

Sample Questions

*The following sample questions are not inclusive and do not necessarily represent all of the types of questions that comprise the exams. The questions **are not** designed to assess an individual's readiness to take the exam.*

Applied Statistics for Machine Learning

Question 1:

You are gathering data on the circumference of tree trunks in centimeters. What is the level of measurement for circumference?

- A. Ratio
- B. Interval
- C. Ordinal
- D. Nominal

Answer: A

Question 2:

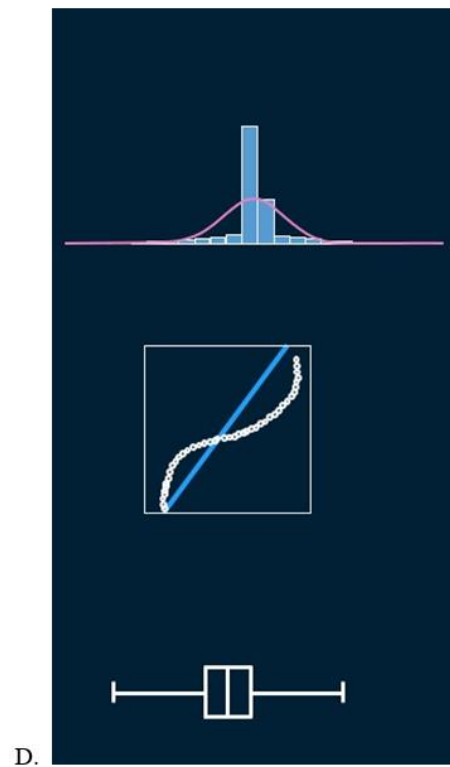
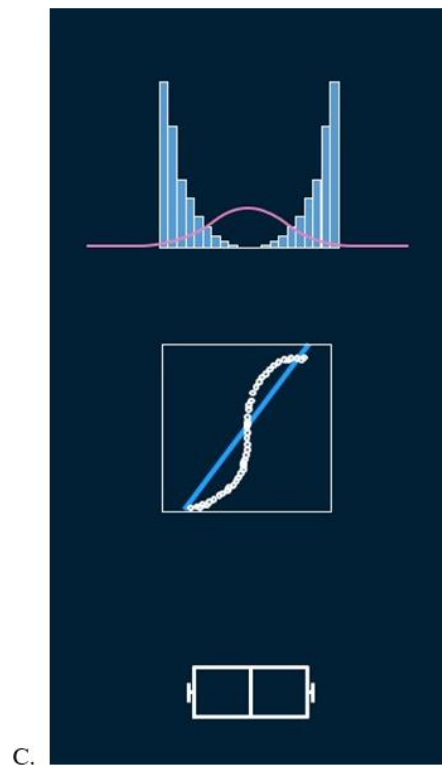
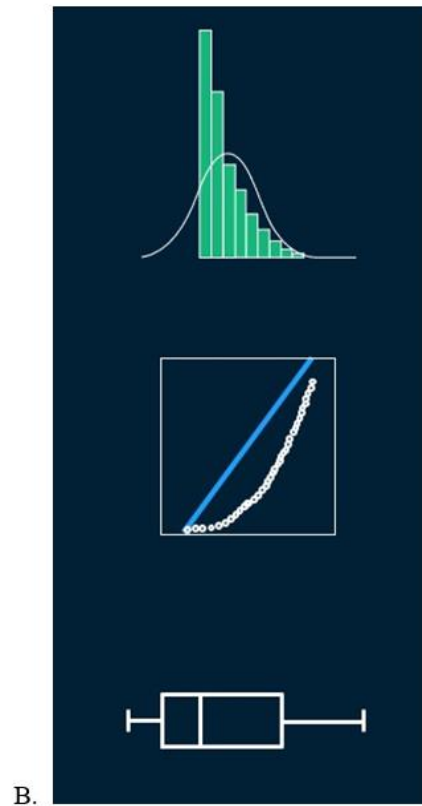
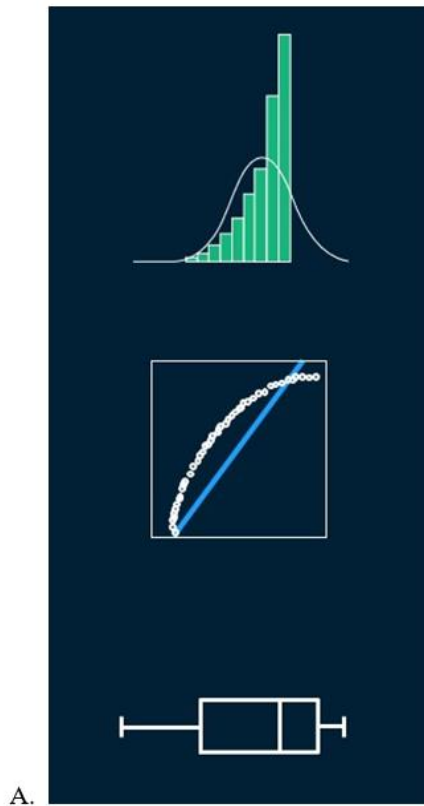
You are gathering data on products purchased by customers. What is the level of measurement of the brand of the product?

