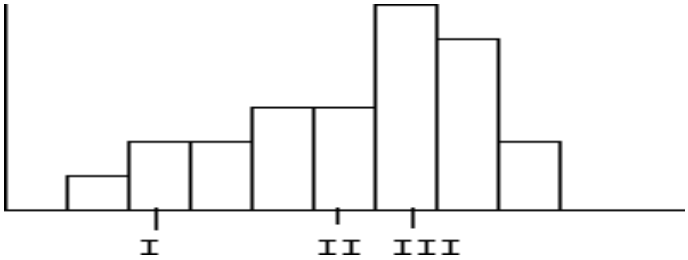
Department of Mathematics and Statistics

Question Bank of entry test for Admission to MPhil Statistics degree program

					Answer Key
1)	Colour and number of flowers is an example of				C
	A.	Quantitative and Qualitative variables	B.	Quantitative and discrete variables	
	C.	Qualitative and Quantitative variables	D.	Qualitative and Continuous variables	
2)	If Z (standard variable) score of a value is 1.5 it means value is				D
	A.	Value is 1.5 units below the mean	B.	Value is 1.5 times S below the mean	
	C.	Value is 1.5 units above the mean	D.	Value is 1.5 times S above the mean	
3)	Why do we use inferential statistics?				B
	A.	Inferential statistics are used to help us to show the difference between the sample and the whole population.	B.	Inferential statistics are used to help us to generalise from the sample to the whole population.	
	C.	Inferential statistics are used to help us to compare the sample to the whole population.	D.	All of the above apply to the use of inferential statistics.	
4)	In any normal distribution, the proportion of observations that are outside 1 standard deviations of the mean is closest to				B
	A.	0.05	B.	0.32	
	C.	0.95	D.	0.68	
5)	The heights in centimeters of 5 students are: 165, 175, 176, 159, 170. The sample median is				C
	A.	176	B.	165	
	C.	170	D.	159	
6)	In general, which of the following statements is FALSE?				B
	A.	The sample mean is more sensitive to extreme values than the median.	B.	The sample standard deviation is a measure of spread around the sample median	
	C.	The sample range is more sensitive to extreme values than the standard deviation.	D.	The sample standard deviation is a measure of variation around the mean	
7)	Which of the following measure may be ZERO				D
	A.	Mean	B.	Variance	
	C.	Range	D.	All of above	
8)	A sample of 99 distances has a mean of 24 feet . Unfortunately, it has just been discovered that an observation which was erroneously recorded as "35" actually had a value of "30". If we make this correction to the data, then:				B
	A.	the mean will increase	B.	the mean will decrease	
	C.	the mean will increase by 5	D.	the mean will decrease by 5	
9)	The term test scores of 15 students enrolled in a Business Statistics class were recorded in ascending order as follows: 4, 7, 7, 9, 10, 11, 13, 15, 15, 15, 17, 17, 19, 19, 20 After calculating the mean, median, and mode, an error is discovered: one of the 15's is really a 17. The measures of central tendency which will change are:				D
	A.	The mean only	B.	The median only	
	C.	The mode only	D.	The mean and mode	

10)	Suppose a frequency distribution is skewed with a median of $75.00 and a mode of $80.00. Which of the following is a possible value for the mean of distribution?				B
	A.	$86	B.	$64	
	C.	$91	D.	$77	
11)	Measurements of nine earthquakes gave the following readings: 4.5 L 5.5 H 8.7 8.9 6.0 H 5.2 where L indicates that the earthquake had an intensity below 4.0 and a H indicates that the earthquake had an intensity above 9.0. The median earthquake intensity of the sample is:				C
	A.	Cannot be computed because all of the values are not known	B.	8.70	
	C.	6.00	D.	5.75	
12)	For the following histogram, what is the proper ordering of the mean, median, and mode? Note that the graph is NOT numerically precise -only the relative positions are important.  				D
	A.	I = median II = mean III = mode	B.	I = median II = mean III = mode	
	C.	I =mode II = median III = mean	D.	I = mean II = median III = mode	
13)	Which of the following statements is NOT true?				D
	A.	In a symmetric distribution, the mean and the median are equal.	B.	In a symmetric distribution, the median is halfway between the first and the third quartiles.	
	C.	The first quartile is equal to the twenty-fifth percentile.	D.	The median is always greater than the mean.	
14)	Which of the following is FALSE:				D
	A.	The numbers 3, 3, 3 have a standard deviation of 0.	B.	The standard deviation is a measure of spread around the mean of the data	
	C.	The numbers 3, 4, 5 have the same standard deviation as 1003, 1004,1005	D.	The numbers 1, 5, 9 have a smaller standard deviation than 101, 105,109.	
15)	Many professional schools require applicants to take a standardized test. Suppose that 1000 students write the test, and you find that your mark of 63 (out of 100) was the 73rd percentile. This means :				B
	A.	At least 73% of the people got 63 or better.	B.	At least 270 people got 63 or better	
	C.	At least 270 people got 73 or better	D.	At least 27% of the people got 73 or worse.	
16)	A student discovers that his grade on a recent test was the 72^{nd} percentile. If 90 students wrote the test, then approximately how many students received a higher grade than he did?				C
	A.	65	B.	72	
	C.	25	D.	71	
17)	Measure of central tendency is also known as				A
	A.	Average	B.	Measure of variation	
	C.	Measure of spread	D.	Both (a) and (c)	

18)	The mean age of 50 students in a bus is 20 years. When the age of conductor is included the mean age is increased by one year. The age of the conductor is				C
	A.	51	B.	55	
	C.	71	D.	73	
19)	If the two observations are 5 and -5 their geometric mean is				D
	A.	5	B.	0	
	C.	-5	D.	Not possible to calculate	
20)	A frequency distribution having one mode is called				A
	A.	Uni-model	B.	Tri-model	
	C.	Bi-model	D.	No mode	
21)	What is one of the distinctions between a population parameter and a sample statistic?				C
	A.	A population parameter is only based on conceptual measurements, but a sample statistic is based on a combination of real and conceptual measurements	B.	A population parameter changes each time you try to measure it, but a sample statistic remains fixed across samples.	
	C.	A sample statistic changes each time you try to measure it, but a population parameter remains fixed.	D.	The true value of a sample statistic can never be known but the true value of a population parameter can be known	
22)	Which of the following is affected by outlier				A
	A.	Mean	B.	Median	
	C.	Q1	D.	Interquartile range	
23)	The sum of the relative frequencies for all classes will always equal				A
	A.	One	B.	Number of items in the study	
	C.	Number of classes	D.	100	
24)	Since the mode is the most frequently occurring data value, it				D
	A.	can never be larger than the mean	B.	is always larger than the mean	
	C.	is always larger than the median	D.	must have a value of at least two	
25)	In a fin Five number summary, which of the following is not used for data summarization?				C
	A.	the smallest value	B.	the 25th percentile	
	C.	the mean	D.	the median	
26)	For a -vely skewed data set which of the following is true				C
	A.	Mean > Median >Mode	B.	Mode < Median <Mean	
	C.	Mode > Median> Mean	D.	None of above	
27)	If all the values in a data set are different then relation between AM, GM and HM is				B
	A.	AM<GM<HM	B.	HM<GM<AM	
	C.	HM>GM>AM	D.	None of above	
28)	If all the values in a data set are same then relation between AM, GM and HM is				B
	A.	AM<GM<HM	B.	AM=GM=HM	
	C.	AM>GM>HM	D.	None of above	
29)	In frequency distribution we assume that each value in the class as at the				B
	A.	Lower boundary	B.	Class mark	

	C.	Upper boundary	D.	Class interval	
30)	HISTOGRAM for equal class interval is constructed by taking ---------on x-axix and -----on y-axix				D
	A.	Mid point , Frequency	B.	Frequency, Class boundaries	
	C.	Class boundaries , Mid point	D.	Class boundaries, Frequency	
31)	A frequency curve with right tail smaller than left tail is called				B
	A.	Symmetric	B.	Skewed to the left	
	C.	Skewed to the right	D.	Normal	
32)	For the data set value 0,0,0 -7,-7,8,8 the values of median and mode are				A
	A.	0 & 0	B.	0 & -7	
	C.	-7 & 0	D.	Not possible to calculate median and mode	
33)	For a data set 5,5,5,5 the values of mean. Median, Mode and standard deviation are				B
	A.	0,5.5,5 and 0	B.	5,5,5 and 0	
	C.	0,5.5,5 and 5	D.	None of above	
34)	Which of the following is not a measure of location				B
	A.	Mean	B.	Standard deviation	
	C.	Median	D.	Mode	
35)	The interquartile range is				D
	A.	the 50th percentile	B.	the difference between the largest and smallest values	
	C.	another name for the quartile deviation	D.	the difference between the upper quartile and the lower quartile	
36)	The variance can never be				B
	A.	Zero	B.	Negative	
	C.	larger than the standard deviation	D.	Both (a) and (c)	
37)	If 84% observations is a data set are less than mean+SD then it indicates that data is				B
	A.	Positively skewed	B.	Symmetric	
	C.	Negatively skewed	D.	None of above	
38)	If 25% observations is a data set are outside the interval mean±2SD then it indicates that data is				D
	A.	Positively skewed	B.	Bell shape Symmetric	
	C.	Negatively skewed	D.	Not bell shape symmetric	
39)	The suitable average for computing average speed of journay is				C
	A.	Geometric Mean	B.	Arithmetic mean	
	C.	Harmonic Mean	D.	Combined Mean	
40)	If mean of variable X is 100 , then the mean of Y=2X-200 will be				A
	A.	0	B.	200	
	C.	100	D.	2	
41)	The most suitable average for qualitative data is				B
	A.	Mean	B.	Mode	
	C.	Median	D.	Harmonic Mean	

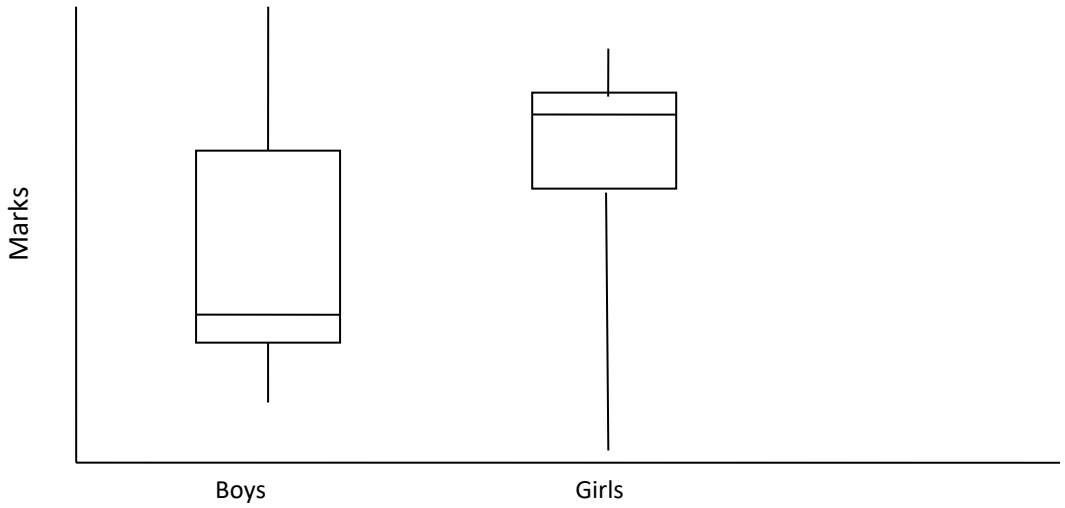
42)	The lower and upper quartiles of a symmetrical distribution are 40 and 60 respectively, then the value of median is				C
	A.	40	B.	60	
	C.	50	D.	20	
43)	The model letter(s) of the word STATISTICS is				B
	A.	S	B.	Both S and T	
	C.	T	D.	Both S and I	
44)	The sum of absolute deviations of the values is least when deviations are taken from				C
	A.	Mean	B.	Mode	
	C.	Median	D.	Q3	
45)	If the smallest observation of the data set is decreased then which of the average will not change				B
	A.	Harmonic Mean	B.	Median	
	C.	Mean	D.	Geometric mean	
46)	Questionnaire is a tool for collecting				C
	A.	Secondary data	B.	Published Data	
	C.	Primary Data	D.	Accurate Data	
47)	The range of -4, -20,-30,-44 and -36 will be				C
	A.	-48	B.	-40	
	C.	40	D.	Not possible to calculate	
48)	Which of the following have same units as units of original data				D
	A.	Mean	B.	Coefficient of Variation	
	C.	Standard Deviation	D.	Both (a) & (c)	
49)	Which of the following is unit free quantity				D
	A.	Coefficient of variation	B.	Coefficient of quartile deviation	
	C.	Z-scores	D.	All of above	
50)	Any measure of dispersion can never be				B
	A.	0	B.	Negative	
	C.	Positive	D.	Equal to 1	
51)	large value of standard deviation will indicate that				C
	A.	All the values are equal to mean	B.	The values are far-away from mean	
	C.	The values are close to mean	D.	All the values are positive	
52)	If standard deviation of the values 2,4,6,8 is 2.236, then standard deviation of the values 4,8,12,16 is				C
	A.	0	B.	19.998	
	C.	4.472	D.	2.236	
53)	The value of standard deviation changes by a change of				C
	A.	Origin	B.	Both origin and scale	
	C.	Scale	D.	Independent of Origin and scale	
54)	If median of the data 4,8,2 and y is 5.5 then value of y must be				A
	A.	7	B.	10	

	C.	1	D.	None of above	
55)	If mode of the data 2,2,2,3,3,,4,and y is 2 then which of the following value of y is not possible				A
	A.	3	B.	1	
	C.	4	D.	2	
56)	If mean marks of students A,B,C is 20 and means mean marks of students A,B,C,D,E is 25 then mean marks of D and E are				A
	A.	32.5	B.	50	
	C.	75	D.	25	
57)	If a variable X has units of measurements as meters then which of the following has no units of measurements				A
	A.	Z scores	B.	Standard deviation	
	C.	Variance	D.	Range	
58)	If variance of a data set is 25 then its standard deviation will be				A
	A.	5	B.	-5	
	C.	±5	D.	None of above	
59)	If every value of data set is increased by adding 10, which of the following remain the same				D
	A.	Median	B.	Range	
	C.	Standard deviation	D.	Both (b) and (c)	
60)	If every value of data set is multiplied by 10, then standard deviation will				D
	A.	Decrease by 100	B.	Increase by 10	
	C.	Decrease by 10	D.	Increase by 10 times	
61)	If standard deviation of the values 2,4,6,8 is 2.236, then standard deviation of -2,-4,-6,-8 is				B
	A.	0	B.	2.236	
	C.	-2.236	D.	Not possible to calculate	
62)	If Range and maximum value of a data set are 30 and 20 respectively then minimum value of the data is				B
	A.	0	B.	-10	
	C.	10	D.	Not possible to calculate	
63)	Find Mode for the data shoe size:- 6 7 8 9 No of students: 15 30 20 5				B
	A.	30	B.	7	
	C.	70	D.	Not possible to calculate	
64)	Geometric mean of two numbers 1/16 and 4/25 is				B
	A.	10	B.	1/10	
	C.	1/100	D.	100	
65)	Harmonic mean of two numbers 5/2 and 15/4 is				C
	A.	1/5	B.	1/3	
	C.	3	D.	5	
66)	If each value of a series is divided by 5, its coefficient of variation is reduced by				A
	A.	0 percent	B.	10 percent	
	C.	5 percent	D.	20 percent	
67)	If each value of a series is decease by 10, its coefficient of variation will be				A
	A.	Increased as compare to original value	B.	Remains the same	

	C.	Decreased as compare to original value	D.	None of above	
68)	If measurement units for data of height is in cm then measurement units of Z score will be				D
	A.	Cm	B.	Feet	
	C.	Cm^2	D.	unit less	
69)	If the data contains an extreme value, the suitable average is				C
	A.	Mean	B.	Weighted mean	
	C.	Median	D.	Geometric mean	
70)	Which of the following statements is always correct?				B
	A.	Mean = Median = Mode	B.	Median = $Q_2 = D_5 = P_{50}$	
	C.	Arithmetic mean = Geometric mean = Harmonic mean	D.	Mode = 2Median - 3Mean	
71)	If the geometric of the two numbers X1 and X2 is 9 if X1=3, then X2 is equal to:				A
	A.	27	B.	9	
	C.	3	D.	81	
72)	The following data represent the number of children per family, find total number of children				D
	A.	10	B.	100	
	C.	23	D.	51	
73)	Half of the difference between upper and lower quartiles is called:				C
	A.	Interquartile range	B.	Mean deviation	
	C.	Quartile deviation	D.	Standard deviation	
74)	The average of squared deviations from mean is called:				C
	A.	Mean deviation	B.	Standard deviation	
	C.	Variance	D.	Coefficient of variation	
75)	If $Y = aX \pm b$, where a and b are any two constants and $a \neq 0$, then Var (Y) is equal to:				D
	A.	a Var(X)	B.	a^2 Var(X) – b	
	C.	a Var(X) + b	D.	a^2 Var(X)	
76)	To compare the variation of two or more than two series, we use				B
	A.	Combined standard deviation	B.	Coefficient of variation	
	C.	Corrected standard deviation	D.	Coefficient of skewness	
77)	Simple bar char is appropriate to graph				D
	A.	Continuous variable	B.	Qualitative Variable	
	C.	Discrete Variable	D.	Both (b) and (c)	
78)	If $n = 5$ $AM = 20$, $\sum X^2 = 9000$ Then variance is				C
	A.	196	B.	96	

# children	0	1	2	3	4
#families	4	4	5	3	7

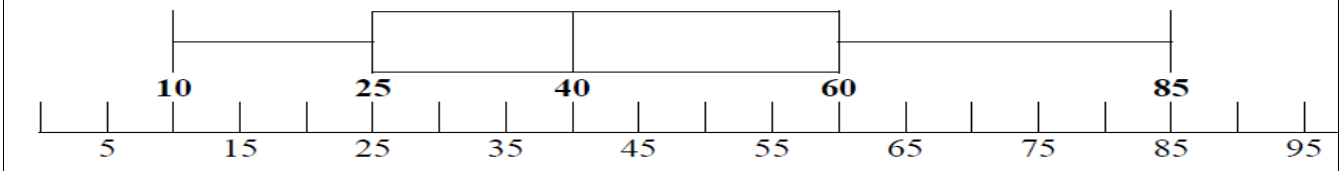
	C.	1400	D.	Not possible to calculate	
79)	If a student got 30 . 20 and 15 marks respectively in two class and one final test , find his mean marks in three tests if weightage of final test is double than each of class test				A
	A.	20	B.	25	
	C.	21.67	D.	15	
80)	If shape of the data is leptokurtic, then it must be				A
	A.	symmetric	B.	Negatively skewed	
	C.	Positively skewed	D.	Normal	
81)	Which of the following average can have more than one value				B
	A.	Harmonic mean	B.	Mode	
	C.	Median	D.	Geometric Mean	
82)	Statistics has two main branches				C
	A.	Test of hypothesis and estimation	B.	Mean and Variance	
	C.	Descriptive and inferential statistics	D.	Interval estimation and point estimation	
83)	Since the population size is always larger than the sample size, then the sample statistic				D
	A.	can never be larger than the population parameter	B.	can never be equal to the population parameter	
	C.	can never be smaller than the population parameter	D.	None of the statement is correct	
84)	If deviation of observations from mean are 3, 0 and -3, the value of biased sample variance is				C
	A.	0	B.	9	
	C.	6	D.	18	
85)	Which measurement scale is useful to measure colour of a flower				C
	A.	Ordinal	B.	Ratio	
	C.	Nominal	D.	interval	
86)	The appropriate average for calculating average percentage increase in population is				C
	A.	AM	B.	HM	
	C.	GM	D.	Mode	
87)	Harmonic mean of the data 1/3 , 0, 1/3 is				D
	A.	1/3	B.	3/2	
	C.	2/3	D.	Not possible to calculate	
88)	Find mean number of children per family No. of .families: 1, 2, 7 No. of children: 0, 1, 2				A
	A.	1.6	B.	5.33	
	C.	2	D.	6	
89)	According to the empirical rule, approximately what percent of the data should lie outside mean ±2SD				C
	A.	95%	B.	68%	
	C.	5%	D.	32%	
90)	You asked five of your classmates about their height. On the basis of this information, you stated that the average height of all students in your university or college is 67 inches. This is an example of:				C
	A.	Descriptive Statistics	B.	Order Statistics	
	C.	Inferential Statistics	D.	Official Statistics	
91)	The height of a student is 60 inches. This is an example of				C
	A.	Qualitative	B.	Discrete	
	C.	Continuous	D.	Categorical	
92)	Grades of students in a paper can be measured by using ----scale				C
	A.	Nominal	B.	Interval	
	C.	Ordinal	D.	Ratio	
93)	Which of the following is based on only two observations in the data				C

