

## **304- Advanced Statistical Methods using R**

**1. What function in R is used to compute basic statistics like mean, median, and standard deviation?**

- A) mean()
- B) sum()
- C) sd()
- D) max()

**Answer: A) mean()**

**2 .In business hypothesis testing, what is the null hypothesis?**

- A) The hypothesis that is accepted
- B) The hypothesis that is rejected
- C) The hypothesis that is not related to business
- D) The hypothesis that there is no effect or difference

**Answer: D) The hypothesis that there is no effect or difference**

**3. What is logistic regression used for?**

- A) Predicting continuous outcomes
- B) Predicting categorical outcomes
- C) Calculating means of two samples
- D) Testing correlations for significance

**Answer: B) Predicting categorical outcomes**

**4. How is the significance of a correlation tested in R?**

- A) t-test
- B) z-test
- C) F-test
- D) Correlation test

**Answer: A) t-test**

**5. What is the purpose of ANOVA?**

- A) Comparing means of two samples
- B) Testing a proportion
- C) Summarizing data
- D) Analyzing variance among groups

**Answer: D) Analyzing variance among groups**

**6. What does Linear Regression in R aim to do?**

- A) Predict categorical outcomes
- B) Predict continuous outcomes
- C) Test correlations for significance
- D) Summarize data

**Answer: B) Predict continuous outcomes**

**7. What is a characteristic of Principal Components Analysis (PCA)?**

- A) It increases the number of variables
- B) It reduces the number of variables

C) It has no impact on dimensionality

D) It creates multicollinearity

**Answer: B) It reduces the number of variables**

**8. What is a fundamental concept in probability?**

A) Mean

B) Median

C) Probability Distributions

D) Standard Deviation

**Answer: C) Probability Distributions**

**9. What distribution is commonly used to model rare events?**

A) Normal distribution

B) Central Limit theorem

C) Poisson distribution

D) Binomial distribution

**Answer: C) Poisson distribution**

**10. What is the purpose of logistic regression?**

A) Predicting continuous outcomes

B) Predicting categorical outcomes

C) Analyzing variance among groups

D) Testing correlations for significance

**Answer: B) Predicting categorical outcomes**

**11. What is the dependent variable in Linear Regression?**

- A) Independent variable
- B) Explanatory variable
- C) Predictor variable
- D) Response variable

**Answer: D) Response variable**

**12. What is the Ordinary Least Squares (OLS) method used for in Linear Regression?**

- A) To maximize the sum of squared errors
- B) To minimize the sum of squared errors
- C) To ignore outliers
- D) To maximize multicollinearity

**Answer: B) To minimize the sum of squared errors**

**13. What is a potential issue in Linear Regression when predictor variables are highly correlated?**

- A) Outliers
- B) Multicollinearity
- C) Heteroscedasticity
- D) Autocorrelation

**Answer: B) Multicollinearity**

**14. What technique is used to reduce the dimensionality of data in Linear Regression?**

- A) Principal Components Analysis (PCA)
- B) Factor Analysis
- C) Linear Discriminant Analysis (LDA)
- D) Both A and B

**Answer: D) Both A and B**

**15. What is the purpose of evaluating assumptions in Linear Regression?**

- A) To maximize outliers
- B) To minimize multicollinearity
- C) To ensure the validity of model results
- D) To increase heteroscedasticity

**Answer: C) To ensure the validity of model results**

**16. What is the definition of probability?**

- A) The likelihood of an event occurring
- B) The sum of all outcomes
- C) The mean of a distribution
- D) The median of a distribution

**Answer: A) The likelihood of an event occurring**

**17. What is Bayes Theorem used for?**

- A) Calculating conditional probability
- B) Calculating marginal probability
- C) Estimating the mean of a distribution
- D) Estimating the standard deviation of a distribution

**Answer: A) Calculating conditional probability**

**18. What theorem states that as the sample size increases, the sample mean approaches the population mean?**

- A) Law of Large Numbers
- B) Central Limit Theorem
- C) Bayes Theorem
- D) Poisson Theorem

**Answer: B) Central Limit Theorem**

**19. What type of events cannot occur simultaneously?**

- A) Mutually Exclusive events
- B) Independent Events
- C) Conditional Events
- D) Marginal Events

**Answer: A) Mutually Exclusive events**

**20. What type of distribution models the number of successes in a fixed number of independent Bernoulli trials?**

- A) Normal distribution

- B) Central Limit theorem
- C) Poisson distribution
- D) Binomial distribution

**Answer: D) Binomial distribution**

**21. In multiple linear regression, what is the purpose of stepwise regression?**

- A) To include all variables in the model
- B) To exclude irrelevant variables from the model
- C) To ignore multicollinearity
- D) To increase the number of predictors

**Answer: B) To exclude irrelevant variables from the model**

**22. What is a key metric used to evaluate logistic regression models?**

- A) Mean squared error
- B) R-squared
- C) Log likelihood ratio
- D) F-statistic

**Answer: C) Log likelihood ratio**

**23. What does the ROC plot measure in logistic regression?**

- A) Sensitivity and specificity
- B) Mean squared error
- C) Variance

D) Bias

**Answer: A) Sensitivity and specificity**

**24. What is a common technique for dimension reduction in predictive modeling?**

A) Ordinary Least Squares (OLS)

B) Principal Components Analysis (PCA)

C) Central Limit Theorem

D) Poisson distribution

**Answer: B) Principal Components Analysis (PCA)**

**25. What is Linear Discriminant Analysis used for in predictive modeling?**

A) Reducing the number of variables

B) Identifying linear combinations of variables

C) Maximizing multicollinearity

D) Increasing model complexity

**Answer: B) Identifying linear combinations of variables**

**26. What is the primary purpose of decomposing a time series?**

A) To identify trends and seasonality

B) To increase heteroscedasticity

C) To maximize autocorrelation

D) To ignore outliers



**Answer: A) To identify trends and seasonality**

**27. What does the autocorrelation function (ACF) plot measure in time series analysis?**

- A) Seasonal variation
- B) Lagged correlation between observations
- C) Stationarity
- D) Multicollinearity

**Answer: B) Lagged correlation between observations**

**29. What is the purpose of the Holt-Winters Method in time series forecasting?**

- A) To identify trends
- B) To estimate seasonal variation
- C) To ignore autocorrelation
- D) To reduce model complexity

**Answer: B) To estimate seasonal variation**

**30. What models are part of the ARIMA family?**

- A) Autoregressive Moving Average Models
- B) Autoregressive Integrated Moving Average Models
- C) Autoregressive Moving Average Integrated Models
- D) Both A and B

**Answer: D) Both A and B**

**31. What is a characteristic of time series objects in R?**

- A) They can only have one column
- B) They can only have numerical values
- C) They must be stationary
- D) They contain time-stamped data

**Answer: D) They contain time-stamped data**