- A. Ratio
- B. Interval
- C. Ordinal
- D. Nominal

Answer: D

Question 3:

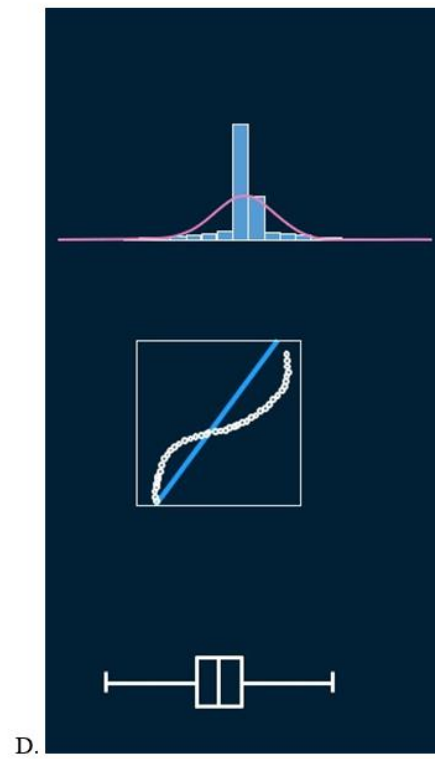
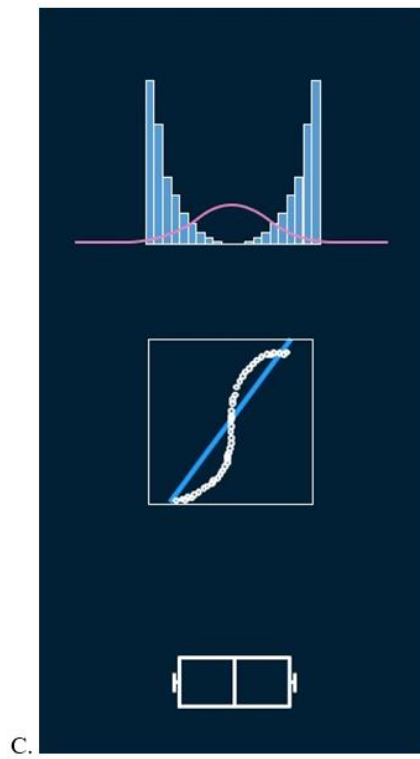
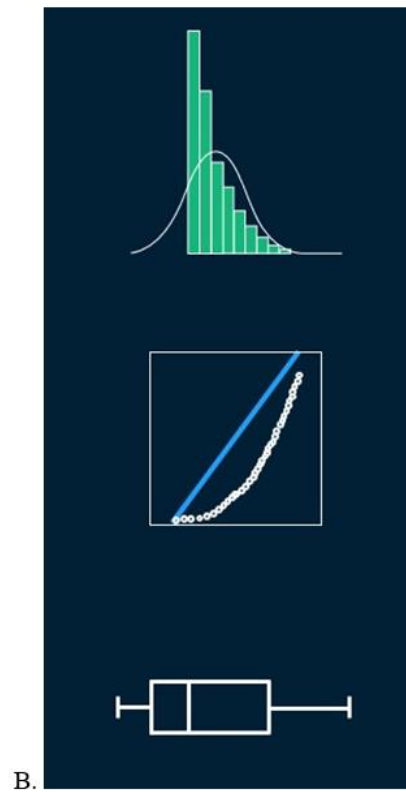
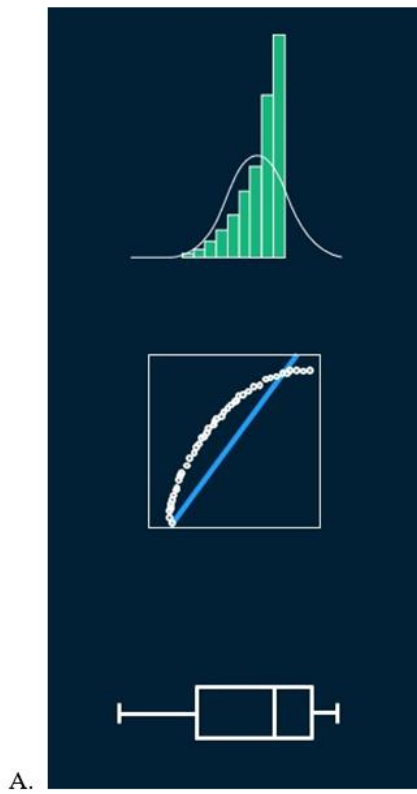
Which exhibit best represents a right-skewed distribution?