	A.	Quartile Deviation	B.	Median	
	C.	Range	D.	Both (a) and (c)	
94)	Which measure of variation is appropriate to measure variation between two data sets measured in different units				B
	A.	Standard deviation	B.	Coefficient of variation	
	C.	Range	D.	All can be used	
95)	If a students wants to collect data from university of Agriculture students regarding services of different mobile companies, population of his study will be				B
	A.	All the students from University of Agriculture	B.	All the mobile user students from University of Agriculture	
	C.	All mobile users in pakistan	D.	All may be population of the above study	
96)	The variance of 5 numbers is 10. If each number is divided by 2, then variance of new number is_____				B
	A.	20	B.	2.5	
	C.	5	D.	0	
97)	If 5X+5Y=40 then arithmetic mean of X and Y is				A
	A.	4	B.	10	
	C.	20	D.	Not possible to calculate	
98)	If any value in the data is less than zero, then which of the following can not be calculated				C
	A.	Arithmatic mean	B.	Harmonic mean	
	C.	Geometric Mean	D.	Both (b) and (C)	
99)	What is an other name of cumulative frequency polygon graph				B
	A.	Frequency Curve	B.	Ogive	
	C.	Cumulative curve	D.	Accumulate	
100)	Graph of five number summary is called				B
	A.	Summary graph	B.	Box whisker plot	
	C.	Bar graph	D.	Line graph	
	Consider following statement for next <u>10</u> questions: The following box plots represent the entry test marks obtained by boys and girls  				

101)	Lowest marks obtained by one of the				C
	A.	Boy	B.	May be boy or girl	
	C.	Girl	D.	None of above	
102)	Data for marks of boys is ---------- as compare to data for marks of girls				B
	A.	More consistent	B.	More Variable	
	C.	Less variable	D.	None of above	
103)	Boys marks are on the average ----------girls marks				C
	A.	Higher than	B.	Both are equal	
	C.	Lower than	D.	Not possible without actual data	
104)	Data for boys is ----------skewed, but girls data is ----------skewed				
	A.	Positively , Positively	B.	Negatively , Positively	
	C.	Positively , Negatively	D.	Negatively, Negatively	
105)	What percent of the values are blow than upper edge of the box				C
	A.	25%	B.	50%	
	C.	75%	D.	100%	
106)	What percent of the values are above than lower edge of the box				C
	A.	25%	B.	50%	
	C.	75%	D.	100%	
107)	What percent of the values are within box				B
	A.	25%	B.	50%	
	C.	75%	D.	100%	
108)	Length of the box represent				C
	A.	Range	B.	Quartile deviation	
	C.	Interquartile range	D.	Mean deviation	
109)	Length of the graph represent				A
	A.	Range	B.	Quartile deviation	
	C.	Interquartile range	D.	Mean deviation	
110)	Position of the line within the box indicates				A
	A.	Shewness of the data	B.	Outlier in the data	
	C.	Kurtosis of the data	D.	None	
	Statement for the next <u>10</u> questions: **A bell shape symmetrical data has mean 50 and standard deviation 10**				
111)	What is the percentage of values lie between 40 and 60				A
	A.	68%	B.	75%	
	C.	95%	D.	89%	
112)	What is the percentage of values lie between 20 and 80				D
	A.	68%	B.	75%	
	C.	95%	D.	99.73%	
113)	What is the percentage of values lie between 30 and 70				C
	A.	68%	B.	68%	
	C.	95%	D.	98%	
114)	What is the percentage of values lie below 70				C
	A.	68%	B.	68%	
	C.	97.5%	D.	95%	

115)	What is the percentage of values lie above 60				B
	A.	68%	B.	16%	
	C.	97.5%	D.	32%	
116)	What is the percentage of values lie below 40				B
	A.	68%	B.	16%	
	C.	97.5%	D.	32%	
117)	What is the percentage of values lie between 30 and 60				A
	A.	81.5%	B.	16%	
	C.	97.5%	D.	32%	
118)	What is the percentage of values lie below 20 or above 80				C
	A.	27%	B.	99.73%	
	C.	0.27%	D.	95%	
119)	What percent of the values below 50				A
	A.	50%	B.	75%	
	C.	25%	D.	90%	
120)	What is the percentage of values lie below 20				D
	A.	27%	B.	0.05%	
	C.	0.27%	D.	0.135%	
121)	Any value which is not consistent with rest of data value is called				A
	A.	Outlier	B.	Normal	
	C.	Abnormal	D.	Order Statistic	
122)	If a data has only two values with variance 9, what is the range of the data				B
	A.	12	B.	6	
	C.	21	D.	Cannot be determine	
123)	Which of the following measure is effected by outlier				A
	A.	Range	B.	Median	
	C.	Quartile Deviation	D.	Deciles	
124)	How much percent the values above 7^{th} decile (D7)				A
	A.	70%	B.	90%	
	C.	30%	D.	10%	
125)	There are ---------values that divides the data into 10 equal parts				B
	A.	10	B.	9	
	C.	8	D.	2	
126)	Measure of central tendency is also called				
	A.	Measure of spread	B.	Average	
	C.	Measure of Location	D.	Both (b) and (c)	
127)	If mean and median of a data re 10 and 30 respectively, find mode of the data				B
	A.	20	B.	70	
	C.	40	D.	Cannot be determine	
128)	If mean and median of a data re 10 and 30 respectively, discuss shape of the data				B
	A.	Positively shewed	B.	Negatively Shewed	
	C.	Symmetrical	D.	May be positive or Negative	
129)	If lower and upper quartiles of a data bell shape symmetrical data are 10 and 40 respectively, find median of the data				B
	A.	20	B.	25	
	C.	30	D.	15	
130)	If lower and upper quartiles of a data bell shape symmetrical data are 10 and 40 respectively, find mode of the data				B
	A.	20	B.	25	
	C.	30	D.	15	
131)	Value of coefficient of skewness less than zero indicates that data is				B

	A.	Positively shewed	B.	Negatively Shewed	
	C.	Symmetrical	D.	May be positive or Negative	
132)	Value of coefficient of skewness greater than zero				A
	A.	Positively shewed	B.	Negatively Shewed	
	C.	Symmetrical	D.	May be positive or Negative	
133)	Value of coefficient of skewness equal than zero				C
	A.	Positively shewed	B.	Negatively Shewed	
	C.	Symmetrical	D.	May be positive or Negative	
134)	Value of coefficient of skewness greater than zero				A
	A.	Right tail of frequency curve is greater than left tail	B.	Left tail of frequency curve is greater than right tail	
	C.	Both tails are equal	D.	Cannot be determined	
135)	With the help of three quartiles we can calculate				A
	A.	Coefficient of Skewness	B.	Coefficient of variation	
	C.	Coefficient of Kurtosis	D.	Coefficient of mean deviation	
136)	With the help of three quartiles we can calculate				B
	A.	Coefficient of Skewness	B.	Coefficient of quartile deviation	
	C.	Coefficient of Kurtosis	D.	Both (b) and (c)	
137)	If bowling speed of a bowler is measured in km per hour then units of coefficient of variation is				D
	A.	Km per minute	B.	Meter per minute	
	C.	Km per hour	D.	Unit free	
138)	Which of the following measure has different units has units of original data				D
	A.	Mean	B.	Standard deviation	
	C.	Median	D.	Variance	
139)	Average of absolute deviation from mean is called				D
	A.	Average Deviation	B.	Quartile deviation	
	C.	Mean deviation	D.	Both (a) and (c)	
140)	Average of Squared deviation from mean is called				B
	A.	Average Deviation	B.	Variance	
	C.	Mean deviation	D.	Standard deviation	
141)	Which of the graph is useful to identify shape of the data				D
	A.	Frequency Curve	B.	Cumulative Frequency polygon	
	C.	Histogram	D.	Both (a) and (c)	
142)	Which of the graph is useful to estimate median and quantile of the data				B
	A.	Frequency Curve	B.	Cumulative Frequency polygon	
	C.	Histogram	D.	Both (a) and (c)	
143)	Which of the graph is useful to estimate mode of the data				C
	A.	Frequency Curve	B.	Cumulative Frequency polygon	
	C.	Histogram	D.	Both (a) and (c)	
144)	Which of the graph is useful to estimate mean of the data				D
	A.	Frequency Curve	B.	Cumulative Frequency polygon	
	C.	Histogram	D.	none	
145)	Which measure of central tendency is use for both qualitative as well as quantitative data				B
	A.	Mean	B.	Mode	
	C.	Median	D.	Standard Deviation	
146)	Which measure of central tendency may have more than one answer				B
	A.	Mean	B.	Mode	
	C.	Median	D.	Standard Deviation	
147)	Which measure of central tendency may not have any answer				B
	A.	Mean	B.	Mode	
	C.	Median	D.	Standard Deviation	
148)	Which measure measures variation in the data relative to size of the data				D

	A.	Coefficient of range	B.	Coefficient of mean deviation	
	C.	Coefficient of Quartile Deviation	D.	Coefficient of variation	
149)	Which measure measures consistency in the data				D
	A.	Coefficient of range	B.	Coefficient of mean deviation	
	C.	Coefficient of Quartile Deviation	D.	Coefficient of variation	
150)	Small value of which measure indicates that values are close to mean in the data				C
	A.	Range	B.	Median	
	C.	Standard Deviation	D.	Mode	
	Consider the following Box plot for the next <u>10</u> questions 				
151)	Maximum value in the data is				B
	A.	95	B.	85	
	C.	60	D.	40	
152)	Minimum value in the data is				C
	A.	5	B.	25	
	C.	10	D.	0	
153)	Median of the data is				C
	A.	42.5	B.	35	
	C.	40	D.	60	
154)	Interquartile quartile range of the data is				B
	A.	42.5	B.	35	
	C.	40	D.	60	
155)	Range of the data is				C
	A.	90	B.	95	
	C.	75	D.	20	
156)	Length of the box measures				D
	A.	Average	B.	kurtosis	
	C.	skewness	D.	variation	
157)	Position of line within box measures				C
	A.	Average	B.	kurtosis	
	C.	skewness	D.	variation	
158)	Shape of the data is				A
	A.	Positively shewed	B.	Negatively Shewed	
	C.	Symmetrical	D.	May be positive or Negative	
159)	Shape of the data is-------kurtic				D
	A.	Lepto	B.	Platy	
	C.	Meso	D.	Cannot be determine	
160)	Mean of data is				D
	A.	40	B.	35	
	C.	42.5	D.	Cannot be determine	
161)	Which coefficient measures the sharpness of the peak of a symmetrical curve				D
	A.	Measure of central tendency	B.	Measure of skewness	
	C.	Measure of variation	D.	Measure of kurtosis	
162)	Sharp peck symmetrical curve is called ------				A
	A.	Lepto	B.	Platy	
	C.	Meso	D.	Normal	

163)	Flat peak symmetrical curve is called------				B
	A.	Lepto	B.	Platy	
	C.	Meso	D.	Normal	
164)	Another name of bell shape normal curve is-------				C
	A.	Lepto	B.	Platy	
	C.	Meso	D.	Normal	
165)	A variable whose mean is zero and variance is one is called				C
	A.	Normal	B.	Unit variable	
	C.	Standard variable	D.	Extreme variable	
166)	A variable X has mean 50 and standard deviation 5, if standard value of an observation of X is 2 find its X value				C
	A.	45	B.	65	
	C.	60	D.	70	
167)	If standard value of a variable is 2 this means value is				A
	A.	Above the mean	B.	At the mean	
	C.	Below the mean	D.	Cannot be determine	
168)	Which of the following is not member of Five number summary				D
	A.	Mean	B.	Range	
	C.	Standard deviation	D.	All are not member of five number summary	
169)	If shape of the data is bell shape normal, which of the following statement must be true				C
	A.	Data is positively skewed	B.	Data is negatively skewed	
	C.	Data is symmetric	D.	Data is positively or negative skewed	
170)	In order to convert a variable X into standard variable we require				
	A.	Mean of X	B.	Both mean and variance	
	C.	Variance of X	D.	Standard deviation od X	
171)	Find unbiased sample variance of the data 1,2,3,4,5				B
	A.	5	B.	2.5	
	C.	2	D.	3	
172)	If variance of the data 2,4,6,8,10 is 8, what is the variance of 102,104,106,108,110				C
	A.	108	B.	4	
	C.	8	D.	6	
173)	If variance of the data 2,4,6,8,10 is 8, what is the variance of -2 ,- 4,-6 ,-8 ,-10				C
	A.	108	B.	-108	
	C.	8	D.	-8	
174)	If variance of the data 2,4,6,8,10 is 8, what is the variance of 4 ,8,12 ,16 ,20				B
	A.	108	B.	32	
	C.	8	D.	16	
175)	If mean and variance of a variable X are 2 and 4 respectively, then standard deviation of Y=2-2X will be				B
	A.	-2	B.	4	
	C.	2	D.	-4	
176)	If mean of a variable increase from 80 to 100, find percentage increase in mean				A
	A.	25%	B.	30%	
	C.	20%	D.	35%	
177)	In calculating average deviation, deviations can be taken from				D
	A.	Mean	B.	Mode	
	C.	Median	D.	From all	
178)	If mean and variance of a variable X are 2 and 4 respectively, then find mean of Y=2-2X will be				A
	A.	-2	B.	4	
	C.	2	D.	-4	
179)	If smallest value of the data is decrease by 2 then which of the following measure will remain the same				B
	A.	Standard Deviation	B.	Median	

	C.	Mean	D.	Range	
180)	Which of the following measure is affected by presence of outlier in the data				D
	A.	Standard Deviation	B.	Range	
	C.	Mean	D.	All are affected by outlier	
181)	If a student got 70 and 60 marks in three tests, what marks he require to get mean marks of 60 in three tests				A
	A.	50	B.	70	
	C.	65	D.	75	
182)	Any numerical quantity calculated by taking sample from population is called				B
	A.	Mean	B.	Statistic	
	C.	Parameter	D.	Both (a) and (b)	
183)	Considering set of observations, the percentage of the values lie within mean ±2SD is				B
	A.	95%	B.	75%	
	C.	68%	D.	89%	
184)	If the arithmetic mean is multiplied by coefficient of variation then resulting value is classified as				B
	A.	Coefficient of deviation	B.	Standard deviation	
	C.	Coefficient of mean	D.	Quartile deviation	
185)	If mean deviation of a set of observations is 8.5 the value of quartile deviation is				A
	A.	7.08	B.	8.08	
	C.	10.08	D.	Cannot be determine	
186)	The kurtosis defines the peakness of the curve in the region which is				A
	A.	Around the mode	B.	Around the mean	
	C.	Around the median	D.	Around standard deviation	
187)	Find mode of the data 1,2,2,2,0,0,0				B
	A.	2	B.	2 and 0 both	
	C.	0	D.	No mode	
188)	Find median of the data 1,2,2,2,0,0,0				C
	A.	0	B.	2	
	C.	1	D.	No median	
189)	Which graph is useful to estimate median				B
	A.	Frequency Curve	B.	Cumulative Frequency polygon	
	C.	Frequency Polygon	D.	All can be used to estimate median	
190)	For the data -1 ,0, 1,2 which of the following cannot be calculated				C
	A.	AM	B.	variance	
	C.	GM	D.	Both (b) and (c)	

					Answer Key
191)	Two variables are said to be uncorrelated if				C
	A.	they tend to change together in the opposite direction	B.	they tend to change together in the same direction	
	C.	when they tend to change with no connection to each other	D.	none of other three	
192)	Two variables are said to be negatively correlated if				A
	A.	they tend to change together in the opposite direction	B.	they tend to change together in the same direction	
	C.	when they tend to change with no connection to each other	D.	none of other three	
193)	Which of the following statements is the property of simple correlation coefficient,				D
	A.	The range of the correlation coefficient is -1 to +1	B.	Correlation coefficient is symmetrical with respect to variables	
	C.	Correlation coefficient is independent of units of measurements	D.	All others	
194)	The correlation coefficient is used to determine				C
	A.	A specific value of the y-variable given a specific value of the x-variable	B.	A specific value of the x-variable given a specific value of the y-variable	
	C.	The strength of the relationship between the x and y variables	D.	None of these	
195)	If there is a very strong correlation between two variables then the correlation coefficient must be				C
	A.	any value larger than 1	B.	much smaller than 0, if the correlation is negative	
	C.	much larger than 0, regardless of whether the correlation is negative or positive	D.	None of these alternatives is correct	
196)	SSE can never be				A
	A.	larger than SSTotal	B.	smaller than SSTotal	
	C.	equal to 1	D.	equal to zero	
197)	Regression modeling is a statistical framework for developing a mathematical equation that describes how				C
	A.	one explanatory and one or more response variables are related	B.	several explanatory and several response variables response are related	
	C.	one response and one or more explanatory variables are related	D.	All of these are correct	
198)	In least squares regression, which of the following is not a required assumption about the error term ε?				A
	A.	The expected value of the error term is one.	B.	The variance of the error term is the same for all values of x.	
	C.	The values of the error term are independent.	D.	The error term is normally distributed	
199)	In a regression analysis if $R^2 = 1$, then				B
	A.	SSE must also be equal to one	B.	SSE must be equal to zero	
	C.	SSE can be any positive value	D.	SSE must be negative	
200)	In regression analysis, the variable that is used to explain the change in the outcome of an experiment, or some natural process, is called				D

	A.	the independent variable	B.	the predictor variable	
	C.	the explanatory variable	D.	all of the others are correct	
201)	In the case of a mathematical model for a straight line, if a value for the x variable is specified, then				A
	A.	the exact value of the response variable can be computed	B.	the computed response to the independent value will always give a minimal residual	
	C.	the computed value of y will always be the best estimate of the mean response	D.	none of these alternatives is correct	
202)	A regression analysis between sales (in Rs 1000) and price (in Rupees) resulted in the following equation: $\hat{Y}$ = 5,000 - 8X. The above equation implies that an				A
	A.	increase of 1Rs in price is associated with a decrease of 8 Rs in sales	B.	increase of 8Rs in price is associated with an increase of 5,000 Rs in sales	
	C.	increase of 1Rs in price is associated with a decrease of 42,00 Rs in sales	D.	increase of 1Rs in price is associated with a decrease of 8000 Rs in sale	
203)	In regression analysis, if the independent variable is measured in kilograms, the dependent variable				D
	A.	must also be in kilograms	B.	must be in some unit of weight	
	C.	cannot be in kilograms	D.	can be any units	
204)	Suppose the correlation coefficient between height (as measured in feet) versus weight (as measured in pounds) is 0.40. What is the correlation coefficient of height measured in inches versus weight measured in ounces? [12 inches = one foot; 16 ounces = one pound]				A
	A.	0.40	B.	0.30	
	C.	0.533	D.	cannot be determined from information given	
205)	Assume the same variables as in question 28 above; height is measured in feet and weight is measured in pounds. Now, suppose that the units of both variables are converted to metric (meters and kilograms). The impact on the slope is				C
	A.	the sign of the slope will change	B.	the magnitude of the slope will change	
	C.	both a and b are correct	D.	neither a nor b are correct	
206)	A residual plot:				B
	A.	displays residuals of the explanatory variable versus residuals of the response variable.	B.	displays explanatory variable versus residuals of the response variable	
	C.	displays the explanatory variable versus the response variable	D.	displays the explanatory variable on the y axis versus the response variable on the x axis	
207)	When the regression line passes through the origin then				C
	A.	The regression coefficient is zero	B.	Correlation is zero	
	C.	Intercept is zero	D.	None of the others	
208)	The range of partial correlation coefficient is				C
	A.	0 to 1	B.	-1 to 0	
	C.	-1 to +1	D.	None of the others	
209)	The range of multiple correlation coefficient is				A
	A.	0 to 1	B.	-1 to 0	
	C.	-1 to +1	D.	None of the others	
210)	A regression analysis is inappropriate when				D
	A.	you have two variables that are measured on an interval or ratio scale.	B.	you want to make predictions for one variable based on information about another variable.	
	C.	the pattern of data points forms a reasonably straight line.	D.	there is heteroscedasticity in the scatter plot.	

