ADDENDUM A: FIRST SEMESTER ASSIGNMENTS

A.1 Assignment 01

ONLY FOR SEMESTER 1 STUDENTS
ASSIGNMENT 01
Unique Nr.: 657656
Fixed closing date: 02 March 2020

Please note that as explained before, examination admission for this module is by obtaining at least 40% in your year mark. You year mark is the average of your best two assignments.

Question 1

What is a summary measure that is calculated from a sample to describe a characteristic of a population?

(1) Boxplot.

(2) Data.

(3) Parameter.

(4) Population.

(5) Statistic.

Question 2

Which one of the following statements is incorrect?

(1) A population is a complete set of objects in the study while a sample is a subset of a population.

(2) A statistic is a property of a population while a parameter is a property of a sample.

(3) Data from a sample are in a form of variables, which can either be numeric or categorical.

(4) Quantitative data are numeric and can either be discrete or continuous.

(5) Qualitative data are categorical and uses labels to identify attributes.

Question 3

Which one of the following statements is incorrect with regard to qualitative or categorical data?

(1) Categorical data is measured on an ordinal or nominal scale.

(2) The frequency or count of each element can be determined.

(3) The mode can be determined.

(4) The mean can be calculated.

(5) None of the above.

QUESTIONS 4 TO 8 ARE BASED ON THE INFORMATION AND TABLE 1.1 BELOW.

Autism Spectrum Disorder is a group of complex brain development disorders. The information on 4 separate varibles below is based on a sample of 500 people living with Autism Spectrum Disoder (ASD) in South Africa. Use this inofrmation to answer questions 4 to 8.

Type of ASD	Age at diagnosis (in months)	Number of people	Dominant form of impairment
A	18.5	108	Social Interaction
B	22	44	Communication
C	20	206	Repetitive behaviour
D	30.5	85	Sensory perception
E	25	57	Blend of impairments

You are tasked with helping Autism South Africa understand this complexity by answering the following questions.

Question 4

Suppose we are interested in the average age of diagnosis for everyone living with ASD in South Africa. Which one of the following statements is incorrect?

(1) A sample refers to the 500 people whose data was collected.

(2) A population refers to everyone living with ASD in South Africa.

(3) The age of diagnosis is a variable.

(4) The mean (average) age of diagnosis calculated from the sample is a parameter.

(5) None of the above.

Question 5

Which one of the following statements is correct?

(1) Type of ASD is a qualitative variable with ordinal scale of measurement.

(2) Type of ASD is quantitative variable with ordinal scale of measurement.

(3) The age at diagnosis is a quantitative variable with ordinal scale of measurement.

(4) The age at diagnosis is a quantitative discrete variable.

(5) The age at diagnosis is a quantitative continuous variable.

Question 6

Which one of the following statements is correct?

(1) Number of people with ASD is a quantitative continuous variable.

(2) Number of people with ASD is a quantitative discrete variable.

(3) Number of people with ASD is a categorical variable with ordinal scale of measurement.

(4) Dominant form of impairment is a categorical variable with ordinal scale of measurement.

(5) Dominant form of impairment is a quantitative discrete variable.

Question 7

Consider statements A to D below.

(A) Type of ASD is measured on an ordinal scale.

(B) Age of diagnosis is measured on a ratio scale.

(C) Number of people is measured on a ratio scale.

(D) Dominant form of impairment is measured on a nominal scale.

Which statement(s) are correct?

(1) Only A and B.

(2) Only A, B and C.

(3) Only B, C and D.

(4) Only B and C.

(5) A, B, C and D.

Question 8

Suppose we are interested in the extent to which children with ASD present repetitive behaviour (mild, moderate, severe and extremely severe). What type of variable is repetitive behaviour?

(1) Qualitative nominal.

(2) Qualitative ordinal.

(3) Quantitative continuous.

(4) Quantitative discrete.

(5) Quantitative ordinal.

Question 9

Consider statements A to D below.

(A) Data can be summarised using both tabular and graphical methods.

(B) Classes in a frequency distribution are non-overlapping.

