Quiz 1: Curve Fitting Concept Check

Instructions: For each multiple choice question, 1 to 3 options may be correct. Select all that apply.

1. **The Loss Function**
   In the context of the regression model, which of the following statements correctly describe the **loss function**?

   A. It serves as the objective measure that the algorithm minimizes to find the optimal parameters.

   B. It calculates the discrepancy (error) between the model's predictions and the actual observed data.

   C. It automatically selects the optimal polynomial degree for the model without human intervention.

   D. A common example used in this context is the Mean Squared Error (MSE).

2. **Understanding Overfitting**
   When a model is described as **overfitting** the data, which of the following characteristics are typically observed?

   A. The model performs exceptionally well on the training data, often capturing the noise rather than just the signal.

   B. The model generalizes poorly, showing high error rates on new, unseen data.

   C. The learned parameters (weights) of the model often exhibit very large magnitudes.

   D. The model is too simple to capture the underlying structure of the data (high bias).

3. **Hyperparameters**
   Distinguishing between parameters and hyperparameters is crucial in machine learning. Which of the following statements regarding **hyperparameters** are true?

   A. They are configuration variables set before the learning process begins (they are not learned via optimization).

   B. They are learned directly from the training data alongside the weights and biases during the optimization loop.

   C. Examples in the context of this lecture include the polynomial degree m and the regularization coefficient λ.

   D. Determining their optimal values typically involves evaluating performance on a validation set.

4. **Model Selection**
   When selecting the best model (e.g., choosing the degree m), what constitute valid criteria for the optimal choice?

   A. The model achieves the lowest possible error on the training set, regardless of complexity.

   B. The model minimizes the error on a held-out validation or test set.

   C. The model strikes a balance, avoiding both high bias (underfitting) and high variance (overfitting).

   D. The model uses the highest degree polynomial available to ensure maximum flexibility.

5. **Regularization**
   How does introducing a **regularization** coefficient (λ) affect the regression model?

   A. It adds a penalty to the loss function based on the magnitude of the model's parameters.

   B. It allows the model to fit the training data more closely by increasing the effective complexity.

   C. It helps prevent overfitting by discouraging extreme parameter values.

   D. It ensures that the training error will reach zero.

6. **The Data**
   Regarding the **data** and the polynomial fitting process:

   A. The observations are assumed to be generated by a deterministic function with no random noise.

   B. A polynomial model with a sufficiently high degree can theoretically pass through every single point in the training set.

   C. Increasing the amount of training data generally helps reduce the risk of overfitting for a fixed model complexity.

   D. The goal is to find a curve that ignores the underlying signal and fits the noise distribution.

7. **Generalization (Open Ended)**
   In a paragraph or two, explain the concept of **generalization** in the context of model building and selection.