211)	In regression analysis, the variable that is being predicted is				B
	A.	the independent variable	B.	the dependent variable	
	C.	usually denoted by x	D.	usually denoted by r	
212)	If the slope of the regression equation y = bo + b1x is positive, then				B
	A.	as x increases y decreases	B.	as x increases so does y	
	C.	Either a or b is correct	D.	as x decreases y increases	
213)	A residual is defined as				B
	A.	The difference between the actual Y values and the mean of Y.	B.	The difference between the actual Y values and the predicted Y values.	
	C.	The predicted value of Y for the average X value.	D.	The square root of the slope	
214)	A linear regression (LR) analysis produces the equation Y = 0.4X + 3. This indicates that				D
	A.	When Y = 0.4, X = 3	B.	When Y = 0, X = 3	
	C.	When X = 3, Y = 0.4	D.	When X = 0, Y = 3	
215)	If the t ratio for testing the significance of slope of a simple linear regression equation is -2.58 and the critical values of the t distribution at the 1% and 5% levels, respectively, are 3.499 and 2.365, then the slope is				D
	A.	not significantly different from zero	B.	significantly different from zero at both the 1% and the 5% levels.	
	C.	significantly different from zero at the 1% level but not at the 5% level.	D.	significantly different from zero at the 5% level but not at the 1% level.	
216)	Ordinary least squares is used to estimate a linear relationship between a firm's total revenue per week (in $1,000s) and the average percentage discount from list price allowed to customers by salespersons. A 95% confidence interval on the slope is calculated from the regression output. The interval ranges from 1.05 to 2.38. Based on this result, the researcher				D
	A.	can conclude that the slope is significantly different from zero at the 5% level of significance.	B.	can be 95% confident that the effect of a 1% increase in the average price discount will increase weekly total revenue by between $1,050 and $2,380.	
	C.	has one chance in twenty of incorrectly concluding that the slope is within the estimated confidence interval.	D.	All of the above are correct.	
217)	Multiple regression analysis is used when				B
	A.	there is not enough data to carry out simple linear regression analysis	B.	the dependent variable depends on more than one independent variable.	
	C.	one or more of the assumptions of simple linear regression are not correct.	D.	the relationship between the dependent variable and the independent variables cannot be described by a linear function.	
218)	The adjusted value of the coefficient of determination				D
	A.	will always increase if additional independent variables are added to the regression model	B.	is equal to the proportion of the sum of the squared deviations of the dependent variable from its mean that is explained by the regression model	
	C.	is always greater than the proportion of the sum of the squared deviations of the dependent variable from its mean that is explained by the regression model.	D.	is always less than the proportion of the sum of the squared deviations of the dependent variable from its mean that is explained by the regression model.	
219)	If the F test statistic for a regression is greater than the critical value from the F distribution, it implies that				C

	A.	none of the independent variables in the regression model have a significant effect on the dependent variable.	B.	all of the independent variables in the regression model have significant effects on the dependent variable.	
	C.	one or more of the independent variables in the regression model have a significant effect on the dependent variable.	D.	None of the above is correct.	
220)	The standard error of the regression measures the				C
	A.	variability of the independent variable(s) relative to its (their) mean	B.	variability of the dependent variable relative to its mean	
	C.	variability of the dependent variable relative to the regression line.	D.	average error that will result if the regression line is used to predict.	
221)	Multicollinearity refers to a situation in which				B
	A.	successive error terms derived from the application of regression analysis to time series data are correlated	B.	there is a high degree of correlation between the independent variables included in a multiple regression model	
	C.	the dependent variable is highly correlated with the independent variable(s) in a regression analysis	D.	the application of a multiple regression model yields estimates that are nonlinear in form	
222)	Autocorrelation refers to a situation in which				A
	A.	successive error terms derived from the application of regression analysis to time series data are correlated.	B.	there is a high degree of correlation between two or more of the independent variables included in a multiple regression model.	
	C.	the dependent variable is highly correlated with the independent variable(s) in a regression analysis.	D.	the application of a multiple regression model yields estimates that are nonlinear in form.	
223)	Heteroskedasticity refers to a situation in which the error terms from a regression analysis				A
	A.	do not have equal variance	B.	are not normally distributed.	
	C.	do not have a mean of zero.	D.	All of the above are correct.	
224)	The Durbin-Watson statistic is used to test for				B
	A.	Multicollinearity	B.	autocorrelation	
	C.	heteroskedasticity	D.	All of the above are correct.	
225)	Autocorrelation may be the result of				D
	A.	the omission of an important explanatory variable	B.	the presence of a trend in the independent variable.	
	C.	nonlinearities in the relationship between the dependent and independent variables	D.	All of the above are correct.	
226)	The relation between F and t test statistics ($F=t^2$) for testing the hypothesis that there is no relation between x and y exists only if				B
	A.	regression degrees of freedom is equal to two	B.	in simple linear regression	
	C.	in multiple linear regression	D.	in polynomial regression	
227)	For a data set for X, Y, and Z as: X= 5, 10, 15, 20. Y=3, 5, 6, 8. And Z=9, 8, 6, 2. The sign of simple linear correlation coefficient between X and Z is				A
	A.	negative	B.	positive	
	C.	not possible to determine	D.	may be positive	

228)	For a data set for Y, X1, and X2 as: Y= 5, 10, 15, 20. X1=3, 5, 6, 8. And X2=9, 8, 6, 2. While performing the multiple linear regression, the partial regression coefficients are (b1=2.74) and (b2=-0.238), what is the value of intercept coefficient?				B
	A.	1.1	B.	-1.1	
	C.	1.74	D.	2.74	
229)	For a data set for Y, X1, and X2 as: Y= 5, 10, 15, 20. X1=3, 5, 6, 8. And X2=9, 8, 6, 2. While performing the multiple linear regression, the regression mean sum of squares is equals to 61.607 and total sum of squares is 125. What is the error sum of squares?				B
	A.	-1.786	B.	1.786	
	C.	3.786	D.	2.786	
230)	For a data set for Y, X1, and X2 as: Y= 5, 10, 15, 20. X1=3, 5, 6, 8. And X2=9, 8, 6, 2. While performing the multiple linear regression, the regression sum of squares is equals to 123.214 and Total sum of squares is 125. What is the F-ratio for regression?				D
	A.	14.50	B.	38.50	
	C.	24.50	D.	34.50	
231)	For a data set for Y, X1, and X2 as: Y= 5, 10, 15, 20. X1=3, 5, 6, 8. And X2=9, 8, 6, 2. While performing the multiple linear regression, the error sum of squares is equals to 1.786 and Total sum of squares is 125. What is the coefficient of determination?				A
	A.	98.57%	B.	24.50%	
	C.	38.57%	D.	94.50%	
232)	In the regression equation Y=10-5X, the slope coefficient is				C
	A.	-10	B.	10	
	C.	-5	D.	-4	
233)	The measurement units for the slope coefficient is equals to				B
	A.	X/Y units	B.	Y/X units	
	C.	Y units only	D.	X units only	
234)	The y-intercept is interpreted as:				B
	A.	it is the average value of y when x=0. And if range of x not contain zero, then b0 has no interpretation	B.	it is the average value of x when y=0. And if range of y not contain zero, then b0 has no interpretation	
	C.	it is the average value of y when x=0. And if range of x contain zero, then b0 has no interpretation	D.	it is the value of y when x=0. And if range of x not contain zero, then b0 has no interpretation	
235)	Following is not a measure of the goodness of fit of the estimated model				D
	A.	Coefficient of determination	B.	Standard error of estimate	
	C.	Residual plot	D.	Scatter plot	
236)	Following one is not the type of Linear Regression				A
	A.	Intrinsically Linear Regression	B.	Curvi Linear Regression	
	C.	Simple Linear Regression	D.	Multiple Linear Regression	
237)	Following one is the type of non-Linear Regression				A
	A.	Intrinsically Linear Regression	B.	Curvi Linear Regression	
	C.	Simple Linear Regression	D.	Multiple Linear Regression	
238)	The coefficient of determination equals to 85% is explained as				B
	A.	about 15% variation in dependent variable has been explained by the linear relationship with X	B.	about 85% variation in dependent variable has been explained by the linear relationship with X	
	C.	about 15% variation in independent variable has been explained by the linear relationship with Y	D.	about 15% variation in independent variable has been explained by the linear relationship with Y	

239)	The 90% confidence interval for population correlation coefficient is explained as				D
	A.	If we take 100 samples of the same size under the same conditions and compute 100 C.I's about parameter, then 10 such C.Is will contain the parameter	B.	If we take 100 samples of the same size under the same conditions and compute 90 C.I's about parameter, then 90 such C.Is will contain the parameter	
	C.	If we take 10 samples of the same size under the same conditions and compute 100 C.I's about parameter, then 100 such C.Is will contain the parameter	D.	If we take 100 samples of the same size under the same conditions and compute 100 C.I's about parameter, then 90 such C.Is will contain the parameter	
240)	The multiple correlation coefficient measures the				A
	A.	strength of linear relationship between one dependent variable and joint effect of all the independent variables	B.	actual linear relationship between all the independent variables	
	C.	actual relationship between two variables while keeping the effect of other variables as constant	D.	strength of linear relationship between many dependent variables and independent variables	
241)	For a data set for X and Y as: X= 7, 6, 4, 5, 3, 2. and Y=10, 20, 30, 40, 50, 60. While performing simple linear regression, the error degrees of freedom in ANOVA table is				C
	A.	5	B.	3	
	C.	4	D.	1	
242)	While performing simple linear regression, testing the hypothesis that there is no linear relation between X and Y can be performed by using				D
	A.	t-test only	B.	t-test and z-test	
	C.	only F-test	D.	t-test and F-test	
243)	For a data set for X and Y as: X= 2, 3, 5, 4, 6, 7 and Y=1, 2, 3, 4, 5, 6. While performing simple linear regression, the Regression SS=15.557 and Total SS=17.5, What will be the F-ratio for regression?				D
	A.	35.03	B.	30.03	
	C.	22.03	D.	32.03	
244)	For a data set for X and Y as: X= 2, 3, 5, 4, 6, 7 and Y=1, 2, 3, 4, 5, 6. While performing simple linear regression, the Regression SS=15.557 and Total SS=17.5, What will be the coefficient of determination?				B
	A.	87.9%	B.	88.9%	
	C.	90.9%	D.	86.9%	
245)	For a data set for X and Y as: X= 42, 33, 25, 24, 16, 7 and Y=1, 12, 23, 14, 8, 3. What is nature of relationship between these two variables?				B
	A.	Simple linear regression	B.	Curvi-linear regression	
	C.	Non-linear regression	D.	Multiple linear regression	
246)	The measurement units for the y-intercept coefficient is equals to				C
	A.	X/Y units	B.	Y/X units	
	C.	Y units only	D.	X units only	
247)	The interpretation of b1 is:				C
	A.	it is the change in average value of x for one unit increase in y	B.	it is the average value of x when y=0. And if range of y not contain zero, then b1 has no interpretation	
	C.	it is the change in average value of y for one unit increase in x	D.	it is the average value of y when x=0. And if range of x not contain zero, then b1 has no interpretation	
248)	Following is used to see the dependence of one variable on one or more variables				B
	A.	Correlation	B.	Regression	

	C.	Sample	D.	Population	
249)	What does the Y intercept (b_0) represent?				A
	A.	The predicted value of Y when X = 0.	B.	The estimated change in average Y per unit change in X.	
	C.	The predicted value of Y.	D.	The mean value of X.	
250)	What does the least squares method do exactly?				C
	A.	Minimizes the distance between the data points	B.	Finds the least problematic regression line	
	C.	Finds those (best) values of the intercept and slope that provide us with the smallest value of the residual sum of squares	D.	Finds those (best) values of the intercept and slope that provide us with the smallest value of the sum of residuals	
251)	Which of the following measures is optimal for comparing the goodness of the fit of competing regression models involving the same dependent variable?				D
	A.	The intercept	B.	The slope coefficient	
	C.	R-square	D.	Standard deviation of the residuals	
252)	What does the following expression (H_0: $\beta_1 = \beta_2 = 0$) mean?				C
	A.	One of the independent variables is useful in predicting the dependent variable	B.	Both of the independent variables are useful in predicting the dependent variable	
	C.	None of the independent variables is useful in predicting the dependent variable	D.	There is a third independent variable predicting the dependent variable	
253)	Which of the following criteria is the most optimal for assessing the goodness of the fit of a multiple linear regression model?				A
	A.	Adjusted R-square	B.	Intercept	
	C.	R-square	D.	Slope	
254)	In which cases are the standardized coefficients suggested to be used to identify the relative importance of the independent variables in a multiple regression model?				B
	A.	When all the independent variables are measured using the same metric	B.	When not all the independent variables are measured using the same metric	
	C.	When not all the dependent variables are measured using the same metric	D.	When all the independent variables are measured using an ordinal scale ranging from 1 to 6	
255)	Which one of these statements is not a Gauss-Markov assumption?				B
	A.	That the error term has a conditional mean of zero	B.	Absence of influential observations.	
	C.	That the error term has constant variance	D.	That the errors are uncorrelated	
256)	Why should we not include irrelevant variables in our regression analysis?				B
	A.	Your R-squared will become too high	B.	Because of data limitations	
	C.	It is bad academic fashion not to base your variables on sound theory	D.	We increase the risk of producing false significant results	
257)	How can we deal with the gap of the assumption of linearity?				A
	A.	Include a squared term	B.	Include an interaction term	
	C.	Use robust regression	D.	Use the margins command	
258)	Which statistics can help us detect multicollinarity?				A
	A.	Variance inflation factor (VIF)	B.	Durbin-Watson	
	C.	F-Statistics	D.	Tolerance values +1	
259)	What does heteroskedasticity mean?				D

	A.	The variance in the residuals are the same regardless of their predicted values.	B.	There is variance in the residuals	
	C.	That we are unable to produce residuals	D.	The variance in the residuals differ depending on their predicted values	
260)	What is the way to check for heteroskedasticity?				A
	A.	We can examine a plot of predicted values vs the residuals	B.	We can run the Hausman test	
	C.	We can see the plot of X against Y	D.	We can compare the *F*-test of two models	
261)	Which one is not a measure of influential (or potentially influential) units?				B
	A.	Leverage	B.	Cook-Weisberg	
	C.	DFBETA	D.	Cook's distance	
262)	The units of correlation coefficient are				C
	A.	Percent	B.	Non existing	
	C.	Kg/ml	D.	None of the above	
263)	Which of the following statements is NOT true regarding linear regression?				B
	A.	It identifies significant predictors for a continuous outcome variable.	B.	It predicts the outcome of a binary variable with continuous variables.	
	C.	It quantifies a relationship between two continuous variables.	D.	It models a linear relationship between two continuous variables	
264)	If the correlation coefficient from linear regression is 0.64. How much of the variation of the Y axis variable is explained by the X axis variable?				C
	A.	64%	B.	6.4%	
	C.	41%	D.	32%	
265)	What is the intercept coefficient of regression equation $\hat{Y} = 8x - 3$				C
	A.	3	B.	-8	
	C.	-3	D.	8	
266)	What is the slope of regression line $\hat{Y} = -8x + 3$				B
	A.	8	B.	-8	
	C.	3	D.	-3	
267)	In a simple regression analysis (where Y is a dependent and X an independent variable), if the Y-intercept is positive, then:				D
	A.	there is a positive correlation between X and Y.	B.	there is a negative correlation between X and Y.	
	C.	if X is increased, Y must also increase.	D.	None of the above answers is correct	
268)	What is the Predicted Y-value if X=10 from regression equation $\hat{Y} = 8x - 3$				D
	A.	-3	B.	80	
	C.	50	D.	77	
269)	What is the slope of regression line $\hat{Y} = 5x + 3$				A
	A.	5	B.	-3	
	C.	3	D.	None of the others	
270)	A study was made on the effect of temperature on the yield of a chemical process. The following data were recorded: Temperature = -5, -3, 0, 1, 4, 6 and Yield= 10, 8, 9, 7, 5, 1. What type of relation exists between X and Y?				A
	A.	Simple linear	B.	Non Linear	
	C.	Curvilinear	D.	Can't be determined	

271)	A study was made on the effect of temperature on the yield of a chemical process. The following data were recorded: Temperature = -5, -3, 0, 1, 4, 6 and Yield= 2, 5, 19, 7, 5, 1. What type of relation exists between X and Y?				A
	A.	Curvi-Linear	B.	Non-Linear	
	C.	Simple Linear	D.	No relation	
272)	A study was made on the effect of temperature on the yield of a chemical process. The following data were recorded: Temperature= -5, -3, 0, 1, 4, 6 and Yield= 2, 5, 19, 7, 5, 1. What will be the sign of b_2 coefficient while fitting second degree curve to this data?				C
	A.	Positive	B.	May be positive or negative	
	C.	Negative	D.	Can't be determined	
273)	A study was made on the effect of temperature on the yield of a chemical process. The following data were recorded: Temperature= -5, -3, 0, 1, 4, 6 and Yield= 2, 5, 19, 7, 5, 1. What will be the sign of b_1 coefficient while fitting second degree curve to this data?				C
	A.	Negative	B.	May be positive or negative	
	C.	Positive	D.	None of the others	
274)	A scatterplot is				D
	A.	one-dimensional graph of randomly scattered data	B.	two-dimensional graph of a straight line	
	C.	two-dimensional graph of a curved line	D.	two-dimensional graph of data values.	
275)	Correlation and regression are concerned with				B
	A.	the relationship between two categorical variables.	B.	the relationship between two quantitative variables.	
	C.	the relationship between a quantitative explanatory variable and a categorical response variable	D.	the relationship between a categorical explanatory variable and a quantitative response variable	
276)	What is the effect of an outlier on the value of a correlation coefficient?				C
	A.	An outlier will always decrease a correlation coefficient	B.	An outlier will always increase a correlation coefficient.	
	C.	An outlier might either decrease or increase a correlation coefficient, depending on where it is in relation to the other points.	D.	An outlier will have no effect on a correlation coefficient.	
277)	Which of the following is a deterministic relationship?				C
	A.	The relationship between hair color and eye color.	B.	The relationship between father's height and son's height.	
	C.	The relationship between height in inches and height in centimeters.	D.	The relationship between height as determined with a ruler and height as determined by a tape measure.	
278)	The regression equation between father's heights and student's heights for the 79 male students is $\hat{y} = 30 + .58x$. One student was 72 inches tall; his father's height was 65 inches. What is the estimated height for this student?				A
	A.	67.7	B.	72	
	C.	30	D.	65	
279)	The regression equation between father's heights and student's heights for the 79 male students is $\hat{y} = 30 + .58x$. One student was 72 inches tall; his father's height was 65 inches. What is the residual for this student?				D
	A.	7	B.	65	
	C.	30	D.	4.3	
280)	The correlation between father's heights and student's heights for the 79 male students is r = 0.9. What is the proportion of variation explained by father's heights?				D
	A.	0.90	B.	0.448	

	C.	0.79	D.	0.810	
281)	Which of the following is a possible value of R-squared and indicates the strongest linear relationship between two quantitative variables?				B
	A.	-90%	B.	80%	
	C.	0%	D.	120%	
282)	The correlation between two variables is given by r = 0.0. What does this mean?				A
	A.	The best straight line through the data is horizontal	B.	There is a perfect positive relationship between the two variables	
	C.	There is a perfect negative relationship between the two variables.	D.	All of the points must fall exactly on a horizontal straight line.	
283)	Which one of the following statements involving correlation is possible and reasonable?				B
	A.	The correlation between hair color and eye color is 0.80.	B.	The correlation between the height of a father and the height of his first son is 0.6	
	C.	The correlation between left foot length and right foot length is 2.35.	D.	The correlation between hair color and age is positive.	
284)	A regression between foot length (response variable in cm) and height (explanatory variable in inches) for 33 students resulted in the following regression equation: $\hat{y} = 10.9 + 0.23\ x$. One student in the sample was 73 inches tall with a foot length of 29 cm. What is the predicted foot length for this student?				B
	A.	17.57 cm	B.	27.69 cm	
	C.	0.23 cm	D.	10.9 cm	
285)	A regression between foot length (response variable in cm) and height (explanatory variable in inches) for 33 students resulted in the following regression equation: $\hat{y} = 10.9 + 0.23\ x$. What is the slope coefficient?				C
	A.	0.23 inches/cm	B.	10.9 cm/inches	
	C.	0.23 cm/inches	D.	10.9 inches/cm	
286)	A regression between foot length (response variable in cm) and height (explanatory variable in inches) for 33 students resulted in the following regression equation: $\hat{y} = 10.9 + 0.23\ x$. What is the estimated average foot length for students who are 70 inches tall?				A
	A.	27 cm	B.	10.9 cm	
	C.	28 cm	D.	30 cm	
287)	In the simple linear regression equation, the symbol $\hat{y}$ represents the				A
	A.	average or predicted response	B.	estimated intercept	
	C.	estimated slope	D.	explanatory variable	
288)	Which one of the following is NOT appropriate for studying the relationship between two quantitative variables?				B
	A.	Scatterplot	B.	Bar chart	
	C.	Correlation	D.	Regression	
289)	Significance of multiple correlation coefficient is tested by				D
	A.	t-test	B.	Chi-square test	
	C.	Z-test	D.	F-test	
290)	Significance of partial correlation coefficient is tested by				C
	A.	F-test	B.	Z-test	
	C.	t-test	D.	None of the others	
291)	If a Durbin Watson statistic takes a value close to zero, what will be the value of the first order autocorrelation coefficient?				B

	A.	Close to zero	B.	Close to plus one	
	C.	Close to minus one	D.	Close to either minus one or plus one	
292)	Suppose that the Durbin Watson test is applied to a regression containing two explanatory variables plus a constant with 50 data points. The test statistic takes a value of 1.53. What is the appropriate conclusion?				D
	A.	Residuals appear to be positively autocorrelated	B.	Residuals appear to be negatively autocorrelated	
	C.	Residuals appear not to be autocorrelated	D.	The test result is inconclusive	
293)	Which of the following are plausible approaches to dealing with a model that exhibits heteroscedasticity? i) Take logarithms of each of the variables ii) Use suitably modified standard errors iii) Use a generalized least squares procedure iv) Add lagged values of the variables to the regression equation.				C
	A.	(ii) and (iv) only	B.	(i) and (iii) only	
	C.	(i), (ii), and (iii) only	D.	(i), (ii), (iii), and (iv)	
294)	What would be then consequences for the OLS estimator if heteroscedasticity is present in a regression model but ignored?				C
	A.	It will be biased	B.	It will be inconsistent	
	C.	It will be inefficient	D.	All of (A), (B) and (C) will be true	
295)	If OLS is used in the presence of autocorrelation, which of the following will be likely consequences? i) Coefficient estimates may be misleading ii) Hypothesis tests could reach the wrong conclusions iii) Forecasts made from the model could be biased iv) Standard errors may inappropriate				A
	A.	(ii) and (iv) only	B.	(i) and (iii) only	
	C.	(i), (ii), and (iii) only	D.	(i), (ii), (iii), and (iv)	
296)	Near multicollinearity occurs when				D
	A.	Two or more explanatory variables are perfectly correlated with one another	B.	The explanatory variables are highly correlated with the error term	
	C.	The explanatory variables are highly correlated with the dependent variable	D.	Two or more explanatory variables are highly correlated with one another	
297)	Which one of the following is NOT an example of mis-specification of functional form?				D
	A.	Using a linear specification when y scales as a function of the squares of x	B.	Using a linear specification when a double-logarithmic model would be more appropriate	
	C.	Modelling y as a function of x when in fact it scales as a function of $1/x$	D.	Excluding a relevant variable from a linear regression model	
298)	Which one of the following is NOT a plausible remedy for near multicollinearity?				D