(C) Frequency refers to the proportion of items in a class while relative frequency refers to the actual count per class.

(D) The modal class is the class with the highest frequency.

Which statement(s) are incorrect?

(1) Only D.

(2) Only B.

(3) Only C.

(4) Only A and B.

(5) Only B, C and D.

Question 10

The sum of relative frequencies for all class will always equal?

(1) 1.

(2) 100.

(3) The number of classes.

(4) The sample size.

(5) The total frequency.

Consider type of ASD and estimated number of people. See Table 1.1a below to answer question 11, 12 and 13.

Table 1.1a

Type of ASD	Number of people
A	108
B	44
C	206
D	85
E	57

Question 11

Table 1.1a is referred to as a _____?

(1) Frequency distribution.

(2) Relative frequency.

(3) Frequency.

(4) Cumulative frequency.

(5) None of the above.

Question 12

What is the appropriate graphical form to summarise the data in Table 1.1a?

(1) Scatter plot.

(2) Stem-and-leave diagram.

(3) Bar or pie chart.

(4) Histogram.

(5) All of the above plots are appropriate.

Question 13

Which one of the following statements is incorrect?

1. The total number of people living with ASD is 500.
2. The frequency of people living with type B ASD is 0.09.
3. The relative frequency of people living with type C ASD is 0.41.
4. The frequency of people living with type D ASD is 85.
5. None of the above.

Question 14

Consider only the age of diagnosis. See Table 1.1b below.

Table 1.1b

Age at diagnosis (in months)
18.5
22
20
30.5
25

Which one of the following stem and leave diagrams best represents the data? **Hint: round up 18.5 and 30.5.**

(1)

1	9
2	0 2 5
3	1

(2)

1	9
2	2 0 5
2	1

(3)

10	9
20	0 2 5
30	1

(4)

1	8.5
2	0 2 5
3	0.5

(5)

1	19
2	20 2 5
3	31

Question 15

Which one of the following descriptive statistics is not a measure of central tendency or location?

(1) Mean.

(2) Median.

(3) Mode.

(4) Range.

(5) All the above measures describe central tendency or location.

Question 16

Which one of the following statements is correct with regard to the coefficient of variation?

(1) It is a measure of location.

(2) It is the difference between the largest and smallest value in a dataset.

(3) It is the sample standard deviation squared

(4) It is the sample standard deviation expressed as a percentage of the mean.

(5) It is the mean squared.

Question 17

What is the observation that occurs most frequently in a dataset?

(1) Mean deviation.

(2) Standard deviation.

(3) Mean.

(4) Median.

(5) Mode.

Consider only the number of people living with ASD and answer questions 18 to 22. See Table 1.1c below.

Table 1.1c

Number of people
108
44
206
85
57

Question 18

Which one of the following statements is incorrect?

(1) The mean and the median are equal.

(2) The mean is less than the median.

(3) The mode is zero.

(4) There is no mode.

(5) None of the above.

Question 19

What is the sample standard deviation?

(1) 64.25.

(2) 57.46.

(3) 4127.50.

(4) 3302.

(5) None of the above.

Question 20

The coefficient of variation is:

(1) 64.25%.

(2) 57.46%.

(3) 4127.50%.

(4) 3302%.

(5) 100%.

Question 21

Which of the following statements is correct?

(1) The median is 85 and equals the mean.

(2) The value of Q_1 is 57.

(3) The position of Q_2 is 2.

(4) The value of Q_3 is 157.

(5) The distribution of the number of people living with ASD is symmetric.

Question 22

The interquartile range is _____?

(1) Equals to the range.

(2) Always less than the range.

(3) 149.

(4) 51.

(5) 0.

Consider the following information to answer question 23 and 24. Let A represent the event that someone lives with ASD and B represent the event that someone does not live with ASD.

Question 23

Which of the following statements is correct?

(1) $P(A \text{ and } B) = P(A) * P(B)$.

(2) $P(A|B) = P(A)$.

(3) $P(B) \neq P(A^C)$.

(4) Events A and B are mutually exclusive.