Answer: B

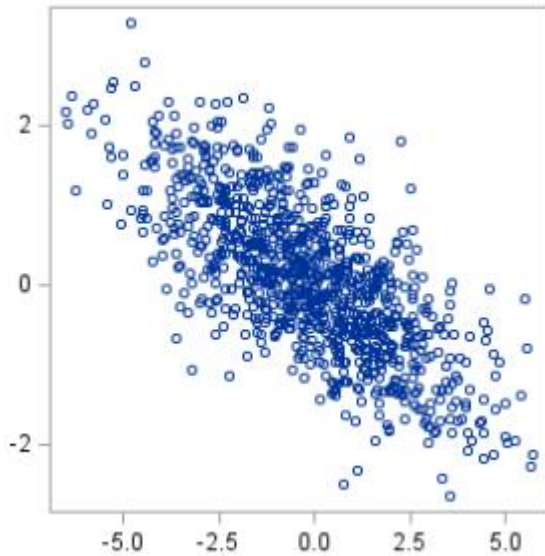
Question 4:

Which exhibit best represents a high positive kurtosis distribution?



Answer: D

Question 5:



Which terms best describe the scatter plot?

- A. Positive linear and strong relationship
- B. Positive linear and weak relationship
- C. Negative linear and strong relationship
- D. Negative linear and weak relationship

Answer: C

Question 6:

What statistics are commonly used for selecting variables during a sequential model selection process? (Choose 2.)

- A. F-statistic
- B. Bayesian Information Criterion (BIC)
- C. Mean Absolute Error (MAE)
- D. Significance Level

Answer: B, D

Question 7:

What types of input and target can be used for logistic regression?

- A. A continuous target and an ordinal input.
- B. An interval target and an ordinal input.
- C. A ratio target and a nominal input.
- D. An ordinal target and an interval input.

Answer: D

Question 8:

You built a logistic regression model to predict if a financial transaction is fraudulent (fraud = 1) or not. Given the values of inputs, you obtained a predicted probability of 0.20. What's the meaning of this predicted probability? (Choose 2.)

- A. There is a 80% chance that the transaction is fraudulent.
- B. There is a 20% chance that the transaction is non-fraudulent.
- C. There is a 80% chance that the transaction is non-fraudulent.
- D. There is a 20% chance that the transaction is fraudulent.

Answer: C, D

Question 9:

Which algorithm uses trial and error to discover which action yields the greatest rewards?

- A. Supervised Learning
- B. Unsupervised Learning
- C. Semi-supervised Learning
- D. Reinforcement Learning

Answer: D

Question 10:

Which of the following distinguishes between model parameters and hyperparameters?

- A. Model parameters govern the learning process and computational aspects, while hyperparameters primarily influence the model's predictions.
- B. The values of model parameters can significantly influence the selection of optimal hyperparameters, while hyperparameters do not directly impact the values of model parameters.
- C. Model parameters and hyperparameters are used interchangeably when referring to the factors affecting the performance of a model.
- D. Model parameters can be learned by the model through the training process, while hyperparameters must be set before training of the model or autotuned using optimization algorithms.

Answer: D