	A.	Use principal components analysis	B.	Drop one of the collinear variables	
	C.	Use a longer run of data	D.	Take logarithms of each of the variables	
299)	Which of the following could result in autocorrelated residuals? i) Slowness of response of the dependent variable to changes in the values of the independent variables ii) Over-reactions of the dependent variable to changes in the independent variables iii) Omission of relevant explanatory variables that are autocorrelated iv) Outliers in the data				C
	A.	(ii) and (iv) only	B.	(i) and (iii) only	
	C.	(i), (ii), and (iii) only	D.	(i), (ii), (iii), and (iv)	
300)	If the residuals from a regression estimated using a small sample of data are not normally distributed, which one of the following consequences may arise?				D
	A.	The coefficient estimates will be unbiased but inconsistent	B.	The coefficient estimates will be biased but consistent	
	C.	The coefficient estimates will be biased and inconsistent	D.	Test statistics concerning the parameters will not follow their assumed distributions	
301)	In multiple linear regression, the square root of the mean square error is called				B
	A.	Standard deviation of Y	B.	Standard error of estimate	
	C.	Correlation coefficient	D.	Standard deviation of X	
302)	Homogeneity of three or more correlation coefficients is tested by				B
	A.	F-test	B.	Chi-square test	
	C.	Z-test	D.	t-test	
303)	For a data set for X and Y as: X= 2, 3, 5, 2, 3, 1 and Y=1, 2, 3, 4, 5, 6. While performing simple linear regression, the estimated regression equation is y-hat = 4.64 - 0.429 x. what is the predicted y-value at X=1				C
	A.	4.64	B.	0.429	
	C.	4.211	D.	-0.429	
304)	For a data set for X and Y as: X= 2, 3, 5, 2, 3, 1 and Y=1, 2, 3, 4, 5, 6. While performing simple linear regression, the estimated regression equation is y-hat = 4.64 - 0.429 x. what is the predicted y-value at X=2				A
	A.	3.782	B.	0.429	
	C.	4.64	D.	37.82	
305)	The correlation coefficient between X and Y is (-0.313) with a p-value (0.515). What conclusion can be about the significance of the correlation coefficient?				C
	A.	Highly significant	B.	Significant	
	C.	Non-significant	D.	None	

306)	For a data set for X and Y as: X= 2, 3, 5, 2, 6, 9 and Y=1, 2, 3, 4, 5, 6. While performing simple linear regression, the estimated regression equation is y-hat = 1.04 + 0.547 X. The regression SS is (11.207). What is the Error SS?				B
	A.	4.293	B.	6.293	
	C.	7.293	D.	5.293	
307)	For a data set for X and Y while performing simple linear regression, the mean square error is 9, what is the value of standard deviation of regression?				C
	A.	9	B.	6	
	C.	3	D.	0	
308)	The correlation coefficient between X and Y is (-0.813) with a p-value (0.01). What conclusion can be about the significance of the correlation coefficient?				B
	A.	Non-significant	B.	Significant	
	C.	Negative but non-significant	D.	Positive and significant	
309)	For a data set for X and Y as: X= 2, 3, 5, 2, 6, 9 and Y=1, 2, 3, 4, 5, 6. While performing simple linear regression, the estimated regression equation is y-hat = 1.04 + 0.547 X. The regression SS is (11.207). What is the Total SS?				C
	A.	1.75	B.	13.207	
	C.	17.50	D.	14.207	
310)	For a data set for X and Y as: X= 2, 3, 5, 2, 3, 1 and Y=1, 2, 3, 4, 5, 6. While performing simple linear regression, the estimated regression equation is y-hat = 4.64 - 0.429 x. The p-value for b1 is 0.546. What is your conclusion about relationship between X and Y?				B
	A.	Linear	B.	Non-significant	
	C.	Non-Linear	D.	Significant	
311)	While performing multiple linear regression on a data set of 20 observations having one dependent and five independent variables, what will be the error degrees of freedom in testing the significance of multiple regression coefficients?				C
	A.	4	B.	5	
	C.	15	D.	19	
312)	The correlation coefficient between X and Y is (0.313) with a p-value (0.515). What conclusion can be about the significance of the correlation coefficient?				C
	A.	significant	B.	Highly significant	
	C.	Non- significant	D.	None	
313)	For a data set for X and Y as: X= 2, 3, 5, 2, 6, 9 and Y=1, 2, 3, 4, 5, 6. While performing simple linear regression, the estimated regression equation is y-hat = 1.04 + 0.547 X. The regression SS is (11.207). What is the F-ratio for regression?				A

	A.	7.12	B.	6.12	
	C.	8.15	D.	5.12	
314)	While performing multiple linear regression on a data set of 15 observations having one dependent and three independent variables, what will be the degrees of freedom for t-test in testing the significance of partial regression coefficients?				C
	A.	10	B.	14	
	C.	11	D.	2	
315)	While performing multiple linear regression on a data set of 20 observations having one dependent and five independent variables, what will be the degrees of freedom for t-test in testing the significance of partial regression coefficients?				A
	A.	14	B.	5	
	C.	19	D.	15	
316)	For a data set for X and Y as: X= 2, 3, 5, 2, 6, 9 and Y=1, 2, 3, 4, 5, 6. While performing simple linear regression, the estimated regression equation is y-hat = 1.04 + 0.547 X. The regression SS is (11.207). What is the Mean square error?				B
	A.	5.733	B.	1.5733	
	C.	2.333	D.	2.5733	
317)	For a data set for X and Y as: X= 2, 3, 4, 5, 6, 9 and Y=11, 10, 8, 6, 5, 1. What type of relation exists between X and Y				A
	A.	Negative	B.	Linear	
	C.	Positive	D.	Non-linear	
318)	For a data set for X and Y as: X= 12, 13, 14, 16, 17, 15 and Y=1, 2, 3, 4, 5, 6. While performing simple linear regression, the estimated slope coefficient is (0.829) and y-intercept is (-8.51). If the TotalSS=17.50 and MSE=1.371, what is the coefficient of determination?				D
	A.	38.65%	B.	8.65%	
	C.	28.65%	D.	68.65%	
319)	While performing multiple linear regression on a data set of 20 observations having one dependent and five independent variables, what will be the regression degrees of freedom in testing the significance of multiple regression coefficients?				D
	A.	19	B.	15	
	C.	14	D.	4	
320)	For a data set for X and Y as: X= 2, 3, 4, 5, 6, 9 and Y=11, 10, 8, 6, 5, 1. What type of relation exists between X and Y				B
	A.	Positive and strong	B.	Negative and linear	
	C.	Negative and weak	D.	Positive and weak	

321)	For a data set for X and Y as: X= 12, 13, 14, 16, 17, 15 and Y=1, 2, 3, 4, 5, 6. While performing simple linear regression, the estimated slope coefficient is (0.829). What is the value of Y-intercept?				C
	A.	8.51	B.	0.729	
	C.	-8.51	D.	85.1	
322)	A scatterplot shows:				A
	A.	Scores on one variable plotted against scores on a second variable.	B.	The frequency with which values appear in the data	
	C.	The average value of groups of data.	D.	The proportion of data falling into different categories.	
323)	For a data set for X and Y as: X= 12, 13, 14, 16, 17, 15 and Y=1, 2, 3, 4, 5, 6. While performing simple linear regression, the estimated slope coefficient is (0.829) and y-intercept is (-8.51). If the TotalSS=17.50 and MSE=1.371, what is the f-ratio for regression?				C
	A.	9.76	B.	10.76	
	C.	8.76	D.	4.76	
324)	For a data set for X and Y as: X= 12, 13, 14, 16, 17, 15 and Y=1, 2, 3, 4, 5, 6. While performing simple linear regression, the estimated slope coefficient is (0.829) and y-intercept is (-8.51). If the TotalSS=17.50 and MSE=1.371, what is the value of regression mean square?				A
	A.	12.014	B.	10.014	
	C.	13.014	D.	6.014	
325)	For a data set for X and Y as: X= 12, 13, 14, 16, 17, 15 and Y=1, 2, 3, 4, 5, 6. While performing simple linear regression, the estimated slope coefficient is (0.829) and y-intercept is (-8.51). If the TotalSS=17.50 and MSE=1.371, what is the value of regression SS?				D
	A.	10.014	B.	14.014	
	C.	11.014	D.	12.014	
326)	The relationship between two variables partialling out the effect that a third variable has on one of those variables can be expressed using a:				B
	A.	Partial correlation	B.	Semi-partial correlation	
	C.	Bivariate correlation	D.	Point-biserial correlation	
327)	When interpreting a correlation coefficient, it is important to look at:				D
	A.	The +/– sign of the correlation coefficient.	B.	The magnitude of the correlation coefficient.	
	C.	The significance of the correlation coefficient.	D.	All of these.	
328)	For a data set for X and Y as: X= 12, 13, 14, 16, 17, 15 and Y=1, 2, 3, 4, 5, 6. While performing simple linear regression, what is the error degrees of freedom in ANOVA table?				B
	A.	1	B.	4	
	C.	6	D.	3	

329)	Which of the following is not an assumption for simple linear regression?				B
	A.	Normally distributed variables	B.	Multicollinearity	
	C.	Linear relationship	D.	Constant variance	
330)	Continuous predictors influence the _____ of the regression line, while categorical predictors influence the _____.				A
	A.	slope, intercept	B.	intercept, slope	
	C.	R^2, *p*-value	D.	*p*-value, R^2	
331)	Which of the following is true about the adjusted R^2?				C
	A.	It is usually larger than the R^2	B.	It is only used when there is just one predictor	
	C.	It is usually smaller than the R^2	D.	It is used to determine whether residuals are normally distributed	
332)	Significance for the multiple regression coefficients is determined by				A
	A.	F-test	B.	R-square	
	C.	Correlation coefficient	D.	t-test	
333)	What is the primary purpose of Pearson's and Spearman's correlation coefficients?				A
	A.	Examining the relationship between two non-categorical variables	B.	Identifying deviations from normality for continuous variables	
	C.	Examining the relationship between two categorical variables	D.	Comparing means across group	
334)	Which of the following would be considered a very strong negative correlation?				C
	A.	0.89	B.	-0.09	
	C.	-0.89	D.	0.09	
335)	Which test is used to determine whether a correlation coefficient is statistically significant?				C
	A.	Paired samples *t*-test	B.	Chi-squared test	
	C.	One-sample *t*-test	D.	Z-test	
336)	The value of Durbin Watson lies between				C
	A.	-4 and 4	B.	0 and 1	
	C.	0 and 4	D.	-1 and 0	
337)	In the least squares line, $\sum(Y - \hat{Y})$ is				B
	A.	Minimum	B.	Maximum	
	C.	Least	D.	Zero	
338)	Significance for the partial regression coefficients is determined by				D
	A.	F-test	B.	R-square	
	C.	Correlation coefficient	D.	t-test	

339)	In the least squares line, $\sum(Y - \hat{Y})^2$ is				D
	A.	Maximum	B.	Zero	
	C.	One	D.	Least	
340)	Which of the following assumptions are required to show the consistency, unbiasedness and efficiency of the OLS estimator? i) $E(u_t) = 0$ ii) $Var(u_t) = \sigma^2$ iii) $Cov(u_t, u_{t-j}) = 0 \; \forall \; j$ iv) $u_t \sim N(0, \sigma^2)$				C
	A.	(ii) and (iv) only	B.	(i) and (iii) only	
	C.	(i), (ii), and (iii) only	D.	(i), (ii), (iii), and (iv)	

					Answer Key
341)	In Sampling with replacement, a sampling unit can be selected				B
	A.	only once	B.	More than once	
	C.	Less than once	D.	None of these	
342)	Which one of the following is not probability sampling?				D
	A.	Simple random sampling	B.	Systematic sampling	
	C.	Stratified sampling	D.	Judgment sampling	
343)	List of all the units of the population is called:				C
	A.	Random sampling	B.	Bias	
	C.	Sampling frame	D.	Probability sampling	
344)	Probability distribution of a statistic is called:				D
	A.	Sampling	B.	Parameter	
	C.	Data	D.	Sampling distribution	
345)	When N=5, n=2 then the number of all possible samples that can be drawn with replacement are				D
	A.	5	B.	7	
	C.	10	D.	25	
346)	A Population of 16 observations having standard deviation 6 and mean 10. The variance of the sampling distribution of mean when a sample of size 4 is taken by with replacement				C
	A.	1.5	B.	5	
	C.	9	D.	4.23	
347)	A population of 5 members is 2, 9, 6, 5, 3. If we construct a sampling distribution of mean by taking samples of size 2 by without replacement what would be the mean of sampling distribution of means.				B
	A.	5.2	B.	5	
	C.	4.8	D.	7	
348)	From the following sampling methods, which is a probability sampling method?				C
	A.	Judgement	B.	Quota	
	C.	Simple random	D.	Convenience	
349)	Which among the following is the benefit of using simple random sampling?				D
	A.	The results are always representative.	B.	Interviewers can choose respondents freely.	
	C.	Informants can refuse to participate.	D.	We can calculate the accuracy of the results.	
350)	Increasing the sample size has the following effect upon the sampling error?				B
	A.	It increases the sampling error	B.	It reduces the sampling error	
	C.	It has no effect on the sampling error	D.	All of the above	
351)	Which of the following is not a type of non-probability sampling?				D

	A.	Quota sampling	B.	Convenience sampling	
	C.	Snowball sampling	D.	Stratified random sampling	
352)	Sample is a subset of?				D
	A.	Data	B.	Set	
	C.	Distribution	D.	Population	
353)	The difference between a statistic and the parameter is called:				C
	A.	Non-random	B.	Probability	
	C.	Sampling error	D.	Random	
354)	The distribution that is formed by taking all possible values of a statistics is known as:				C
	A.	Hypergeometric distribution	B.	Normal distribution	
	C.	Sampling distribution	D.	Binomial distribution	
355)	Among these, which sampling is based on equal probability?				A
	A.	Simple random sampling	B.	Stratified random sampling	
	C.	Systematic sampling	D.	Probability sampling	
356)	The difference between the expected value of a statistic and the value of the parameter is called a				B
	A.	Standard error	B.	Bias	
	C.	Sampling error	D.	Non-sampling error	
357)	Interviewing all members of a given population is called				C
	A.	A sample	B.	Sampling	
	C.	Census	D.	Sample survey	
358)	A sampling frame is:				C
	A.	A summary of the various stages involved in designing a survey	B.	An outline view of all the main clusters of units in a sample	
	C.	A list of all the units in the population from which a sample will be selected	D.	A wooden frame used to display tables of random numbers	
359)	It is helpful to use a multi-stage cluster sample when				D
	A.	The population is widely dispersed geographically	B.	You have limited time and money available for travelling	
	C.	You want to use a probability sample in order to generalize the results	D.	All of the above	
360)	Non-Sampling error is reduced by				D
	A.	Increasing Sample Size	B.	Decreasing Sample Size	
	C.	Reducing the Amount of Data	D.	None of these	
361)	Simple random sample can be drawn with the help of				D
	A.	random number table	B.	chit method	
	C.	roulette wheel	D.	all the above	
362)	As a normal practice, sampling fraction is considered to be negligible if it is:				B
	A.	less than 10%	B.	less than or equal to 5%	
	C.	more than 5%	D.	more than 10%	
363)	Under equal allocation in stratified sampling the sample from each stratum is:				B
	A.	proportional to stratum size	B.	of same size from each stratum	
	C.	is proportional to per unit cost of survey	D.	all the above	

364)	Selected units of a systematic sample are:				B
	A.	not easily locatable	B.	easily locatable	
	C.	not representing the whole population	D.	all the above	
365)	Which of the following statements is true?				D
	A.	population mean increases with the increase in sample size	B.	population mean decreases with increase in sample size	
	C.	population mean decreases with the decrease in sample size	D.	population mean is a constant value	
366)	If $E(\bar{X}) = 10$ and $\mu = 10$ then bias is equal to:				B
	A.	0	B.	10	
	C.	20	D.	Difficult to tell	
367)	The standard error increases when sample size is				A
	A.	Increased	B.	Decreased	
	C.	Fixed	D.	More than 30	
368)	Which of the following statement is true?				B
	A.	more the S.E, better it is	B.	less the S.E, better it is	
	C.	S.E is always zero	D.	S.E is always unity	
369)	Which of the following statement is not true?				D
	A.	S.E cannot be zero	B.	S.E cannot be one	
	C.	S.E can be negative	D.	all the above	
370)	When N=5, n=2 then the number of all possible samples that can be drawn by without replacement are				C
	A.	5	B.	7	
	C.	10	D.	25	
371)	A population of 5 members is 2, 9, 6, 5, 3. If we construct a sampling distribution of means by taking samples of size 3 by without replacement what would be the Standard error of mean.				A
	A.	1	B.	2	
	C.	1.414	D.	5	
372)	A population of 5 members is 2, 9, 6, 5, 3. If we construct a sampling distribution of means by taking samples of size 3 by with replacement what would be the Standard error of mean.				C
	A.	1	B.	2	
	C.	1.414	D.	5	
373)	A selection procedure of a sample having no involvement of probability is known as				D
	A.	Purposive sampling	B.	Judgment sampling	
	C.	Subjective sampling	D.	all the above	
374)	If $E(\bar{X}) = \mu$, then bias is:				C
	A.	Positive	B.	Negative	
	C.	Zero	D.	100%	
375)	Standard deviation of sampling distribution of a statistic is called				C
	A.	Serious error	B.	Dispersion	
	C.	Standard error	D.	Difference	
376)	Any measure of the population is called:				B
	A.	Finite	B.	Parameter	
	C.	Without replacement	D.	Random	
377)	Under proportional allocation, the size of the sample from each stratum is proportional to				B
	A.	total sample size	B.	size of the stratum	

	C.	population size	D.	all the above	
378)	Systematic sampling means:				B
	A.	selection of n contiguous units	B.	selection of n units situated at equal distances	
	C.	selection of n largest units	D.	selection of n middle units in a sequence	
379)	If population variance of an infinite population is σ^2 and a sample of n items is selected from this population, the standard error of sample mean is equal to				C
	A.	$\frac{\sigma^2}{n}$	B.	σ/n	
	C.	$\sigma/\sqrt{n}$	D.	σ	
380)	Sampling error can be reduced by:				D
	A.	choosing a proper probability sampling	B.	selecting a sample of adequate size	
	C.	using a suitable formula for estimation	D.	all the above	
381)	The magnitude of the standard error of an estimate is an index of its:				B
	A.	accuracy	B.	precision	
	C.	efficiency	D.	all the above	
382)	Probability of selection varies at each subsequent draw in				A
	A.	Sampling without replacement	B.	Sampling with replacement	
	C.	both (a) and (b)	D.	neither (a) nor (b)	
383)	If **N** is the size of the population and **n** is size of the sample, then sampling fraction is:				C
	A.	n^N	B.	N^n	
	C.	n/N	D.	$\binom{N}{n}$	
384)	The standard error increases when sample size is				B
	A.	Increased	B.	Decreased	
	C.	Fixed	D.	More than 30	
385)	The sampling procedure in which the population is first divided into homogenous groups and then a sample is drawn from each group is called:				C
	A.	Probability sampling	B.	Simple random sampling	
	C.	Stratified random sampling	D.	Sampling with replacement	
386)	A population about which we want to get some information is called				D
	A.	Finite population	B.	Infinite population	
	C.	Sampling population	D.	Target population	
387)	For making voters list in Pakistan, we need:				C
	A.	Sampling error	B.	Standard error	
	C.	Census	D.	Simple random sampling	
388)	Sampling based upon equal probability is called:				C
	A.	Probability sampling	B.	Systematic sampling	
	C.	Simple random sampling	D.	Stratified random sampling	
389)	In probability sampling, the probability of selecting an item from the population is				B
	A.	Equal to zero	B.	Non zero	
	C.	Equal to one	D.	All of the above	
390)	50. A Population of 16 observations having standard deviation 6 and mean 10. The mean of the sampling distribution of mean when a sample of size 4 is taken by with replacement:				B
	A.	3	B.	10	
	C.	9	D.	4.23	