(5) Events A and B are independent.

Question 24

Suppose further that $P(A) = 0.96$ and $P(B) = 0.04$. $P(A \text{ or } B) =?$

(1) 0.92.

(2) 0.0384.

(3) 1.

(4) 0.96.

(5) 0.04.

Autism South Africa collected the following information on specialists consulting with children living with Autism Spectrum Disorder (ASD). Table 1.2 below shows the number of boys and girls consulting with different specialists. Complete Table 1.2 below to answer question 25 to 30.

Table 1.2 *Specialist consulting with children with ASD.*

	Boy	Girl	Total
Speech Therapist (ST)	90	30	
Neuropsychologist (NP)	45	15	
Psychiatrist (P)	30		
Total		**55**	

Question 25

What is the probability that a randomly chosen child who consulted with a Neuropsychologist is a boy i.e $P(NP$ and Boy$)$?

(1) $0.975.$

(2) $0.818.$

(3) $0.750.$

(4) $0.273.$

(5) $0.205.$

Question 26

What is the probability that a randomly chosen child either consulted with a Neuropsychologist or is a boy i.e $P(NP$ or Boy$)$?

(1) $0.975.$

(2) $0.818.$

(3) $0.750.$

(4) 0.273

(5) $0.205.$

Question 27

What is the probability that a randomly chosen child is a boy given that he consulted with a Neuropsychologist?

(1) $0.975.$

(2) $0.818.$

(3) $0.750.$

(4) $0.273.$

(5) $0.205.$

Question 28

Which one of the following statements is incorrect?

(1) $P(\text{Boy}) = 0.75.$

(2) $P(\text{Boy or Girl}) = 1.$

(3) $P(\text{Girl or } ST) = 0.795.$

(4) $P(ST) = 0.545.$

(5) $P(ST \text{or } NP) = 0.818.$

Question 29

Which one of the following statements is incorrect?

(1) $P(ST \text{ and Boy}) = 0.409.$

(2) $P(P \text{ and Girl}) = 0.045.$

(3) $P(ST|\text{Boy}) = 0.545.$

(4) $P(NP|\text{Girl}) = 0.068.$

(5) $P(NP|\text{Boy}) = P(NP|\text{Girl}) = P(NP).$

Question 30

Which one of the following statements is incorrect?

(1) The events Neuropsychologist and Boy are independent.

(2) The events Neuropsychologist and Girl are independent.

(3) The events Psychiatrist and Boy are dependant.

(4) Autism Spectrum Disorder in more common in boys than in girls.

(5) The events Speech Therapist and Psychiatrist are mutually exclusive.

A.2 Assignment 02

ONLY FOR SEMESTER 1 STUDENTS
ASSIGNMENT 02
Unique Nr.: 645093
Fixed closing date: 16 March 2020

Please note that as explained before, examination admission for this module is by obtaining at least 40% in your year mark. You year mark is the average of your best two assignments.

Consider the information below to answer questions 1 and 2.
A Speech Therapist knows that the number of children that consults with her on any given day is given by the following discrete probability distribution. Let X be the number of children consulting with the Speech Therapist on any given day and $P(X = x)$ be the associated probability.

x	1	2	3	4	5
$P\,(X = x)$	0.1	0.15	0.2	0.3	0.25

Question 1

Which of the following statements is incorrect?

(1) The probability that the Speech Therapist will consult with at least one child is 1.

(2) The probability that the Speech Therapist will consult with 1 child on any given day is 0.1.

(3) The probability that the Speech Therapist will consult with no children on any given day is 0.

(4) The probability that the Speech Therapist will consult with more than 5 children on any given day is 0.

(5) The expected number of children consulting with the speech Therapist on any given day is 4.

Question 2

What is the standard deviation of X?

(1) 11.01.

(2) 3.45.

(3) 3.32.

(4) 1.65.

(5) 1.28.

Question 3

Which one of the following is not a property of a binomial distribution?

(1) Each experiment has n trials.

(2) The n trials are independent of each other.

(3) Each trial has two possible outcomes that are mutually exclusive i.e. a success and a failure

(4) The probability of success remains the same for all trials.

