

## CENTRE FOR DEVELOPMENT OF ADVANCED COMPUTING Advanced Computing Training School



#### **PG-DBDA** Sample Paper Data analytics using statistics

$\sim$	3 T	- 4
1	NIA	
ι,	. 1 N ( )	. І

What is a hypothesis?

- A. A statement that the researcher wants to test through the data collected in a study
- B. A research question the results will answer
- C. A theory that underpins the study
- D. A statistical method for calculating the extent to which the results could have happened by chance

#### Q.No.2

Least Square Method uses \_\_\_\_.

- A. Linear polynomial
- B. Linear regression
- C. Linear sequence
- D. None of the mentioned above

#### Q. No. 3

Alternative Hypothesis is also called as?

- A. Null Hypothesis
- B. Research Hypothesis
- C. Simple Hypothesis
- D. None of the mentioned above

## Q. No. 4

Data Analysis is defined by the statistician?

- A. John Tukey
- B. Hans Peter Luhn
- C. Gregory Lon
- D. None of the mentioned above

## Q. No. 5

All pandas data structures are \_\_\_ mutable but not always \_\_\_\_ mutable.

- a) size, value
- b) semantic, size
- c) value, size
- d) none of the mentioned

## Q. No. 6

Which of the following object you get after reading CSV file?

- a) DataFrame
- b) Character Vector
- c) Panel
- d) All of the mentioned

#### Q. No. 7

Which of the following testing is concerned with making decisions using data?

a) Probability

PG-DBDA 1



# CENTRE FOR DEVELOPMENT OF ADVANCED COMPUTING Advanced Computing Training School



PG-DBDA Sample Paper Data analytics using statistics

b) Hypothesis
c) Causal
d) None of the mentioned
Q. No. 8  The expected value or of a random variable is the center of its distribution.  a) mode b) median c) mean d) bayesian inference
Q.No.9 Which of the following of a random variable is a measure of spread? a) variance b) standard deviation c) empirical mean d) all of the mentioned
Q.No.10  The square root of the variance is called the deviation.  a) empirical b) mean c) continuous d) standard
Q. No. 11 Which of the following inequality is useful for interpreting variances? a) Chebyshev b) Stautaory c) Testory d) All of the mentioned
Q. No. 12 Chebyshev's inequality states that the probability of a "Six Sigma" event is less than
a) 10% b) 20% c) 30% d) 3%
Q. No. 13 The function returns its argument with a modified shape, whereas the method modifies the array itself.  a) reshape, resize b) resize, reshape c) reshape2, resize

PG-DBDA 2

d) all of the mentioned



## CENTRE FOR DEVELOPMENT OF ADVANCED COMPUTING **Advanced Computing Training School**



## **PG-DBDA** Sample Paper Data analytics using statistics

$\sim$	TA T	1 4
"	No	1 /1
v.	No.	14

ndarray is also known as the alias array.

- a) True
- b) False

#### Q. No. 15

Which of the following things can be accomplished with linear model?

- a) Flexibly fit complicated functions
- b) Uncover complex multivariate relationships
- c) Build accurate prediction models
- d) All of the mentioned

#### O. No. 16

Which of the following statement is incorrect with respect to outliers?

- a) Outliers can have varying degrees of influence
- b) Outliers can be the result of spurious or real processes
- c) Outliers cannot conform to the regression relationship
- d) None of the mentioned

Q.No.17	
The	of a collection of data is the joint density evaluated as a function of the
parameters	with the data fixed.
a) probabili	ty
b) likelihoo	d
c) poisson d	listribution
d) all of the	mentioned
Q.No.18	
Which of th	e following object has a method cov to compute covariance between series?
a) Series	
b) DataFran	ne
c) Panel	
d) None of	the mentioned
Q. No. 19	
The	of the Chi-squared distribution is twice the degrees of freedom.
a) variance	
b) standard	deviation

- c) mode
- d) none of the mentioned

#### Q. No. 20

Gosset's distribution is invented by which of the following scientist?

- a) William Gosset
- b) William Gosling
- c) Gosling Gosset
- d) All of the mentioned

PG-DBDA 3