					Answer Key
391	Which one provide the estimate of experimental error in design of experiment				C
	A.	Randomization	B.	Local Control	
	C.	Replication	D.	Treatment	
392	There are _____ basic principles of sound statistical design				B
	A.	2	B.	3	
	C.	4	D.	5	
393	If the experimental material is not homogenous and there is one source of variation in experiment then we use				B
	A.	CRD	B.	RCBD	
	C.	LSD	D.	FACTORIAL EXPERIMENTS	
394	An experiment is performed in CRD with 3 replication to compare four treatments. Then total experimental units will be				D
	A.	3	B.	11	
	C.	4	D.	12	
395	An experiment is performed in CRD with 3 replication to compare four treatments. Then what will be degree of freedom for total				B
	A.	3	B.	11	
	C.	4	D.	12	
396	An experiment is performed in CRD with 3 replication to compare four treatments. Then what will be degree of freedom for error				C
	A.	9	B.	11	
	C.	8	D.	12	
397	An experiment is performed in CRD with 3 replication to compare four treatments. The treatments sum of square = 8. If Error sum of Square = 12 then what will be total sum of square				B
	A.	4	B.	20	
	C.	9	D.	12	
398	An experiment is performed in CRD with 3 replications to compare four treatments. The treatments sum of square = 9. If Error sum of Square = 12 then what will be mean sum of square for treatment				C
	A.	4	B.	9	
	C.	3	D.	11	
399	What will be the F ratio if an experiment is performed in CRD with 3 replication to compare four treatments. The treatments mean Sum of Square = 96, Error mean sum of square = 12				C
	A.	9	B.	11	
	C.	8	D.	5	
400	Analysis of variance is a statistical method of comparing the _____ of several populations				C
	A.	One mean only	B.	Two means only	
	C.	More than two means	D.	All of the above	
401	To test the hypothesis about one population variance, the test statistic will be				D
	A.	F	B.	Z	
	C.	T	D.	Chi-square	
402	Under One way variability in environmental conditions the appropriate design for conducting an experiment will be				C
	A.	LSD	B.	CRD	
	C.	RCBD	D.	None of A, B & C	
403	The scientific method for the construction of statistical layout plan for an experiment is				A
	A.	Random Number Table	B.	Lottery Method	
	C.	Scientist's preference	D.	None of A, B & C	
404	MANOVA stands for				B
	A.	Multiple Analysis of Variance	B.	Multivariate Analysis of Variance	

	C.	Many Analysis of Variance	D.	Most Applied Analysis of Variance	
405	The data is recorded with “p” variables on “n” objects in array shape known as:				C
	A.	Data matrix	B.	Sample matrix	
	C.	Multivariate matrix	D.	None of the above	
406	For testing mean vector in multivariate normal distribution we use				D
	A.	Multivariate Normal Distribution	B.	F-Distribution	
	C.	T-Distribution	D.	Hotelling’s T square Distribution	
407	The determined the appropriate number of principal components in multivariate analysis, the following method could be use				B
	A.	Box Plot	B.	Scree Plot	
	C.	Scatter plot	D.	Q-Q plot	
408	The variances and covariance in 2×2 covariance matrix will be				D
	A.	2,2	B.	1,2	
	C.	2,3	D.	2,1	
409	Which of the following technique is useful for data reduction				C
	A.	Cluster Analysis	B.	Multivariate analysis of variance	
	C.	Principal components	D.	Eigen Analysis	
410	Total sample variance is equivalent to				A
	A.	Generalized variance	B.	Transpose of Variance Covariance matrix	
	C.	Sum of all off diagonal elements	D.	Non of the Above	
411	What will be degree of freedom (df) for column-wise blocking in 5×5 Latin Square Design				A
	A.	4	B.	14	
	C.	2	D.	15	
412	In Completely Randomized Design (CRD) for 7 treatments with 4 replications of each; the total df				D
	A.	6	B.	3	
	C.	28	D.	27	
413	In Randomized Complete Block Design (RCBD) with 5 treatments & 4 blocks; the error df will be				B
	A.	13	B.	12	
	C.	11	D.	10	
414	The Error sum of squares in ANOVA table could be				C
	A.	Negative	B.	Positive but greater than treatment sum of squares	
	C.	Positive but less than treatment sum of squares	D.	Non of the Above	
415	The RCBD design control the source of variation in				B
	A.	None of the directions	B.	One directions	
	C.	Two directions	D.	Three directions	
416	The mean sum of square (MS) is the sum of squares divided by				A
	A.	Its corresponding degree of freedom	B.	Total number of observations	
	C.	Total number of Treatments	D.	None of the above	
417	In case of studying two sources of variation; Complete Randomized Design is				C
	A.	Efficient	B.	Less efficient	
	C.	Not efficient	D.	Very efficient	
418	If the number of replications equals the number of varieties, which plot design is appropriate				A
	A.	Latin square design	B.	Random design	
	C.	Systematic design	D.	CRD with unequal replication	
419	The implementation of ANOVA required -------- assumptions				B
	A.	1	B.	4	
	C.	2	D.	3	
420	How many main effects would be calculated in four factor factorial experiment?				C
	A.	3	B.	2	
	C.	4	D.	1	
421	For rejecting the null hypothesis the P-value should be				D

	A.	Greater than level of significance	B.	Greater than power of test	
	C.	Less than power of test	D.	Less than level of significance	
422	Randomized complete block design is not recommended when material is				A
	A.	Homogenous	B.	Out of control	
	C.	Heterogenous	D.	More than 5 treatment	
423	Tuckey's Honestly significant difference (HSD) is applied when				B
	A.	F ratio is non significant	B.	F ratio is significant	
	C.	F ratio is very small	D.	Non of the other given option	
424	An experiment is performed in RCBD with 3 blocks to compare five treatments. Then total experimental units will be				B
	A.	14	B.	15	
	C.	5	D.	3	
425	An experiment is performed in RCBD with 3 blocks to compare five treatments. Then error degree of freedom will be				C
	A.	14	B.	4	
	C.	8	D.	2	
426	If Relative efficiency (RE) of RCBD = 3.25 then RCBD is -------- efficient than CRD				C
	A.	Less	B.	Equal	
	C.	More	D.	Not Known	
427	An experiment is performed in which four seed treatment were compared with no treatment (check) on soybean seeds using RCBD with 5 blocks. what will be degree of freedom for treatment				B
	A.	5	B.	4	
	C.	3	D.	2	
428	The F ratio in ANOVA for a Completely Randomized Design (CRD) is the ratio of				B
	A.	MST/MST (Mean Square Treatments÷ Mean Square Total)	B.	MST/MSE (Mean Square Treatments÷ Mean Square Error)	
	C.	SST/SSE (Sum of Square Treatments÷ Sum of Square Error)	D.	MSB/MSE (Mean Square Blocks ÷ Mean Square Error)	
429	An ANOVA procedure for CRD is applied to data obtained from 6 samples, where each sample contains 9 observations. The degrees of freedom for the critical value of F are				C
	A.	5 numerator and 8 denominator degrees of freedom	B.	53 degrees of freedom	
	C.	5 numerator and 48 denominator degrees of freedom	D.	54 degrees of freedom	
430	When conducting ANOVA Fdata will always fall within the range?				A
	A.	Between o and infinity	B.	Between negative infinity and positive infinity	
	C.	Between 0 and 1	D.	Between 1 and infinity	
431	The ------- sum of squares measure the variation of sample treatment mean around the overall mean				D
	A.	Error	B.	Total	
	C.	Interaction	D.	Treatment	
432	If the true means of p population are same then MSTR/MSE should be				A
	A.	Close to 1	B.	Close to -1	
	C.	A negative value between 0and 1	D.	Greater than 1	
433	There are 30 students in class and each experimental condition in 5*2 between group design how many participants would be needed				C
	A.	30	B.	10	
	C.	300	D.	40	
434	What we must include when reporting ANOVA				D
	A.	S.D	B.	Mean	
	C.	d.f	D.	All of these	
435	In a factorial experiment				B

	A.	Testing one factor at a time	B.	All possible combination of factor levels are tested	
	C.	Cannot estimate interaction	D.	All of these	
436	Regression + ANOVA is called				B
	A.	ANOVA	B.	ANCOVA	
	C.	MANOVA	D.	Regression	
437	ANCOVA stands for				B
	A.	Analysis of variance	B.	Analysis of covariance	
	C.	Multivariate analysis of variance	D.	Multivariate analysis of covariance	
438	Approximately what percentage of scores fall within one standard deviation of the mean in a normal distribution				D
	A.	34%	B.	99%	
	C.	95%	D.	68%	
439	What is the Error d.f formula in Latin square design if there are p treatments				C
	A.	p-1	B.	P^2-1	
	C.	(p-1)(p-2)	D.	p-2	
440	If there are four treatments in a Latin square design then the error d.f will be				B
	A.	3	B.	6	
	C.	12	D.	15	
441	Which one control the source of variation locally				C
	A.	Randomization	B.	Replication	
	C.	Blocking	D.	Treatment	
442	The first principle of statistical design is				A
	A.	Randomization	B.	Replication	
	C.	Blocking	D.	Treatment	
443	The third principal of statistical design is				C
	A.	Randomization	B.	Replication	
	C.	Local control	D.	Treatment	
444	An experiment is performed in CRD with 3 replication to compare four treatments. The treatments total are 23, 26,30 and 20. Then what will be the value of Correction factor (CF)				B
	A.	99	B.	816.75	
	C.	3267	C.	33	
445	An experiment is performed in CRD with 3 replication to compare four treatments. The treatment total are 23,26, 30 and 20. Then what will be Error degree of freedom				D
	A.	11	B.	3	
	C.	4	D.	8	
446	An experiment is performed in CRD with 3 replication to compare four treatments. The treatments SS = 80, total SS = 120, then F ratio will be				D
	A.	23.67	B.	26.66	
	C.	5.00	D.	5.33	
447	The variable of interest in an ANOVA procedure is called				C
	A.	A Block	B.	A correction factor	
	C.	A Treatment	D.	An Error	
448	The ANOVA procedure is a statistical approach for determining whether or not the				D
	A.	Means of two samples are equal	B.	Means of two populations are equal	
	C.	Means of at least three samples are equal	D.	Means of at least three populations are equal	
449	The number of times each experimental condition is observed in an experimental design is known as				A
	A.	A replication	B.	An experimental condition	
	C.	A factor	D.	A block	
450	The single replication is called				A
	A.	Replicate	B.	Block	

	C.	Treatment	D.	None of these	
451	Which design yield maximum degree of freedom for error				A
	A.	CRD	B.	Latin square	
	C.	RCBD	D.	None of these	
452	Which design is not suitable for field				A
	A.	CRD	B.	Latin square	
	C.	RCBD	D.	None of these	
453	Error sum of square never be				A
	A.	Negative	B.	Positive	
	C.	Negative and Positive	D.	None of these	
454	If Trt=4 and TrtMS=35 then TrtSS will be				C
	A.	101	B.	103	
	C.	105	D.	107	
455	If TrSS=116.25 with df=3, ErrorSS=28.5 with df=6, TotalSS=164.25 with df=11. The Block F ratio will be				C
	A.	0.053	B.	1.053	
	C.	2.053	D.	3.053	
456	The sum of residual is always				A
	A.	0	B.	0.5	
	C.	1	D.	2	
457	Bivariate analysis has				B
	A.	One variable	B.	Two variable	
	C.	Three variable	D.	Multiple variables	
458	In multivariate analysis there is				C
	A.	Single response variable	B.	Two response variables	
	C.	More than one response variable	D.	None of these	
459	Bivariate statistics is ____in nature				B
	A.	Descriptive	B.	Inferential	
	C.	Descriptive and inferential	D.	None of these	
460	Univariate is a				A
	A.	Descriptive statistics	B.	Inferential statistics	
	C.	Descriptive and inferential	D.	None of these	
461	MANOVA is an extension of ANOVA when there are ----------				D
	A.	One dependent variable	B.	Two dependent variables	
	C.	More than one independent variable	D.	More than one dependent variable	
462	Which Multivariate techniques reduce the number of variables				C
	A.	Factor analysis	B.	Principal component analysis	
	C.	Both factor analysis and principal component analysis	D.	Neither factor analysis nor principal component analysis	
463	Principal component analysis is used when				D
	A.	The number of variables is large	B.	The variables are highly correlated	
	C.	The sample size is large	D.	all of these	
464	Wilks lambda statistic is based on the principal of				B
	A.	Least square	B.	Likelihood ratio	
	C.	Variance ratio	D.	Lagrange multiplier	
465	An experiment design is				C
	A.	A map	B.	An architect	
	C.	A plan of experiment	D.	All of these	
466	Which one of the following is not a contrast				A
	A.	T1+2T2-T3	B.	T1-T3	
	C.	T1-2T2+T3	D.	-T1+2T2-T3	
467	Which one of the following is contrast				C

	A.	3T1+T2-3T3+T4	B.	T1+3T2-3T3+T4	
	C.	-3T1-T2+T3+3T4	D.	T1+T2+T3-T4	
468	An experiment is performed in CRD with 10 replication to compare two treatments. Then total experimental units will be				D
	A.	10	B.	12	
	C.	8	D.	20	
469	A teacher use the different teaching ways on different groups in his class to look which yields the best results. In this example a treatment is				C
	A.	class	B.	teacher	
	C.	Different teaching ways	D.	Different groups	
470	If the total degrees of freedom and between treatments in a CRD are 15 and 4 respectively, the degrees of freedom for error will be				A
	A.	11	B.	18	
	C.	14	D.	19	
471	If there are 6 treatments with 3 blocks in a RCBD then the degrees of freedom for error are				C
	A.	3	B.	6	
	C.	10	D.	15	
472	For a 7x7 Latin Square design there will be observations				D
	A.	0	B.	7	
	C.	14	D.	49	
473	In a Latin Square design, the SSE can be obtained as				B
	A.	SSE=SST+SSTr+SSR+SSC	B.	SSE=SST-SSTr-SSR-SSC	
	C.	SSE=SST-SSTr+SSR+SSC	D.	SSE=SST+SSTr-SSR-SSC	
474	The natural variability of process is measured by				B
	A.	Process mean	B.	Process standard deviation	
	C.	Sample standard deviation	D.	Sample maean	
475	Which type of the chart will be used to plot the number of defective in the output of any process				D
	A.	X bar chart	B.	R chart	
	C.	C chart	D.	P chart	
476	Process control is carried out				B
	A.	Before production	B.	During production	
	C.	After production	D.	All of these	
477	Central tendency of a process is observed through				B
	A.	Range chart	B.	Mean chart	
	C.	U chart	D.	C hart	
478	The process capability is calculated as				C
	A.	(USL-LSL)/3sigma	B.	(USL+LSL)/3sigma	
	C.	(USL-LSL)/6sigma	D.	(USL+LSL)/6sigma	
479		Process capability uses			B
	A.	Specification	B.	Control limits	
	C.	Process standard deviation	D.	Mean of any one sample	
480		In one way ANOVA if total number of observation is 15 with 5 treatments then total d.f is			D
	A.	75	B.	10	
	C.	3	D.	14	
481		If there are four treatments in a Latin square design then total d.f will be			D
	A.	3	B.	6	
	C.	12	D.	15	
482		In a RCBD if there are 8 treatments and 4 blocks then df of total will be			D
	A.	7	B.	21	
	C.	3	D.	31	

483		In RCBD if there are 8 treatments and 4 blocks having TrtSS is 3243.5 and BlockSS is 9.8547 then what will be the ErrorSS			C
	A.	60.02	B.	60.45	
	C.	63.176	D.	90.34	
484		In RCBD if there are 8 treatments and 4 blocks having trtSS is 3243.5 and blockSS is 9.8547 then what will be the MS of Treatment			C
	A.	1.1	B.	1.7	
	C.	3.3	D.	4.5	
485		What will be the F value in case of RCBD experiment if Mean square for treatment is 36 and MSError is 3			B
	A.	11	B.	12	
	C.	2	D.	15	
486		If there are 8 treatments and 4 blocks contain one missing observation in RCBD, then the error degrees of freedom is			C
	A.	22	B.	26	
	C.	20	D.	28	
487		In RCBD all restrictions are imposed only on			A
	A.	Complete block	B.	Random block	
	C.	Average	D.	calculations	
488		Which design of experiment is suitable for laboratory experiment			A
	A.	CRD	B.	RCBD	
	C.	LSD	D.	Factorial experiment	
489		A design in which the treatments are assigned to the experimental unit completely at random			B
	A.	ANCOVA	B.	CRD	
	C.	RCBD	D.	Factorial experiment	
490		The highest value that a quality characteristic can take before the process becomes out-of-control is called			B
	A.	Central line	B.	Upper control limit	
	C.	Lower control limit	D.	Control limit	

					Answer Key
491)	The feed of a certain type of hormone increases the mean weight of chicks by 0.3 ounces. A sample of 25 eggs has a mean increase of 0.4 ounces with its standard deviation as 0.20 ounces. What is the value of t-statistic?				A
	A.	2.5	B.	-10	
	C.	-2.5	D.	10	
492)	Scientists claim that a diet will increase the mean weight of eggs at least by 0.3 ounces. A sample of 25 eggs has a mean increase of 0.4 ounces with a S.D. of 0.20. What will be the null hypothesis for testing this claim about diet?				A
	A.	$\mu \geq 0.3$	B.	$\mu \leq 0.3$	
	C.	$\mu > 0.3$	D.	$\mu > 0.4$	
493)	Scientists claim that a diet will increase the mean weight of eggs by 0.3 ounces. A sample of 25 eggs has a mean increase of 0.4 ounces with a S.D. of 0.20. What will be the null hypothesis for testing this claim about diet?				C
	A.	$\mu \geq 0.3$	B.	$\mu \leq 0.3$	
	C.	$\mu > 0.3$	D.	$\mu > 0.4$	
494)	The type-I error occurs when				A
	A.	Rejecting the TRUE null hypothesis	B.	Rejecting the FALSE null hypothesis	
	C.	Don't rejecting the FALSE null hypothesis	D.	Don't rejecting the TRUE null hypothesis	
495)	The type-II error occurs when				C
	A.	Rejecting the TRUE null hypothesis	B.	Rejecting the FALSE null hypothesis	
	C.	Don't rejecting the FALSE null hypothesis	D.	Don't rejecting the TRUE null hypothesis	
496)	The POWER of test occurs when				B
	A.	Rejecting the TRUE null hypothesis	B.	Rejecting the FALSE null hypothesis	
	C.	Don't rejecting the FALSE null hypothesis	D.	Don't rejecting the TRUE null hypothesis	
497)	Null and Alternate hypothesis are the statements about				A
	A.	Population parameters	B.	Sample statistics	
	C.	Sampling Distribution with replacement	D.	Sampling Distribution without replacement	
498)	To test the Average Marks of the whole class in a statistics course, if the researcher has no knowledge about the population variance and s/he selects a sample size less than 30; then s/he must use				D
	A.	Chi-square test	B.	F-test	
	C.	Z-test	D.	t-test	
499)	To test the Average Marks of the whole class in a statistics course, if the researcher has the knowledge about the population variance and s/he selects a sample size less than 30; then s/he must use				C
	A.	Chi-square test	B.	F-test	
	C.	Z-test	D.	t-test	
500)	To test the Average Marks of the whole class in a statistics course, if the researcher has no knowledge about the population variance and s/he selects a sample size more than 30; then s/he must use				C
	A.	Chi-square test	B.	F-test	
	C.	Z-test	D.	t-test	
501)	To test the Average Marks of the whole class in a statistics course, if the researcher has the knowledge about the population variance and s/he selects a sample size more than 30; then s/he must use				C
	A.	Chi-square test	B.	F-test	
	C.	Z-test	D.	t-test	
502)	A sample is used to				C
	A.	Increase time	B.	Reduced Efficiency	
	C.	Reduced Cost	D.	Least Accuracy	
503)	When applying a certain rule or formula to the collected data from a random experiment; then the obtained quantity is known as				C
	A.	Statistics	B.	Estimator	
	C.	Estimate	D.	Parameter	
504)	A manufacturer claims that the average life of his light bulbs is more than 3000 hrs. A random sample of 36 bulbs is tested and found to have an average lifetime of 2985 and standard deviation 15. Choose the value of the test statistic.				B
	A.	9	B.	6	
	C.	8	D.	7	
505)	Inference about population parameters can be obtained through				D
	A.	Estimation	B.	Testing of Hypothesis	