(5) The probability of success is always $\frac{1}{2}$ because of there are two outcomes.

Autism South Africa has found that 50% of people with Autism Spectrum Disorder (ASD) struggle with social interaction. Assume we randomly select 6 people living with ASD and use the given information to answer questions 4 to 6.

Question 4

What is the probability that only 3 people with ASD will struggle with social interaction?

(1) 0.5.

(2) 0.25.

(3) 0.3125.

(4) 0.2344.

(5) None of the above.

Question 5

What is the probability that at most 1 person with ASD will struggle with social interaction?

(1) 0.1093.

(2) 0.8907.

(3) 0.0937.

(4) 0.0156.

(5) None of the above.

Question 6

Consider statements A to C below.

(A) The expected number of people with ASD struggling with social interaction is 3.

(B) The variance of number of people with ASD struggling with social interaction is 1.5.

(C) The standard deviation of number of people with ASD struggling with social interaction is 1.225.

Which statement(s) are correct?

(1) Only A.

(2) Only B.

(3) Only C.

(4) Only A and B.

(5) A, B and C.

The average number of adults with ASD consulting with a Neuropsychologist per day is Poisson distributed with the mean of 1.5. Use this information to answer questions 7 to 9.

Question 7

What is the probability that on any given day a Neuropsychologist will consult with only 1 adult with ASD?

(1) 0.3230.

(2) 0.3452.

(3) 0.2510.

(4) $\frac{1.5^1 \times e^{-1.5}}{1!}$.

(5) Need more information to calculate the probability.

Question 8

What is the probability that on any given day a Neuropsychologist will consult with at least 7 adults with ASD?

(1) 0.0013.

(2) 0.0009.

(3) 0.0008.

(4) 0.0006.

(5) None of the above.

Question 9

What is the variance of the number of adults with ASD consulting with a Neuropsychologist on any given day?

(1) 0.

(2) 1.5.

(3) 3.

(4) 4.5.

(5) None of the above.

Question 10

Which one of the following statements is incorrect with regard to the standard normal distribution?

(1) The standard normal distribution has a median of zero.

(2) The standard normal distribution is symmetric around the mean of zero.

(3) The standard normal distribution has a standard deviation of 1.

(4) The standard normal distribution has a variance of 1.

(5) The area to the right side of the mean is 1 and the area to the left side of the mean is also 1.

Question 11

Which if the following measures controls the spread or flatness of a normal distribution?

(1) Mean.

(2) Median.

(3) Mode.

(4) Variance.

(5) None of the above.

Consider the standard normal random variable $(\mathbf{Z})$ and answer questions 12 to 14.

Question 12

The area to the right of $Z = 0$ is _____.

(1) $1.$

(2) The same as the area to the left of $Z = 0.$

(3) Greater than the area to the left of $Z = 0.$

(4) Less than the area to the left of $Z = 0.$

(5) $0.$

Question 13

$P(Z \geq -2.88) = ?$

(1) $0.0020.$

(2) $0.001\,4.$

(3) $0.9986.$

(4) $0.9980.$

(5) None of the above.

Question 14

$P(-1.1 \leq Z \leq 1.9) = ?$

(1) 0.7285.

(2) 0.8126.

(3) 0.8356.

(4) 0.9713.

(5) None of the above.

QUESTIONS 15 TO 17 ARE BASED ON THE FOLLOWING INFORMATION.
Mmabatho the Social Scientist is interested in the reading speed of adults with ASD. She found out that the reading speed is normally distributed with a mean of 90 words per minutes (wpm) and a standard deviation of 9 wpm.

Question 15

What is the probability that a randomly selected adult with ASD will read at a speed of 110 wpm at most?

(1) 0.9868.

(2) 0.0132.

(3) 0.0139.

(4) -2.2.

(5) 2.2.

Question 16

What is the probability that the reading speed of an adult with ASD is between 90 wpm and 117 wpm?

(1) 0.9987.

(2) 0.9988.

(3) 0.5.

(4) 0.4987.

