis (LDA)
- (b) Quadratic Discriminant Analysis (QDA)
- (c) Naive Bayes (where features are modeled as conditionally independent Gaussian distributions given the class)

Provide explicit expressions for the number of parameters in terms of p and K for each method.