	C.	Interval Estimation	D.	Estimation and Testing of Hypothesis	
506)	A manufacturer claimed that the average life of its product is at least 50 days. Choose the appropriate alternative hypothesis that population mean will be				B
	A.	More than 50	B.	Less than 50	
	C.	Equal to 50	D.	At most 50	
507)	A manufacturer claimed that the average life of its product will not be more than 50 days. Choose the appropriate alternative hypothesis that population mean will be				A
	A.	More than 50	B.	Less than 50	
	C.	Equal to 50	D.	At most 50	
508)	A manufacturer claimed that the average life of its product will be more than 50 days. Choose the appropriate alternative hypothesis that population mean will be				A
	A.	More than 50	B.	Less than 50	
	C.	Equal to 50	D.	At most 50	
509)	A manufacturer claimed that the average life of its product is fewer than 50 days. Choose the appropriate alternative hypothesis that population mean will be				B
	A.	More than 50	B.	Less than 50	
	C.	Equal to 50	D.	At most 50	
510)	The numerical values computed from sample data randomly drawn from a population are called				A
	A.	Parameters	B.	Sampling Unit	
	C.	Sampling Frame	D.	Statistics	
511)	The numerical values computed from sample data randomly drawn from a sample are called				D
	A.	Parameters	B.	Sampling Unit	
	C.	Sampling Frame	D.	Statistics	
512)	The test used for testing the significance in an analysis of variance table is				B
	A.	Chi-square test	B.	F-test	
	C.	Z-test	D.	t-test	
513)	For testing the equality of two means using t-test; there is an assumption that population means are equal. This assumption about equality of variances will be tested using the				B
	A.	Chi-square test	B.	F-test	
	C.	Z-test	D.	t-test	
514)	For computing the confidence interval about a single population variance; the following test will be used				A
	A.	Chi-square test	B.	F-test	
	C.	Z-test	D.	t-test	
515)	A researcher is interested to test a certain value of variability among the plants' heights; the following test will be used				A
	A.	Chi-square test	B.	F-test	
	C.	Z-test	D.	t-test	
516)	The point where the Null Hypothesis being rejected is called as				D
	A.	Significant value	B.	Rejection Value	
	C.	Acceptance Value	D.	Critical Value	
517)	For valid statistical inference there is importance of sampling. select the order of sampling from best to worst				A
	A.	simple random, stratified, convenience	B.	simple random, convenience, stratified	
	C.	stratified, simple random, convenience	D.	stratified, convenience, simple random	
518)	What will be the value of test-statistic when testing the hypothesis that average temperature of a location is 2^0C. The recorded temperatures (^{0}C) of randomly selected days are -4,0,12,0				A
	A.	Zero	B.	$-\infty$	
	C.	$+\infty$	D.	Not possible to find a value	
519)	An estimator "q" is an unbiased estimator of the population mean "Q" if				B
	A.	E(x)=μ	B.	E(q)=Q	
	C.	E(Q)=Q	D.	Each value of sample points be the same	
520)	A test is said to be most powerful test of size α, if				
	A.	Among all other test of size α or less it has the largest power	B.	Among all other test of size α or greater it has the largest 1- α	
	C.	Among all other test of size α or greater it has the smallest power	D.	Among all other test of size α or greater it has the largest β	
521)	A set of jointly sufficient statistics is defined to be minimal sufficient if and only if				B
	A.	It is a function of some other set of sufficient statistics	B.	It is a function of every other set of sufficient statistics	

	C.	It is a function of any sufficient statistics in the set	D.	It is not a function of every other set of sufficient statistics	
522)	For a biased estimator $\hat{\theta}$ of θ, which one of the following is correct				B
	A.	$MSE(\hat{\theta})=SD(\hat{\theta})+[Bias(\hat{\theta})]$	B.	$MSE(\hat{\theta})=Var(\hat{\theta})+[Bias(\hat{\theta})]^2$	
	C.	$MSE(\hat{\theta})=SD(\hat{\theta})+[Bias(\hat{\theta})]^2$	D.	$MSE(\hat{\theta})=Var(\hat{\theta})+[Bias(\hat{\theta})]$	
523)	The critical value of a test statistic is determined from the following				C
	A.	The sampling distribution of the statistics assumption the Null Hypothesis	B.	Calculations based on many actual reptations of the same Experiment	
	C.	The sampling distribution of the statistic assuming Alternative hypothesis	D.	None of these	
524)	What is the probability of a type II error when α=0.07				D
	A.	0.049	B.	0.93	
	C.	0.03	D.	Cannot be determined without more information	
525)	The appropriate statistical method for comparing the equality of more than TWO population means is				D
	A.	Z-test	B.	t-test	
	C.	F-test	D.	ANOVA	
526)	What will be the unbiased point estimator for population mean if the recorded temperatures (^{0}C) of randomly selected days were the -4,0,12,0.				B
	A.	Zero	B.	2	
	C.	4	D.	6	
527)	What is the most appropriate to say about the statistical inference for population mean if the recorded temperatures (^{0}C) of randomly selected days were the -4,0,12,0.				C
	A.	Possible to find statistical inference	B.	Not possible to find statistical inference	
	C.	Limited information is given	D.	Data are not related to what is required	
528)	What will be the unbiased point estimator for population variance if the recorded temperatures (^{0}C) of randomly selected days were the -4,0,12,0.				C
	A.	Zero	B.	2	
	C.	48	D.	6	
529)	A parameter is a ------ quantity whereas statistic is a ----- quantity.				B
	A.	Constant, Constant	B.	Constant, Variable	
	C.	Variable, Constant	D.	Variable, Variable	
530)	1- α is the probability of ---------				A
	A.	Acceptance Region	B.	Type-I error	
	C.	Rejection Region	D.	Type-II error	
531)	Power of a test is denoted by------				C
	A.	β	B.	α	
	C.	1-β	D.	1-α	
532)	Confidence interval is denoted by------				D
	A.	β	B.	α	
	C.	1-β	D.	1-α	
533)	A value(s) that separates the rejection region and the acceptance region(s) is called as				A
	A.	Critical Value	B.	Confidence Interval	
	C.	Width of the Confidence Interval	D.	Confidence Boundaries	
534)	The significance level of a test is denoted by------				B
	A.	β	B.	α	
	C.	1-β	D.	1-α	
535)	Which of the following is simple hypothesis				B
	A.	$\mu \geq 30$	B.	$\mu = 30$	
	C.	$\mu \neq 30$	D.	$\mu \leq 30$	
536)	The one who only has one POSSIBLE value in any circumstances				C
	A.	Confidence Interval	B.	Point Estimator	
	C.	Point Estimate	D.	Parameter	
537)	When variance of an estimator approaches to zero as samples gets larger then such an estimator will be called as				D
	A.	Sufficient	B.	Efficient	
	C.	Unbiased	D.	Consistent	
538)	The process of making estimates about the population parameter from a sample is called				A
	A.	Statistical inference	B.	Statistical independence	

	C.	Statistical decision	D.	Statistical hypothesis	
539)	There are two main branches of statistical inference, namely				D
	A.	Biased estimator and unbiased estimator	B.	Level of significance and degree of freedom	
	C.	Point estimate and interval estimate	D.	Estimation of parameter and testing of hypothesis	
540)	The process of using sample data to estimate the values of unknown population parameters is called				D
	A.	Estimate	B.	Interval Estimate	
	C.	Estimator	D.	Estimation	
541)	'Statistic' is an estimator, and its computed value(s) is called				B
	A.	Interval estimate	B.	Estimate	
	C.	Estimator	D.	Estimation	
542)	The end points of a confidence interval are called				B
	A.	Confidence coefficient	B.	Confidence limits	
	C.	Parameters	D.	Width of the confidence interval	
543)	The difference between the two end points of a confidence interval is called				D
	A.	Confidence coefficient	B.	Confidence limits	
	C.	Parameters	D.	Width of the confidence interval	
544)	A set (range) of the values calculated from the sample data and it is likely to contain the true value of the parameter with some probability is called				D
	A.	Level of Confidence	B.	Confidence limits	
	C.	Point Estimate	D.	Interval Estimate	
545)	The estimator is said to be ----- if the mean of the estimator is not equal to the mean of the population parameter.				C
	A.	Positively Biased	B.	Negatively Biased	
	C.	Biased	D.	Unbiased	
546)	Estimation can be classified into				D
	A.	Biased and Unbiased	B.	One sided and sided testing	
	C.	Type-I and Type-II	D.	Point estimation and interval estimation	
547)	The estimate is the observed value of an				C
	A.	Interval estimation	B.	Estimation	
	C.	Estimator	D.	Unbiased estimator	
548)	A single value used to estimate the value of population parameter is called				B
	A.	Level of significance	B.	Point estimate	
	C.	Confidence limits	D.	Interval estimate	
549)	The probability associated with confidence interval is called				A
	A.	Confidence coefficient	B.	Confidence limits	
	C.	Degrees of freedom	D.	Width of the Confidence Interval	
550)	Each of the following increases the width of a confidence interval except				C
	A.	Increased confidence level	B.	Increased variability	
	C.	Increased sample size	D.	Decreased sample size	
551)	If a researcher takes a large enough sample, then he/she will almost always obtain				D
	A.	virtually significant results	B.	practically significant results	
	C.	consequentially significant results	D.	statistically significant results	
552)	Which of the following is true of the null and alternative hypotheses				A
	A.	Exactly one hypothesis must be true	B.	both hypotheses must be true	
	C.	It is possible for both hypotheses to be true	D.	It is possible for neither hypothesis to be true	
553)	A type II error occurs when				A
	A.	the null hypothesis is incorrectly accepted when it is false	B.	the null hypothesis is incorrectly rejected when it is true	
	C.	the sample mean differs from the population mean	D.	the test is biased	
554)	The hypothesis that an analyst is trying to prove is called the				B
	A.	elective hypothesis	B.	alternative hypothesis	
	C.	optional hypothesis	D.	null hypothesis	
555)	Which of the following is true about chi-square distribution				D
	A.	It is skewed distribution	B.	Its shape depends on number of degrees of freedom (df)	
	C.	As the degrees of freedom increases its shape becomes more symmetrical	D.	It is skewed, its shape depends on df and it becomes symmetrical as df increases	

556)	To test independence between two attributes in contingency table sum of observed frequencies must be ------ expected frequency				B
	A.	Greater than	B.	Equal to	
	C.	Less than	D.	Less than or equal to	
557)	To test independence between two attributes in contingency table, Test is always				A
	A.	Right tail	B.	Left tail	
	C.	Two tailed	D.	May be Right tail or Left tail	
558)	Width of confidence interval for population mean becomes large if				D
	A.	Sample size becomes large	B.	Standard deviation becomes small	
	C.	Level of significance becomes large	D.	Level of confidence becomes large	
559)	To test the equality of several population means, the appropriate test statistics is -----				A
	A.	F-test	B.	Chi-Square test	
	C.	t-test	D.	Z-test	
560)	To test the equality of several population variances, the appropriate test statistics is -----				B
	A.	F-test	B.	Chi-Square test	
	C.	t-test	D.	Z-test	
561)	To test the equality of two normally distributed population means and two population variances; the appropriate test statistics are ----- and ------ respectively				C
	A.	t-test and Z-test	B.	t-test and chi-square test	
	C.	t-test and F-test	D.	F-test and t-test	
562)	To test equality of two population proportion for large sample sizes appropriate test statistic is				C
	A.	F-test	B.	Chi-Square test	
	C.	Z-test and Chi-Square test	D.	Z-test	
563)	If random variable Y is distributed as normal with mean 0 and variance equal to 1 then Y^2 will be distributed as				B
	A.	Standard Normal	B.	Chi-square	
	C.	Normal	D.	t	
564)	If X and Y are two independently distributed standard normal variables, then $X^2+ Y^2$ will be distributed as ----------				B
	A.	Standard Normal	B.	Chi-Square	
	C.	F	D.	Normal	
565)	If X and Y are two independently distributed standard normal variables, then X^2/Y^2 will be distributed as ----------				C
	A.	Standard Normal	B.	Chi-Square	
	C.	F	D.	Normal	
566)	Which of the following is a good definition of standard error				D
	A.	The variability of scores	B.	The typical amount by which sample variances deviate from the population variance	
	C.	The estimated standard deviation of scores	D.	The typical amount by which sample means deviate from the population mean	
567)	The sample standard deviation of a sample of 9 scores is 8.3. What is the best estimate of the standard deviation of the population from which the sample was taken				B
	A.	0.27	B.	2.77	
	C.	27.7	D.	27	
568)	The paired t-test is really				B
	A.	Two one sample tests	B.	A one-sample test based on the difference scores	
	C.	A two samples test ignoring of the samples	D.	None of the A, B and C options	
569)	One circumstance in which researcher should not use the t-test is				A
	A.	If the scores from both groups are very skewed	B.	If the scores from both groups are normally distributed	
	C.	If the data comes from questionnaire	D.	If researcher wants to generalize from a sample	
570)	The simplest form of inferential statistics, which uses known sample evidence (statistic) to draw conclusions regarding unknown population characteristics (parameter) is known as				D
	A.	Descriptive Statistics	B.	Inferential Statistics	
	C.	Testing of Hypothesis	D.	Estimation	
571)	If an estimator achieves improved reliability and precession as the sample size becomes larger then such an estimator is called				A
	A.	Consistent	B.	Efficient	
	C.	Sufficient	D.	Unbiased	
572)	Of all possible unbiased estimators of some parameter the one with the smallest variance is said to be				B

	A.	Consistent	B.	Efficient	
	C.	Sufficient	D.	Unbiased	
573)	A single numerical quantity used to estimate the population parameter is called				C
	A.	One estimate	B.	Single Estimate	
	C.	Point estimate	D.	Random Estimate	
574)	The value obtained by subtracting the number of parameters to be estimated from the number of independent values in a sample is called				D
	A.	Type-I Error	B.	Type-II Error	
	C.	Level of Significance	D.	Degrees of freedom	
575)	The hypothesis against which we hope to gather evidence is called				A
	A.	Null Hypothesis	B.	Alternative Hypothesis	
	C.	Statistical Hypothesis	D.	Composite Hypothesis	
576)	An estimator based on maximum available information in a sample is called				C
	A.	Consistent	B.	Efficient	
	C.	Sufficient	D.	Unbiased	
577)	To prove that one teaching method is superior to another, the null hypothesis "there is no difference in the two methods" would (Choose the most appropriate and more comprehensive)				A
	A.	be corrected one	B.	be incorrected one	
	C.	not be tested because of limited given information	D.	be irrelevant	
578)	The sum of squares of a sequence of independent normal variates with mean μ and variance σ^2 is said to be				D
	A.	Standard normal variate	B.	t variate	
	C.	Normal Variate	D.	Chi-square variate	
579)	When we want to test the hypothesis concerning population variance the distribution for statistical inference will be used is				B
	A.	Normal	B.	Chi-Square	
	C.	t	D.	F	
580)	A technique by means of which we test the hypothesis whether the sample distribution is in agreement with the theoretical distribution is called				B
	A.	Normal	B.	Chi-Square	
	C.	t	D.	F	
581)	Analysis of variance is a statistical method of testing hypothesis for comparing the ---------- of several populations				A
	A.	Means	B.	Proportions	
	C.	Standard Deviations	D.	Probabilities	
582)	The p-value in hypothesis testing represents which of the following. (Select the best answer among given choices)				D
	A.	The probability of failing to reject the null hypothesis, given the observed results	B.	The probability that the null hypothesis is true, given the observed results	
	C.	The probability that the observed results are statistically significant, given that the null hypothesis is true	D.	The probability of observing results as extreme or more extreme than currently observed, given that the null hypothesis is true	
583)	What is one of the distinctions between a population parameter and a sample statistic?				B
	A.	A population parameter is only based on conceptual measurements, but a sample statistic is based on a combination of real and conceptual measurements.	B.	A sample statistic changes each time you try to measure it, but a population parameter remains fixed.	
	C.	A population parameter changes each time you try to measure it, but a sample statistic remains fixed across samples	D.	The true value of a sample statistic can never be known but the true value of a population parameter can be known	
584)	A chi-square test involves a set of counts called "expected counts." What are the expected counts?				B
	A.	Hypothetical counts that would occur of the alternative hypothesis were true.	B.	Hypothetical counts that would occur if the null hypothesis were true.	
	C.	The actual counts that did occur in the observed data.	D.	The long-run counts that would be expected if the observed counts are representative.	
585)	The upper and lower boundaries of interval of confidence are classified as				D
	A.	Error Biased Limits	B.	Marginal Limits	
	C.	Estimate Limits	D.	Confidence Limits	
586)	For a parameter whose value is unknown, the belief or claim for that parameter is classified as				C
	A.	Parameter claim Testing	B.	Expected Belief Testing	
	C.	Hypothesis Testing	D.	Primary Limit Testing	

587)	To develop interval, estimate of any parameter of population, the value which is added or subtracted from the point estimate is classified as				D
	A.	Margin of Efficiency	B.	Margin of Consistency	
	C.	Margin of Biasedness	D.	Margin of Error	
588)	Considering the sample size for doing statistical inference, the sampling distribution of standard error decreases when the				A
	A.	Size of sample increases	B.	Size of sample decreases	
	C.	Margin of error increases	D.	Margin of error decreases	
589)	For testing the ratio of defective items out of randomly selected items from a shipment; the appropriate parameter will be				B
	A.	Mean	B.	Proportion	
	C.	Variance	D.	Mode	
590)	For a sample to be truly representative of the population, it must be				C
	A.	Fixed	B.	Specific	
	C.	Random	D.	Casual	
591)	Statistical Inference can be made using ---- data set(s)				D
	A.	Only Primary	B.	Only Secondary	
	C.	Both Primary and Secondary	D.	Either of Primary or Secondary	
592)	The mean of 85 sample points is 510 and standard error of mean is 4.99. The 95% confidence limits will be				A
	A.	500, 520	B.	400, 520	
	C.	200, 700	D.	470,570	
593)	A research firm conducted a survey to determine the mean amount smokers spend on cigarettes during a day. A sample of 100 smokers revealed that the sample mean is $5 and sample standard deviation is $2. Assume that the sample was drawn from a normal population. The point estimate of the population mean is				C
	A.	2	B.	4	
	C.	5	D.	100	
594)	A research firm conducted a survey to determine the mean amount smokers spend on cigarettes during a day. A sample of 100 smokers revealed that the sample mean is $5 and sample standard deviation is $2. Assume that the sample was drawn from a normal population. The point estimate of the population variance is				B
	A.	2	B.	4	
	C.	5	D.	100	
595)	A research firm conducted a survey to determine the mean amount smokers spend on cigarettes during a day. A sample of 100 smokers revealed that the sample mean is $5 and sample standard deviation is $2. Assume that the sample was drawn from a normal population. The point estimate of the population variance is				A
	A.	2	B.	4	
	C.	5	D.	100	
596)	Which of the following is not one of the assumptions made in the analysis of variance?				C
	A.	Each sample is an independent random sample	B.	The distribution of the response variable is a normal curve within each population	
	C.	The different populations all have the same mean	D.	The different populations all have the same standard deviation σ	
597)	Which one of the following choices describes a problem for which an analysis of variance would be appropriate?				C
	A.	Comparing the proportion of successes for three different treatments of anxiety. Each treatment is tried on 100 patients	B.	Analyzing the relationship between high school GPA and college GPA	
	C.	Comparing the mean birth weights of newborn babies for three different racial groups	D.	Analyzing the relationship between gender and opinion about capital punishment (favor or oppose)	
598)	Ninety people with high cholesterol are randomly divided into three groups of thirty, and a different treatment program for decreasing cholesterol is assigned to each group. The response variable is the change in cholesterol level after two months of treatment. An analysis of variance will be used to compare the three treatments. What null hypothesis is tested by this F-test?				D
	A.	The sample variances are equal for the three treatment groups	B.	The population variances are equal for the three treatments	
	C.	The sample means are equal for the three treatment groups	D.	The population means are equal for the three treatments	
599)	A shopper wanted to test whether there was a difference in the average waiting times at the check-out counter among 5 different supermarkets. She selected a random sample of 20 shoppers from each of the five supermarkets. What is the null hypothesis for this situation?				B

	A.	The average waiting time to check out is 25 minutes for all five supermarkets.	B.	The average waiting time to check out is the same for all five supermarkets.	
	C.	The average waiting time for each of the 100 shoppers is different.	D.	The average waiting time to check out is not the same for all five supermarkets.	
600)	A shopper wanted to test whether there was a difference in the average waiting times at the check-out counter among 5 different supermarkets. She selected a random sample of 20 shoppers from each of the five supermarkets. What is the alternative hypothesis for this situation?				D
	A.	The average waiting time to check out is 25 minutes for all five supermarkets.	B.	The average waiting time to check out is the same for all five supermarkets.	
	C.	The average waiting time for each of the 100 shoppers is different.	D.	The average waiting time to check out is not the same for all five supermarkets.	
601)	A student wanted to test whether there was a difference in the mean daily hours of study for students living in four different dormitories. She selected a random sample of 50 students from each of the four dormitories. What is the null hypothesis for this situation?				B
	A.	The mean daily hours of study is 3 hours for each dormitory	B.	The mean daily hours of study is the same for each dormitory	
	C.	The mean daily hours of study is different for each of the 200 students in the sample.	D.	The mean daily hours of study is not the same for all four dormitories.	
602)	A student wanted to test whether there was a difference in the mean daily hours of study for students living in four different dormitories. She selected a random sample of 50 students from each of the four dormitories. What is the alternative hypothesis for this situation?				D
	A.	The mean daily hours of study is 3 hours for each dormitory	B.	The mean daily hours of study is the same for each dormitory	
	C.	The mean daily hours of study is different for each of the 200 students in the sample.	D.	The mean daily hours of study is not the same for all four dormitories.	
603)	A study compared grade point averages (GPA) for students in a class: students were divided by 6 locations where they usually sat during lecture (i.e. left or right front, left or right center, left or right rear). A total sample size of 12 students was studied (2 students from each section) using one-way analysis of variance. What are the numerator and denominator degrees of freedom for the F-test?				C
	A.	6 for numerator and 12 for denominator.	B.	5 for numerator and 11 for denominator.	
	C	5 for numerator and 6 for denominator.	D	None of the given options A, B and C	
604)	A randomly selected sample of 1,000 college students was asked whether they had ever used the drug Ecstasy. Sixteen percent (16% or 0.16) of the 1,000 students surveyed said they had. Which one of the following statements about the number 0.16 is correct?				A
	A.	It is a sample proportion.	B.	It is a population proportion.	
	C.	It is a margin of error.	D.	It is a randomly chosen number.	
605)	A study compared grade point averages (GPA) for students in a class: students were divided by 6 locations where they usually sat during lecture (i.e. left or right front, left or right center, left or right rear). A total sample size of 12 students was studied (2 students from each section) using one-way analysis of variance. The p-value for the F-test is 0.46. If the significance level = 0.05, what is the conclusion?				A
	A.	The null hypothesis is not rejected so we cannot say the population means are different.	B.	The null hypothesis is not rejected so we can say the population means are different.	
	C.	The null hypothesis is rejected so we cannot say the population means are different.	D.	The null hypothesis is rejected so we can say the population means are different.	
606)	In a random sample of 1000 students, $\hat{p} = 0.80$ (or 80%) were in favor of longer hours at the school library. The standard error of $\hat{p}$ (the sample proportion) is				A
	A.	0.013	B.	0.160	
	C.	0.640	D.	0.800	
607)	For a random sample of 9 women, the average resting pulse rate is $x = 76$ beats per minute, and the sample standard deviation is $s = 5$. The standard error of the sample mean is				C
	A.	0.557	B.	0.745	
	C.	1.667	D.	2.778	
608)	Assume the cholesterol levels in a certain population have mean $\mu = 200$ and standard deviation $\sigma = 24$. The cholesterol levels for a random sample of $n = 9$ individuals are measured and the sample mean x is determined. What is the z-score for a sample mean $x = 180$?				B
	A.	-3.75	B.	-2.50	
	C.	- 0.83	D.	2.50	