(5) 0.

Question 17

Mmabatho knows that the probability for a reading speed of at least x wpm for an adult with ASD is 0.2514 i.e. $P(X \geq x) = 0.2514$. What is the value of x?

(1) 83.97

(2) 96.03

(3) $90.$

(4) $0.67.$

(5) $-0.67.$

Question 18

Calculate the missing value for the probability $P(Z < ?) = 0.5.$

(1) $1.$

(2) $-1.$

(3) $0.$

(4) $0.5.$

(5) $-0.5.$

Question 19

A normally distributed population with 1000 observations has a mean of 100 and a standard deviation 10. A sample of size 200 is randomly drawn from the population.
What is the standard error of the sampling distribution?

(1) $10.$

(2) $0.707.$

(3) $0.05.$

(4) $0.1.$

(5) $0.316.$

Question 20

Consider a population with a mean of 70 and a standard deviation of 6. A random sample of size 36 is drawn. What is the probability that the sample mean is between 70.5 and 71.5 ?

(1) 0.9332.

(2) 0.5000.

(3) 0.6915.

(4) 0.2415.

(5) None of the above.

Question 21

Consider the information provided in question 20. What is the probability that the sample mean is at least 67.5?

(1) 2.5.

(2) 0.9938.

(3) 0.0062.

(4) -2.5.

(5) None of the above

Question 22

Remember Mmabatho the Social Scientist? Suppose she knows that reading time of adults with ASD is normally distributed with a population mean and standard deviation of 90 wpm and 18 wpm respectively. If she selects a sample size of 30, what is the probability that the average reading speed of an adult with ASD will be at least 95 wpm?

(1) 0.9357.

(2) 0.0643.

(3) 0.3897.

(4) 0.6103.

(5) None of the above.

Question 23

In a random sample of 100 people, 25 are classified as meeting a characteristic of interest i.e. success. Suppose the proportion of success is known to be 0.3. What is the sample proportion p and the standard error of the proportion σ_p?

(1) $p = 0.3$ and $\sigma_p = 0.0458$.

(2) $p = 0.3$ and $\sigma_p = 0.0433$.

(3) $p = 0.25$ and $\sigma_p = 0.0433$.

(4) $p = 0.25$ and $\sigma_p = 0.0458$.

(5) None of the above.

Autism South Africa knows the true population proportion of children with ASD in special need school is 0.74. Assume a sample size of 100 children with ASD is selected and use this information to answer question 24 and 25.

Question 24

What is the probability that the sample proportion of children with ASD in special need schools is at least 0.7?

(1) 0.5000.

(2) 0.1848.

(3) 0.8186.

(4) 0.6554.

(5) 0.3446.

Question 25

What is the probability that the sample proportion is between 0.7 and 0.84 i.e. $P(0.7 \leq p \leq 0.84) = ?$

(1) 0.9887.

(2) 0.9940.

(3) 0.9887.

(4) 0.7996.

(5) 0.1922.

A.3 Assignment 03

ONLY FOR SEMESTER 1 STUDENTS
ASSIGNMENT 03
Unique Nr.: 807054
Fixed closing date: 26 March 2020

Please note that as explained before, examination admission for this module is by obtaining at least 40% in your year mark. You year mark is the average of your best two assignments.

QUESTIONS 1 TO 4 ARE BASED ON THE FOLLOWING INFORMATION.

Consider the random variable X with a mean 100 and a **population standard deviation** of 9. Assume a sample of size 30 was used.

Question 1

Consider the statements below.

(A) The 90% confidence interval estimate for the population mean is $(97.2970;\ 102.7030)$.

(B) The 95% confidence interval estimate for the population mean is $(96.7794;\ 103.2206)$.

(C) The 99% confidence interval estimate for the population mean is $(95.7606;\ 104.2394)$.

Which statement(s) are correct?

(1) Only A.

(2) Only B.

(3) Only C.

(4) Only A and B.

(5) A, B and C.

Question 2

Assuming the sample size stays the same, what happens to the confidence interval estimate when the level of confidence increases?