609)	In a past General Social Survey, a random sample of men and women answered the question "Are you a member of any sports clubs?" Based on the sample data, 95% confidence intervals for the population proportion who would answer "yes" are .13 to .19 for women and .247 to .33 for men. Based on these results, you can reasonably conclude that				C
	A.	At least 25% of American men and American women belong to sports clubs	B.	At least 16% of American women belong to sports clubs	
	C.	There is a difference between the proportions of American men and American women who belong to sports clubs.	D.	There is no conclusive evidence of a gender difference in the proportion belonging to sports clubs	
610)	Suppose a 95% confidence interval for the proportion of Americans who exercise regularly is 0.29 to 0.37. Which one of the following statements is FALSE?				B
	A.	It is reasonable to say that more than 25% of Americans exercise regularly.	B.	It is reasonable to say that more than 40% of Americans exercise regularly.	
	C.	The hypothesis that 33% of Americans exercise regularly cannot be rejected.	D.	It is reasonable to say that fewer than 40% of Americans exercise regularly.	
611)	Null and alternative hypotheses are statements about:				A
	A.	population parameters.	B.	sample parameters.	
	C.	sample statistics.	D.	it depends - sometimes population parameters and sometimes sample statistics.	
612)	A hypothesis test is done in which the alternative hypothesis is that more than 10% of a population is left-handed. The p-value for the test is calculated to be 0.25. Which statement is correct?				D
	A.	We can conclude that more than 10% of the population is left-handed.	B.	We can conclude that more than 25% of the population is left-handed.	
	C.	We can conclude that exactly 25% of the population is left-handed	D.	We cannot conclude that more than 10% of the population is left-handed	
613)	Which of the following is NOT true about the standard error of a statistic?				D
	A.	The standard error measures, roughly, the average difference between the statistic and the population parameter.	B.	The standard error is the estimated standard deviation of the sampling distribution for the statistic.	
	C.	The standard error can never be a negative number.	D.	The standard error increases as the sample size(s) increases.	
614)	A prospective observational study on the relationship between sleep deprivation and heart disease was done by Ayas, et. al. (Arch Intern Med 2003). Women who slept at most 5 hours a night were compared to women who slept for 8 hours a night (reference group). After adjusting for potential confounding variables like smoking, a 95% confidence interval for the relative risk of heart disease was (1.10, 1.92). Based on this confidence interval, a consistent conclusion would be				A
	A.	Sleep deprivation is associated with a modestly increased risk of heart disease.	B.	Sleep deprivation is associated with a modestly decreased risk of heart disease.	
	C.	There was no evidence of an association between sleep deprivation and heart disease.	D.	Lack of sleep causes the risk of heart disease to increase by 10% to 92%.	
615)	Consider a random sample of 100 females and 100 males. Suppose 15 of the females are left-handed and 12 of the males are left-handed. What is the estimated difference between population proportions of females and males who are left-handed (females − males)? Select the choice with the correct notation and numerical value.				D
	A.	$\pi_1 - \pi_2 = 3$	B.	$\pi_1 - \pi_2 = 0.03$	
	C.	$p_1 - p_2 = 3$	D.	$p_1 - p_2 = 0.03$	
616)	A result is called "statistically significant" whenever				C
	A.	The null hypothesis is true	B.	The alternative hypothesis is true.	
	C.	The p-value is less or equal to the significance level.	D.	The p-value is larger than the significance level.	
617)	The confidence level for a confidence interval for a mean is				D
	A.	The probability the procedure provides an interval that covers the sample mean.	B.	The probability of making a Type 1 error if the interval is used to test a null hypothesis about the population mean.	
	C.	The probability that individuals in the population have values that fall into the interval	D.	the probability the procedure provides an interval that covers the population mean.	
618)	It is known that for right-handed people, the dominant (right) hand tends to be stronger. For left-handed people who live in a world designed for right-handed people, the same may not be true. To test this, muscle strength was measured on the right and left hands of a random sample of 15 left-handed men and the difference (left - right) was found. The alternative hypothesis is one-sided (left hand stronger). The resulting t-statistic was 1.80. This is an example of				B
	A.	A two-sample t-test.	B.	A paired t-test.	
	C.	A pooled t-test	D.	An unpooled t-test.	

619)	It is known that for right-handed people, the dominant (right) hand tends to be stronger. For left-handed people who live in a world designed for right-handed people, the same may not be true. To test this, muscle strength was measured on the right and left hands of a random sample of 15 left-handed men and the difference (left - right) was found. The alternative hypothesis is one-sided (left hand stronger). The resulting t-statistic was 1.80. The d.f. will be				A
	A.	14	B.	28	
	C.	18	D.	15	
620)	A test of H0: $\mu = 0$ versus Ha: $\mu > 0$ is conducted on the same population independently by two different researchers. They both use the same sample size and the same value of $\alpha = 0.05$. Which of the following will be the same for both researchers?				B
	A.	The p-value of the test.	B.	The power of the test if the true $\mu = 6$.	
	C.	The value of the test statistic.	D.	The decision about whether or not to reject the null hypothesis.	
621)	Which of the following is not a correct way to state a null hypothesis?				A
	A.	H_0: $p_1 - p_2 = 0$	B.	H_0: $\mu_1 - \mu_2 = 0$	
	C.	H_0: $\mu_d = 10$	D.	$\pi = 0.03$	
622)	A test to screen for a serious but curable disease is similar to hypothesis testing, with a null hypothesis of no disease, and an alternative hypothesis of disease. If the null hypothesis is rejected treatment will be given. Otherwise, it will not. Assuming the treatment does not have serious side effects, in this scenario it is better to increase the probability of:				A
	A.	making a Type I error, providing treatment when it is not needed.	B.	making a Type I error, not providing treatment when it is needed.	
	C.	making a Type II error, providing treatment when it is not needed.	D.	making a Type II error, not providing treatment when it is needed.	
623)	A random sample of 25 college males was obtained and each was asked to report their actual height and what they wished as their ideal height. A 95% confidence interval for μd = average difference between their ideal and actual heights was 0.8" to 2.2". Based on this interval, which one of the null hypotheses below (versus a two-sided alternative) can be rejected?				A
	A.	H_0: $\mu_d = 0.5$	B.	H_0: $\mu_d = 1.0$	
	C.	H_0: $\mu_d = 1.5$	D.	H_0: $\mu_d = 2.0$	
624)	The average time in years to get an undergraduate degree in computer science was compared for men and women. Random samples of 100 male computer science majors and 100 female computer science majors were taken. Choose the appropriate parameter(s) for this situation.				D
	A.	One population proportion p.	B.	Difference between two population proportions $p_1 - p_2$.	
	C.	One population mean μ_1	D.	Difference between two population means $\mu_1 - \mu_2$	
625)	If the word significant is used to describe a result in a news article reporting on a study				D
	A.	The p-value for the test must have been very large	B.	The effect size must have been very large.	
	C.	The sample size must have been very small.	D.	It may be significant in the statistical sense, but not in the everyday sense.	
626)	A random sample of 5000 students were asked whether they prefer a 10-week quarter system or a 15 week semester system. Of the 5000 students asked, 500 students responded. The results of this survey _____				C
	A.	can be generalized to the entire student body because the sampling was random.	B.	can be generalized to the entire student body because the margin of error was 4.5%.	
	C.	should not be generalized to the entire student body because the non-response rate was 90%.	D.	should not be generalized to the entire student body because the margin of error was 4.5%.	
627)	A significance test based on a small sample may not produce a statistically significant result even if the true value differs substantially from the null value. This type of result is known as				D
	A.	the significance level of the test	B.	the power of the study	
	C.	a Type I error	D.	a Type II error	
628)	An observational study found a statistically significant relationship between regular consumption of tomato products (yes, no) and development of prostate cancer (yes, no), with lower risk for those consuming tomato products. Which of the following is not a possible explanation for this finding?				D
	A.	Something in tomato products causes lower risk of prostate cancer	B.	There is a confounding variable that causes lower risk of prostate cancer, such as eating vegetables in general, that is also related to eating tomato products.	
	C.	A large number of food products were measured to test for a relationship, and tomato products happened to show a relationship just by chance	D.	A large sample size was used, so even if there were no relationship, one would almost certainly be detected.	

629)	An observational study found a statistically significant relationship between regular consumption of tomato products (yes, no) and development of prostate cancer (yes, no), with lower risk for those consuming tomato products. Which of the following is a valid conclusion from this finding?				B
	A.	Something in tomato products causes lower risk of prostate cancer.	B.	Based on this study, the relative risk of prostate cancer, for those who do not consume tomato products regularly compared with those who do, is greater than one	
	C.	If a new observational study were to be done using the same sample size and measuring the same variables, it would find the same relationship.	D.	Prostate cancer can be prevented by eating the right diet.	
630)	The best way to determine whether a statistically significant difference in two means is of practical importance is to				A
	A.	find a 95% confidence interval and notice the magnitude of the difference.	B.	repeat the study with the same sample size and see if the difference is statistically significant again	
	C.	see if the p-value is extremely small.	D.	see if the p-value is extremely large.	
631)	A large company examines the annual salaries for all of the men and women performing a certain job and finds that the means and standard deviations are $32,120 and $3,240, respectively, for the men and $34,093 and $3521, respectively, for the women. The best way to determine if there is a difference in mean salaries for the population of men and women performing this job in this company is				B
	A.	to compute a 95% confidence interval for the difference.	B.	to subtract the two sample means.	
	C.	to test the hypothesis that the population means are the same versus that they are different.	D.	to test the hypothesis that the population means are the same versus that the mean for men is higher.	
632)	One problem with hypothesis testing is that a real effect may not be detected. This problem is most likely to occur when				A
	A.	the effect is small and the sample size is small.	B.	the effect is large and the sample size is small.	
	C.	the effect is small and the sample size is large.	D.	the effect is large and the sample size is large.	
633)	If we do not reject the null hypothesis, we conclude that:				B
	A.	There is enough statistical evidence to infer that the alternative hypothesis is true.	B.	There is not enough statistical evidence to infer that the alternative hypothesis is true.	
	C.	There is enough statistical evidence to infer that the null hypothesis is true.	D.	There is not enough statistical evidence to infer that the null hypothesis is true.	
634)	The p-value of a test is the:				C
	A.	Smallest significance level at which the null hypothesis cannot be rejected.	B.	Largest significance level at which the null hypothesis cannot be rejected.	
	C.	Smallest significance level at which the null hypothesis can be rejected.	D.	Largest significance level at which the null hypothesis can be rejected.	
635)	To determine the p-value of a hypothesis test, which of the following is not needed?				D
	A.	Whether the test is one-tail or two-tail	B.	The value of the test statistic	
	C.	The form of the null and alternative hypotheses	D.	The level of significance	
636)	Which of the following p-values will lead us to reject the null hypothesis if the significance level of the test is 5%?				B
	A.	0.150	B.	0.034	
	C.	0.051	D.	0.550	
637)	Suppose we reject a null hypothesis at the 5% level of significance. For which of the following levels of significance do we also reject the null hypothesis?				A
	A.	6%	B.	2.5%	
	C.	4%	D.	2%	
638)	Which of the following statements about hypothesis testing is true?				A
	A.	If the p-value is greater than the significance level, we fail to reject H_0.	B.	A type II error is rejecting the null hypothesis when it is true.	
	C.	If the alternative hypothesis is that the population mean is greater than a specified value, the test is a two-tailed test.	D.	The significance level equals one minus the probability of a type-I error.	
639)	The purpose of hypothesis testing is to:				B
	A.	test how far the mean of a sample is from zero	B.	determine the appropriate value of the significance level	
	C.	determine the appropriate value of the significance level	D.	Derive the standard error of the data	
640)	To test a hypothesis involving proportions, both np and n(1-p) should				B
	A.	Be at least 30	B.	Be greater than 5	
	C.	Lie in the range from 0 to 1	D.	Be greater than 50	

					Answer Key
641)	In a Binomial Distribution, if 'n' is the number of trials and 'p' is the probability of success, then the mean value is given by _____				A
	A.	np	B.	n	
	C.	p	D.	np(1-p)	
642)	In a Binomial Distribution, if p, q and n are probability of success, failure and number of trials respectively then variance is given by _____				B
	A.	np	B.	npq	
	C.	np^2q	D.	npq^2	
643)	Nature of the binomial random variable X is:				C
	A.	Categorical	B.	Qualitative	
	C.	Discrete	D.	Continuous	
644)	In a binomial probability distribution, the sum of probability of failure and probability of success is always:				D
	A.	Zero	B.	Less than 0.5	
	C.	Greater than 0.5	D.	One	
645)	In a binomial experiment, the successive trials are:				B
	A.	Dependent	B.	Independent	
	C.	Mutually exclusive	D.	Fixed	
646)	In a binomial experiment with three trials, the variable can take:				C
	A.	2 values	B.	3 values	
	C.	4 values	D.	5 values	
647)	The shape of the binomial probability distribution depends upon the values of its:				D
	A.	Mean	B.	Variance	
	C.	Random variable	D.	Parameters	
648)	In binomial distribution the numbers of trials are:				D
	A.	Very small	B.	Very large	
	C.	Random	D.	Fixed	
649)	In a binomial probability distribution, relation between mean and variance is:				C
	A.	Mean < Variance	B.	Mean = Variance	
	C.	Mean > Variance	D.	All A, B and C are possible	
650)	Binomial distribution becomes _____ if n = 1.				B
	A.	Hypergeometric distribution	B.	Bernoulli distribution	
	C.	Uniform distribution	D.	Normal distribution	
651)	Which of the following is not property of a binomial experiment?				C
	A.	Probability of success remains constant	B.	n is fixed	
	C.	Successive trials are dependent	D.	It has two parameters	
652)	The binomial probability distribution is symmetrical when				A
	A.	p = q	B.	p < q	
	C.	p > q	D.	np > npq	
653)	The binomial distribution is negatively skewed if:				C
	A.	p < 0.5	B.	p = 0.5	
	C.	p > 0.5	D.	p = 1	
654)	If a binomial probability distribution has parameters (n, p)= (10, 0.3), the probability of x = 11 is:				A
	A.	0	B.	1	
	C.	0.03	D.	0.3	
655)	If a binomial probability distribution if n = 6, p = 0.9, then P(X = 4.5) is:				A
	A.	Zero	B.	Less than Zero	
	C.	Greater than Zero but less than One	D.	One	
656)	If three coins are tossed, the probability of two heads is:				B
	A.	1/8	B.	3/8	
	C.	2/3	D.	0	
657)	The hypergeometric distribution has _____ parameters.				B
	A.	2	B.	3	
	C.	4	D.	None of these	

658)	The probability of a success changes from trial to trial in:				C
	A.	Binomial distribution	B.	Negative binomial distribution	
	C.	Hypergeometric distribution	D.	Poisson distribution	
659)	In hypergeometric probability distribution, the relation between mean and variance is:				A
	A.	Mean > Variance	B.	Mean < Variance	
	C.	Mean = Variance	D.	All A, B and C are possible	
660)	Which of the following is the property of hypergeometric experiment?				D
	A.	**p** remains constant from trial to trial	B.	Successive trials are independent	
	C.	Sampling is performed with replacement	D.	Sampling is performed without replacement	
661)	For larger values of 'n' with small values of 'p', Binomial Distribution _____				D
	A.	Stays as it is	B.	Tends to Bernoulli distribution	
	C.	Tends to hypergeometric distribution	D.	tends to Poisson Distribution	
662)	Let X be the number of heads obtained in 40 independent tosses of a fair coin. Then X is a Binomial random variable with				B
	A.	n = 40, p = 0	B.	n = 40, p = 0.5	
	C.	n = 0.5, p = 40	D.	n = 41, p = 0.5	
663)	What is the range of a Geometric random variable				C
	A.	All integers	B.	All positive integers	
	C.	All non-negative integers	D.	All negative integers	
664)	If 'λ' is the mean of a Poisson Distribution, then variance is given by _____				A
	A.	λ	B.	λ^2	
	C.	σ	D.	σ^2	
665)	Poisson distribution is applied for _____				B
	A.	Continuous random variable	B.	Discrete random variable	
	C.	Irregular Random Variable	D.	Uncertain Random Variable	
666)	In a Poisson distribution				A
	A.	Mean = Variance	B.	Mean = Standard deviation	
	C.	Mean < Standard deviation	D.	Mean < Variance	
667)	Two random variables X and Y are said to be independent if:				C
	A.	E(XY)=1	B.	E(XY)=0	
	C.	E(XY)=E(X).E(Y)	D.	E(XY)= Any constant value	
668)	A random variable assuming only a finite number of values is called:				A
	A.	Discrete random variable	B.	Continuous random variable	
	C.	Distributional random variable	D.	None of these	
669)	A random variable is also called				D
	A.	Constant	B.	Variable	
	C.	Attribute	D.	Chance variable	
670)	A quantity which can vary from one individual to another is called				B
	A.	Constant	B.	Variable	
	C.	Data	D.	None of these	
671)	The lifetime of a motor bike battery is				A
	A.	Continuous variable	B.	Discrete variable	
	C.	Qualitative variable	D.	Random Variable	
672)	A variable whose value is determined by the outcome of a random experiment is called				B
	A.	Randomized	B.	Random variable	
	C.	Experimental variable	D.	None of these	
673)	The sum of probabilities of a discrete random variable is				D
	A.	Zero	B.	Four	
	C.	Three	D.	One	
674)	While tossing 3 coins, the values that a random variable (number of heads) can take				B
	A.	1, 2, 3	B.	0, 1, 2, 3	
	C.	1, 2, 3, 4	D.	0, 1, 2, 3, 4	
675)	The speed of the vehicle is an example of				C
	A.	Discrete variable	B.	Qualitative variable	
	C.	Continuous variable	D.	None of these	
676)	If ***Var(X) = 4***, then ***Var(3X+5)*** is equal to				A
	A.	36	B.	12	

	C.	41	D.	17	
677)	If ***X*** is random variable, then ***Var(2 – 3X)*** is equal to				D
	A.	*Var(2) – 3Var(X)*	B.	*2 – 3Var(X)*	
	C.	*2 – 9Var(X)*	D.	*9Var(X)*	
678)	If ***X*** and ***Y*** are independent then ***Var(X – Y)***				C
	A.	*Var(X) – Var(Y)*	B.	*Var(X) . Var(Y)*	
	C.	*Var(X) + Var(Y)*	D.	None of these	
679)	If ***Var(X/3)=*** _____				A
	A.	*1/9 Var(X)*	B.	*1/3 Var(X)*	
	C.	*1/6 Var(X)*	D.	None of these	
680)	For a continuous random variable, the area under the probability distribution curve between any two points is always				C
	A.	Greater than one	B.	Less than zero	
	C.	In the range zero and one	D.	Equal to one	
681)	A continuous random variable is a random variable that can				B
	A.	Assess only countable values	B.	Asses any value in one or more intervals	
	C.	Have no random sample	D.	Assume no continuous frequency	
682)	Var (X+4) = _____				C
	A.	Var(X) + 16	B.	Var(X) + 4	
	C.	Var(X)	D.	None of these	
683)	The probability that a continuous random variable assumes a single value is				A
	A.	Equal to zero	B.	Between zero and one	
	C.	Greater than one	D.	Less than one	
684)	The probability that a discrete random variable assumes a single value is				B
	A.	Equal to zero	B.	Between zero and one	
	C.	Greater than one	D.	Less than one	
685)	In generating random numbers the probability of each digit/number is				A
	A.	Equal	B.	Unequal	
	C.	Remains constant	D.	None of these	
686)	The set of all possible outcomes of a random experiment is called				C
	A.	Population	B.	Sample	
	C.	Sample space	D.	Empty set	
687)	Discrete data is usually generated by the process				B
	A.	Measurements	B.	Counting	
	C.	Both by counting and measurements	D.	None of these	
688)	The number of deaths in a road accident is an example of _____ variable				A
	A.	Discrete	B.	Continuous	
	C.	Constant	D.	None of these	
689)	Random numbers are generated by some				B
	A.	Continuous process	B.	Random process	
	C.	Automatically generated	D.	None of these	
690)	A variable which takes measurable values is called a				C
	A.	Constant	B.	Discrete variable	
	C.	Continuous variable	D.	None of these	
691)	In a family with two children, how many can be girls?				D
	A.	0,1	B.	2	
	C.	0,1,2,3	D.	0,1,2	
692)	Usually measurements give rise to _____ data				B
	A.	Discrete	B.	Continuous	
	C.	Constant	D.	Qualitative	
693)	If "a" is constant than ***Var(a)*** is				A
	A.	0	B.	1	
	C.	a	D.	a^2	
694)	If ***Var(X) = 2*** and ***Var(Y) = 5***, and if ***X*** and ***Y*** are independent variables, then ***Var(2X – Y)*** = _____				B
	A.	1	B.	13	
	C.	3	D.	-1	
695)	$E(x - \mu)$ = _____				A

	A.	Zero	B.	Mean	
	C.	Variance	D.	Standard deviation	
696)	$E(X-\mu)^2$ = _____				C
	A.	Zero	B.	Mean	
	C.	Variance	D.	Standard deviation	
697)	If $P(X=10) = 1/10$ then $E(X)$ is				B
	A.	10	B.	1	
	C.	1/100	D.	Zero	
698)	If X and Y are random variable then E(X – Y) is equal to				A
	A.	$E(X) - E(Y)$	B.	$E(X) + E(Y)$	
	C.	$E(X) \cdot E(Y)$	D.	$E(X) - Y$	
699)	If "a" is any constant, then $E(a)$ is				B
	A.	a	B.	Zero	
	C.	a/n	D.	None of these	
700)	What does Poisson Distribution describe?				D
	A.	Future events	B.	Total number of events	
	C.	Common events	D.	Rare events	
701)	If X has a binomial distribution with parameter **n** and **p** then **X/n** has the variance:				D
	A.	npq	B.	n^2pq	
	C.	pq/2	D.	pq/n	
702)	The distribution in which the probability of each successive draw varies is:				A
	A.	Hypergeometric	B.	Geometric	
	C.	Binomial	D.	Normal	
703)	Total area under the Normal curve is:				B
	A.	Undefined	B.	Unity	
	C.	Zero	D.	0.5	
704)	The approximate relation between M.D. about mean and S.D. of a normal distribution:				A
	A.	$M.D. = \frac{4}{5}\sigma$	B.	$M.D. = \frac{5}{4}\sigma$	
	C.	$M.D. = \sigma$	D.	$M.D. = \frac{2}{3}\sigma$	
705)	The area under the standard normal curve between the lines z = ±1.96 is:				A
	A.	95 percent	B.	90 percent	
	C.	5 percent	D.	10 percent	
706)	If Z is standard normal variate, the proportion of items lying above Z=0 is				B
	A.	1	B.	0.5	
	C.	1.645	D.	0.95	
707)	If X~N(8, 64), the standard normal variate Z will be:				D
	A.	$Z = \frac{X-64}{8}$	B.	$Z = \frac{X-8}{64}$	
	C.	$Z = \frac{8-X}{8}$	D.	$Z = \frac{X-8}{8}$	
708)	For distribution Function $F(X)$, $F(-\infty)$=_____				A
	A.	0	B.	-1	
	C.	1	D.	Undefined	
709)	For distribution Function $F(X)$, $F(+\infty)$=_____				C
	A.	0	B.	-1	
	C.	1	D.	Undefined	
710)	The probability function is always				B
	A.	Negative	B.	Non-negative	
	C.	Infinity	D.	None of these	
711)	The distribution function $F(X)$ is represented by				D
	A.	P(X)	B.	$P(X \geq x)$	
	C.	P(X=x)	D.	$P(X \leq x)$	
712)	For a random variable X, E(X) is				C
	A.	Harmonic mean	B.	Geometric mean	
	C.	Arithmetic mean	D.	None of these	

713)	The range of the random variable follows normal distribution is:				D
	A.	0 to n	B.	0 to ∞	
	C.	-1 to +1	D.	-∞ to +∞	
714)	In normal distribution				A
	A.	Mean = Median = Mode	B.	Mean < Median < Mode	
	C.	Mean > Median > Mode	D.	None of these	
715)	Which of the following is true for Normal curve:-				D
	A.	Symmetrical	B.	Unimodal	
	C.	Bell-shaped	D.	All of these	
716)	In a normal curve, the highest point on the curve occurs at:-				D
	A.	Mean	B.	Median	
	C.	Mode	D.	All of these	
717)	The normal curve is symmetrical and for symmetrical distribution, the values of all odd order moments about mean will always be:				A
	A.	Zero	B.	Undefined	
	C.	One	D.	None of these	
718)	If $X \sim N(\mu, \sigma^2)$, the points of inflection of normal distribution are:-				C
	A.	± σ	B.	± μ	
	C.	μ ± σ	D.	σ ± μ	
719)	The Quartile deviation of the Normal distribution is:-				C
	A.	4/5	B.	4/5 σ	
	C.	2/3 σ	D.	2/3	
720)	The value of **"e"** in the pdf of Normal distribution is approximately equal to				A
	A.	2.7183	B.	2.1783	
	C.	2.8173	D.	2.1416	
721)	The value of **"π"** in the pdf of Normal distribution is approximately equal to:				B
	A.	3.4116	B.	3.1416	
	C.	3.1614	D.	3.6416	
722)	If $X \sim N(\mu, \sigma^2)$, the standard normal variate is distributed as:-				B
	A.	$N(1,0)$	B.	$N(0,1)$	
	C.	$N(\mu, 0)$	D.	$N(0, \sigma^2)$	
723)	The coefficient of skewness of a normal distribution is:				C
	A.	Positive	B.	Negative	
	C.	Zero	D.	Three	
724)	The Mode of the Normal distribution is				A
	A.	Equal to zero	B.	Less than zero	
	C.	Greater than zero	D.	Exactly one	
725)	The normal probability density curve is symmetrical about the mean, This means that P(X<μ) = P(X>μ) is equal to				B
	A.	0	B.	0.5	
	C.	1	D.	None of these	
726)	The skewness and kurtosis of the normal distribution are respectively:				D
	A.	Zero and One	B.	One and One	
	C.	Zero and Three	D.	Zero and Zero	
727)	If $X \sim N(100,64)$, then the standard deviation σ is				C
	A.	100	B.	64	
	C.	8	D.	10	
728)	If $Z \sim N(0,1)$, the coefficient of variation is equal to:				D
	A.	Zero	B.	One	
	C.	100%	D.	Infinity	
729)	The points of inflection of the standard normal distribution lie at:				C
	A.	-1 to 0	B.	0 and +1	
	C.	-1 to +1	D.	μ and σ	
730)	If $Z \sim N(0,1)$, then μ_4 is equal to:				C
	A.	0	B.	1	
	C.	3	D.	σ^4	
731)	If X is a normal random variable having mean μ , then E | X - μ | is equal to:-				A

	A.	Mean deviation	B.	Standard deviation	
	C.	Quartile deviation	D.	Variance	
732)	The range of standard normal distribution is:				B
	A.	0 to n	B.	$-\infty$ to $+\infty$	
	C.	-1 to +1	D.	None of these	
733)	If $Z \sim N(0,1)$, then β_2 is equal to:				B
	A.	0	B.	3	
	C.	$3\sigma^4$	D.	σ^2	
734)	The median of a normal distribution corresponds to a value of Z is:				A
	A.	0	B.	1	
	C.	0.5	D.	-0.5	
735)	For a normal distribution with $\mu = 10$, $\sigma = 2$, the probability of a value greater than 10 is:				B
	A.	0	B.	0.50	
	C.	1	D.	None of these	
736)	Which of the following mentioned standard Probability density functions is applicable to discrete Random Variables?				D
	A.	Normal distribution	B.	Rayleigh distribution	
	C.	Exponential distribution	D.	Poisson distribution	
737)	What is the area under a conditional Cumulative density function?				C
	A.	0	B.	Infinity	
	C.	1	D.	Changes with Cumulative Distribution Function	
738)	When do the conditional density functions get converted into the marginally density functions				B
	A.	Only if random variables exhibit statistical dependency	B.	Only if random variables exhibit statistical independency	
	C.	Only if random variables exhibit deviation from its mean value	D.	If random variables do not exhibit deviation from its mean value	
739)	A table with all possible value of a random variable and its corresponding probabilities is called _____				D
	A.	Probability Mass Function	B.	Probability Density Function	
	C.	Cumulative distribution function	D.	Probability distribution	
740)	The expected value of a discrete random variable '**x**' is given by				B
	A.	x P(x)	B.	Σ x P(x)	
	C.	Σ P(x)	D.	1	
741)	Out of the following values, which one is not possible in probability?				D
	A.	P(x) = 1/4	B.	Σ x P(x) = 2	
	C.	P(x) = 0.5	D.	P(x) = -0.5	
742)	The standard normal curve is symmetric about the value _____				D
	A.	0.5	B.	1	
	C.	∞	D.	0	
743)	Normal Distribution is also known as _____				C
	A.	Cauchy's Distribution	B.	Laplacian Distribution	
	C.	Gaussian Distribution	D.	Lagrangian Distribution	
744)	In Normal distribution, the highest value of ordinate occurs at _____				A
	A.	Mean	B.	Variance	
	C.	Extremes	D.	Same value occurs at all points	
745)	In a normal distribution, about 95% of observations are				B
	A.	within one standard deviation of the mean	B.	within two standard deviations of the mean	
	C.	within three standard deviations of the mean	D.	included in computing the mean.	
746)	The z-score is				D
	A.	The number of standard errors between the mean and some observation	B.	The difference between the sample mean and population mean	
	C.	The width of the 95% confidence interval	D.	The number of standard deviations an observation is from the mean	
747)	Approximately what area is covered under the Normal distribution curve between ± 3 standard deviation?				D
	A.	68.27%	B.	95.45%	
	C.	99.99%	D.	99.73%	
748)	Which of the following is false about binomial probabilities?				B
	A.	Their distributions may be approximately symmetric	B.	The probability of success must be 0.50	
	C.	Events musts be independent	D.	Trials must be fixed	

749)	Which one of these variables is a continuous random variable?				C
	A.	The number of tattoos a randomly selected person has	B.	The number of women taller than 68 inches	
	C.	The time it takes a randomly selected student to complete an exam.	D.	The number of correct guesses on a multiple choice test	
750)	Which one of these variables is a binomial random variable?				C
	A.	Time it takes a randomly selected student to complete a multiple choice exam	B.	Number of textbooks a randomly selected student bought this term	
	C.	Number of women taller than 68 inches in a random sample of 5 women	D.	Number of CDs a randomly selected person owns	
751)	A medical treatment has a success rate of 0.8. Two patients will be treated with this treatment. Assuming the results are independent for the two patients, what is the probability that neither one of them will be successfully cured?				D
	A.	0.5	B.	0.16	
	C.	0.64	D.	0.04	
752)	If Y is a random variable with mean μ, then $\boldsymbol{E(Y-\mu)^r}$ is known as				C
	A.	Variance	B.	r^{th} raw moment	
	C.	r^{th} central moment	D.	None of these	
753)	The moment generating function of Binomial distribution is:				B
	A.	$(q+pe^{tn})$	B.	$(q+pe^{t})^{n}$	
	C.	$(q+pe^{t})^{-n}$	D.	$(q+pe^{t})$	
754)	If the joint p.d.f of two random variables X and Y is defined as, $f(x,y) = x + y$, $0 \leq x,y \leq 1$ and zero otherwise. What is the marginal distribution of X?				A
	A.	$x + 1/2$	B.	$x + 1/4$	
	C.	$x + y + 1$	D.	None of these	
755)	If the random variable takes negative values, then the negative values will have				D
	A.	Constant probabilities	B.	Negative probabilities	
	C.	Zero probabilities	D.	None of these	
756)	If $\boldsymbol{X}$ is a random variable, then $\boldsymbol{E(e^{tX})}$ is known as				B
	A.	Probability generating function	B.	Moment generating function	
	C.	Characteristic function	D.	None of these	
757)	If F(x) is distribution function of a discrete random variable X, then F(5) – F(2) is equal to				C
	A.	$P(2 < X \leq 5)$	B.	$P(2 \leq X < 5)$	
	C.	$P(2 \leq X \leq 5)$	D.	$P(2 < X < 5)$	
758)	Let X is Poisson(a) and Y is Poisson(b) be two independent random variables. Consider a random variable Z = X + Y. Then Z is				D
	A.	Poisson(a/b)	B.	Poisson(a - b)	
	C.	Poisson(a.b)	D.	Poisson(a + b)	
759)	Which of the following is true?				B
	A.	Negative binomial is special case of Geometric distribution	B.	Geometric is special case of Negative binomial distribution	
	C.	Both A and B	D.	Neither A nor B	
760)	When can we use a normal distribution to approximate a binomial distribution?				D
	A.	When n is greater than 30	B.	When *np* is greater than or equal to 5	
	C.	When *nq* is greater than or equal to 5	D.	When both *np* and *nq* are greater than or equal to 5	
761)	Match the following binomial probability with its corresponding normal distribution probability statement after a continuity correction. **P(x > 25)**				A
	A.	$P(x \geq 25.5)$	B.	$P(x \leq 25.5)$	
	C.	$P(x \geq 24.5)$	D.	$P(x \leq 24.5)$	
762)	An oil company conducts a geological study that indicates that an exploratory oil well should have a 20% chance of striking oil. The company is interested to find the probability that the first strike comes on the third well drilled. Which distribution will be used?				B
	A.	Negative binomial distribution	B.	Geometric distribution	
	C.	Binomial distribution	D.	Bernoulli distribution	
763)	An oil company conducts a geological study that indicates that an exploratory oil well should have a 0.25 probability of striking oil. The company is interested to find the probability that the 3^{rd} strike comes on the 6^{th} well drilled. Which distribution will be used?				A
	A.	Negative binomial distribution	B.	Geometric distribution	

	C.	Binomial distribution	D.	Bernoulli distribution	
764)	If X follows Geometric distribution with parameter p (probability of success) then the Mean of X is				C
	A.	P	B.	np	
	C.	1/p	D.	p^2	
765)	A continuous probability can be represented by				B
	A.	Constant	B.	Graph	
	C.	Table	D.	None of these	
766)	In normal distribution, the proportion of observations that lies between 1 standard deviations of the mean is closest to				B A. C.
	A.	0.5-	A.	0.5-	
	C.	0.99	C.	0.99	
767)	The distribution of square of standard normal random variable will be				B A. C.
	A.	F	A.	F	
	C.	Standard Normal	C.	Standard Normal	
768)	In a binomial experiment with three trials, the binomial random variable can take				A A. C.
	A.	4 values	A.	4 values	
	C.	2 values	C.	2 values	
769)	A random variable X has a binomial distribution with n = 9, the variance of X is				B A. C.
	A.	3pq	A.	3pq	
	C.	$3\sqrt{pq}$	C.	$3\sqrt{pq}$	
770)	The hyper geometric distribution has ---- parameters				C A. C.
	A.	1	A.	1	
	C.	3	C.	3	
771)	Let X be a random variable with Var(X)=9 then SD(2X)= ------				D
	A.	18	B.	36	
	C.	3	D.	6	
772)	The exponential curve is also like ---- curve				A
	A.	Power	B.	Logarithmic	
	C.	Semi Logarithmic	D.	Inverse	
773)	If P(A and B) = P(A/B) . P(B) then both events are				A
	A.	Dependent	B.	Independent	
	C.	Mutually exclusive	D.	Not Known	
774)	In any normal distribution, the proportion of observations that are outside ±1 standard deviation of the mean is closest to				B
	A.	0.05	B.	0.32	
	C.	0.68	D.	0.95	
775)	If $P(A \cap B) = \phi$ then $P(A \cup B) =$_____				B
	A.	0	B.	$P(A)+P(B)-P(AB)$	
	C.	$P(A)+P(B)$	D.	None of these	
776)	When an event is certain to occur, its Probability is				B
	A.	0	B.	1	
	C.	0.5	D.	None of these	
777)	In binomial probability distribution, the dependents of standard deviations must includes				D
	A.	probability of q	B.	probability of p	
	C.	Number of trials	D.	All of these	
778)	In binomial distribution, the formula of calculating standard deviation is				C
	A.	square root of p	B.	square root of pq	
	C.	square root of npq	D.	square root of np	

779)	The formula of mean of uniform or rectangular distribution is as				B
	A.	mean = 4(b + a)/2b	B.	mean = (b + a)/2	
	C.	mean = (b - 2a)/4	D.	mean = (2a + 2b)/2a	
780)	The normal distribution is also classified as				A
	A.	Gaussian distribution	B.	Poisson distribution	
	C.	Bernoulli's distribution	D.	weighted average distribution	
781)	The mean deviation of a normal distribution is				B
	A.	$\frac{5}{4}\sigma$	B.	$\frac{4}{5}\sigma$	
	C.	$\frac{2}{5}\sigma$	D.	None of these	
782)	The chi-square distribution is a special case of				B
	A.	Beta distribution	B.	Normal distribution	
	C.	Exponential distribution	D.	Gamma distribution	
783)	Which of the distribution have larger variance than its mean				C
	A.	Binomial	B.	Hypergeometric	
	C.	Negative binomial	D.	None of these	
784)	For Cauchy distribution which of the following is true.				A
	A.	Mean does not exist	B.	Variance does not exist	
	C.	2^{nd} moment does not exist	D.	None of these	
785)	For Beta distribution of 2^{nd} kind, the range of X is				C
	A.	$X \in (0,1)$	B.	$X \in (1,0)$	
	C.	$X \in (-\infty, \infty)$	D.	$X \in (0, \infty)$	
786)	Mathematical simulation techniques use to generate the _____ number				D
	A.	Prime	B.	Odd	
	C.	Even	D.	Random	
787)	If a random variable X has probability density function $f(x) = \begin{cases} \frac{3}{50}\left(x^2 - 4x + 5\right) & 0 \le x \le 5 \\ 0 & x < 0 \text{ or } x > 5 \end{cases}$ then the Mode of X is:				B
	A.	0	B.	1	
	C.	2.5	D.	3.125	
788)	The amount of time a patient waits in a doctor's office is an example of _____.				D
	A.	the normal distribution	B.	the binomial distribution	
	C.	A discrete random variable	D.	A continuous random variable	
789)	The dispersion of the distribution of a random variable is measured by the:				D
	A.	Mean	B.	Median	
	C.	Expected value	D.	Standard deviation	
790)	The temperature in a Faisalabad city has a uniform distribution with a range from 78 degrees to 95 degrees. What is the mean of this distribution?				B

	A.	17	B.	86.5	
	C.	95	D.	84.5	
791)	When two events can not occur at the same time they are said to be _____ event.				B
	A.	Independent	B.	Mutually exclusive	
	C.	Random	D.	Both A and B	
792)	Let X be a random variable with Var(X)=7 then Var(2X)=____.				C
	A.	13.69	B.	15.70	
	C.	28.00	D.	17.40	
793)	If $P(A \cap B) = \phi$ then $P(A \cup B) =$_____				C
	A.	0	B.	$P(A) + P(B) - P(AB)$	
	C.	$P(A) + P(B)$	D.	None of these	
794)	The mean of the Poisson distribution is 9 then its Standard deviation is				A
	A.	3	B.	81	
	C.	74.6	D.	1.31	
795)	In normal distribution, the proportion of observations that lies between 1 standard deviations of the mean is closest to				B
	A.	0.5-	B.	0.68	
	C.	0.99	D.	0.95	
796)	For Beta distribution of 1st kind, the range of X is				D
	A.	$X \in (0,1)$	B.	$X \in (1,0)$	
	C.	$X \in (-\infty, \infty)$	D.	$X \in (0, \infty)$	
797)	The term "sample space" is used for				A
	A.	All possible outcomes	B.	All possible successes	
	C.	probability	D.	sample	
798)	The parameters of hypergeometric distribution are ----- Note: ***N*** is population size, ***n*** is sample size, ***p*** is the successes' probability K is number of success states in the population, ***k*** is the number of observed successes				C
	A.	***N, n,*** and ***p***	B.	***n*** and ***p***	
	C.	***N, K,*** and ***n***	D.	***n*** and ***k***	
799)	If ***N*** is population size, ***n*** is sample size, ***p*** is successes' probability, ***K*** is number (#) of success states in population, and ***k*** is the # of observed successes. Then parameters of binomial distribution are -----				B
	A.	***N, n,*** and ***p***	B.	***n*** and ***p***	
	C.	***N, K,*** and ***n***	D.	***n*** and ***k***	
800)	Bayes' theorem				D
	A.	Is an example of subjective probability	B.	Can assume of value less than 1	
	C.	is used to revise probability based on additional information	D.	All of these	