(1) The confidence interval estimate doesn't change.

(2) The confidence interval estimate becomes wider.

(3) The confidence interval estimate becomes narrower.

(4) The confidence interval estimate converges to zero.

(5) None of the above.

Question 3

Which of the following statements is incorrect?

(1) When the sample size is 60 the 95% confidence interval estimate for the population mean is $(97.7227;\ 102.2773)$.

(2) When the sample size is 100 the 95% confidence interval estimate for the population mean is $(98.2360;\ 101.7640)$.

(3) When the sample size is 200 the 95% confidence interval estimate for the population mean is $(98.7527;\ 101.2473)$.

(4) When the sample size is 500 the 95% confidence interval estimate for the population mean is $(99.2111;\ 100.7889)$.

(5) None of the above.

Question 4

Assuming the level of significance stays the same, what happens to the confidence interval estimate when the sample size increases?

(1) The confidence interval estimate doesn't change.

(2) The confidence interval estimate becomes wider.

(3) The confidence interval estimate becomes narrower.

(4) The confidence interval estimate converges to zero.

(5) None of the above.

QUESTIONS 5, 6 AND 15 ARE BASED ON THE FOLLOWING INFORMATION.

Mmabatho the Social Scientist took a random sample of 30 adults with Autism Spectrum Disorder (ASD) and found their reading time to be normally distributed with sample mean and sample standard deviation of 90 wpm and 18 wpm respectively.

Question 5

What is the 99% confidence interval estimate of the population mean?

(1) $(80.9626; 99.0374)$.

(2) $(80.9429; 99.0571)$.

(3) $(81.9090; 98.0910)$.

(4) $(81.5213; 98.4787)$.

(5) $(82.3428; 97.6572)$.

Question 6

What is the upper limit for the 95% confidence interval estimate of the population mean?

(1) 96.7206.

(2) 96.4412.

(3) 83.5588.

(4) 83.2794.

(5) None of the above.

QUESTIONS 7, 8 AND 16 ARE BASED ON THE FOLLOWING INFORMATION.

Mmabatho randomly selected a sample of 100 children with ASD and found that only 70 of them are in special need schools.

Question 7

What is the 90% confidence interval estimate of the proportion of children with ASD in special need schools?

(1) (0.6102; 0.7898).

(2) (0.6246; 0.7754).

(3) (0.6038; 0.7962).

(4) (0.6678; 0.8122).

(5) None of the above.

Question 8

What is the lower limit for the 95% confidence interval estimate of the proportion of children in special need schools?

(1) 0.6102.

(2) 0.6246.

(3) 0.7754.

(4) 0.7898.

(5) None of the above.

Question 9

Which of the following is not a step in hypothesis testing?

(1) Formulate the null and alternative hypothesis.

(2) Specify the level of significance.

(3) Determine the test statistic, critical value and p-value.

(4) Determine the rejection region.

(5) None of the above.

Question 10

Botlhale, a practicing Statistician is tasked with investigating whether the population mean of a random variable X is different from 100. She has the following additional information to help her with the investigation: $\alpha = 0.01,\ \sigma = 3,\ \bar{X} = 99$ and $n = 100$.
Consider statements A to E below.

(A) $H_0 : \mu = 100$ against $H_1 : \mu \neq 100$

(B) $H_0 : \mu = 100$ against $H_1 : \mu < 100$

(C) The level of significance is 0.01.

(D) A Z–statistic is applicable.

(E) A t–statistics is applicable.

Which statement(s) are correct?

(1) Only A and B.

(2) Only A and C.

(3) Only A, C and D.

(4) Only A, C and E.

(5) Only B, C and B.

Question 11

What is the value of the test statistic?

(1) -3.33.

(2) 2.58.

(3) 1.96.

(4) 1.645.

(5) 0.0012.

Question 12

What is the critical value?

(1) -3.33.

(2) 2.58.

(3) 1.96.

(4) 1.645.

(5) 0.0012.

Question 13

What is the p-value?

(1) 0.05.

(2) 0.01.

(3) 0.0024.

(4) p–value < 0.002.

(5) None of the above.

Question 14

What is the decision with regard to the hypothesis and the conclusion about the population mean?

(1) We reject the null hypothesis and conclude that the population mean is equal to 100.

(2) We reject the null hypothesis and conclude that the population mean is not equal to 100.

(3) We do not reject the null hypothesis and conclude that the population mean is equal to 100.

(4) We do not reject the null hypothesis and conclude that the population mean is not equal to 100.

(5) We accept the null hypothesis.

Question 15

Botlhale and Mmabatho are collaborating to test the hypothesis that the mean reading time of adults with ASD is less that 100. **Consider information provided for question 5 and 6 and assume a 5% level of significance.**
Which of the following statements is incorrect?

(1) A one sided test is applicable.

(2) The critical value is 2.045.

(3) The value of the test statistic is -3.04.

(4) The p-value is less than 0.005.

(5) We reject H_0 and conclude that the mean reading time of adults with ASD is less than 100.

Question 16

Botlhale and Mmabatho are at it once again. This time they want to determine whether the true proportion of ASD children in special need schools in the population is 0.75. **Consider information provided for questions 7 and 8 and assume a 5% level of significance.**
Which of the following statements is incorrect?

(1) $H_0 : \pi = 0.75$ against $H_1 : \pi \neq 0.75$

(2) The critical value is 1.96.

(3) The value of the test statistic is -1.09.

(4) The p-value is 0.2502.

(5) We do not reject H_0.

Question 17

Consider a Chi-square random variable with 15 degrees of freedom and 0.1 level of significance. Which of the following test statistic values will result in rejection of the null hypothesis?

(1) 21.1.

(2) 18.5.

(3) 19.8.

(4) 23.5.

(5) 2.7.

Botlhale and Mmabatho's final project is to determine if specialist type and gender of children are independent of each other. Use 5% level of significance. Consider the contingency table below and answer questions 18 to 20.

Table 3.1 Specialist consulting with children with ASD.

	Boys	Girls	**Total**
Speech Therapist (ST)	90	30	
Neuropsychologist (NP)	45	15	
Psychiatrist (P)	30	10	
Total			

Question 18

Which of the following statements is incorrect?

(1) The null hypothesis is that specialist type and gender of children are independent of each other.

(2) The alternative hypothesis is that specialist type and gender of children are dependent.

(3) The expected frequency of boys consulting with Speech Therapists is 90.

(4) The observed frequency of boys consulting with Speech Therapists is 90.

(5) There are 3 degrees of freedom.

Question 19

What is the value of the test statistic and the critical value?

(1) $\chi^2 = 1$ and $\chi^2_{(\alpha,df)} = 5.991$.

(2) $\chi^2 = 1$ and $\chi^2_{(\alpha,df)} = 7.378$.

(3) $\chi^2 = 1$ and $\chi^2_{(\alpha,df)} = 7.815$.

(4) $\chi^2 = 0$ and $\chi^2_{(\alpha,df)} = 5.991$.

(5) $\chi^2 = 0$ and $\chi^2_{(\alpha,df)} = 7.378$.

Question 20

What is the decision with regard to the hypothesis and the conclusion about the two variables?

(1) We reject the null hypothesis and conclude that specialist type and gender of children are dependent.

(2) We reject the null hypothesis and conclude that specialist type and gender of children are independent of each other.

(3) We do reject the null hypothesis and conclude that specialist type and gender of children are independent of each other.

(4) We do reject the null hypothesis and conclude that specialist type and gender of children are dependent.

(5) We accept then null hypothesis.

ADDENDUM B: SECOND SEMESTER ASSIGNMENTS

B.1 Assignment 01

ONLY FOR SEMESTER 2 STUDENTS
ASSIGNMENT 01
Unique Nr.: 843102
Fixed closing date: 03 August 2020

Please note that as explained before, examination admission for this module is by obtaining at least 40% in your year mark. You year mark is the average of your best two assignments.

Consider the measures below and answer questions 1 and 2.

(A) μ

(B) σ

(C) $\bar{x}$

(D) s

Question 1

Which measures are considered as population parameters?

(1) Only A.

(2) Only B.

(3) Only A and B.

(4) Only A, B and C.

(5) Only C and D.

Question 2

Which measures are considered as sample statistics?

(1) Only A.

(2) Only B.

(3) Only A and B.

(4) Only A, B and C.

(5) Only C and D.

Question 3

Which one of the following statement is incorrect?

(1) Quantitative variables can either be discrete or continuous.

(2) Qualitative variables used labels to describe attributes.

(3) The mean, median and mode can be determined for a quantitative variable.

(4) The mean cannot be determined for a qualitative variable.

(5) None of the above.

Researchers from Africa Check (a fact checking organisation) are tasked with investigating fake news posts on Social Media. Questions 4 and 5 relates to these researchers.

Question 4

Which one of the following statements is incorrect?

(1) The position (Junior, Mid-Level and Senior) of the researcher is a qualitative nominal variable

(2) The gender of the researchers is a qualitative nominal variable.

(3) The age if the researchers is a quantitative continuous variable.

(4) The salary of the researchers is a quantitative continuous variable.

(5) None of the above.

Question 5

Which one of the following statements is incorrect with regard to scales of measurements?

(1) Researcher's position has an ordinal scale of measurement.

(2) Researcher's gender has a nominal scale of measurement.

(3) Researcher's age has a ratio scale of measurement.

(4) Researcher's salary has a ratio scale of measurement.

(5) None of the above.

Question 6

Consider the two variables below. (News Medium and Social Media Platform).

News Medium	Social Media Platform
Traditional Media	Facebook
Social Media	Twitter
	WhatsApp

Which one of the following statements is incorrect?

(1) News medium is qualitative variable.

(2) News medium has an nominal scale of measurement.

(3) Social Media platform is a quantitative variable.

(4) Social Media platform has a nominal scale of measurement.

(5) News Medium and Social Media Platform are both categorical variables.

Question 7

Which one of the following statements in incorrect?

(1) The number of likes a fake news post received is a quantitative discrete variable.

(2) The number of retweets a fake news post received is a quantitative discrete variable.

(3) The number of mentions of fake news tweet is a quantitative continuous variable.

(4) The duration in hours it takes a fake news post to go viral is a quantitative continuous variable.

(5) None of the above.

Consider Table 2.1 below and use the information to answer questions 8 to 11.

Table 2.1

Social Media Platform	Fake news post count
Facebook	1424
Twitter	988
WhatsApp	888

Question 8

Which one of the following statements is correct?

(1) Fake news post count is a qualitative nominal variable.

(2) Fake news post count is a categorical variable measured on an ordinal scale.

(3) Fake news post count is a quantitative discrete variable.

(4) Fake news post count is a quantitative continuous variable measured on a ratio.

(5) None of the above.

Question 9

Table 2.1 is referred to as a _____?

(1) Frequency distribution table.

(2) Relative frequency table.

(3) Frequency table.

(4) Cumulative frequency table.

(5) None of the above.

Question 10

The appropriate graphical form to summarise the data in Table 2.1 is_____?

(1) Histogram.

(2) Scatter plot.

(3) Bar Chart.

(4) Frequency distribution.

(5) None of the above.

Question 11

Which one of the following statements is incorrect?

(1) The relative frequency of facebook is 1424.

(2) The percentage frequency if facebook is 43.15%.

(3) The relative frequency of Twitter is 0.30.

(4) The percentage frequency of WhatsApp is 26.91%.

(5) None of the above.

Question 12

Consider graphical methods A to D below.

(A) Bar Chart.

(B) Histogram.

(C) Pie chart.

(D) Scatter plot.

(E) Stem-and-leave display.

Which graphical method(s) are most appropriate for quantitative data?

(1) Only A and B.

(2) Only B and D.

(3) Only B,D and E.

(4) A only.

(5) B only.

Question 13

Table 2.2 below shows a numeric frequency distribution of number of likes on a Facebook fake news post. **Use this information to answer questions 13 and 14.**

Likes	Cumulative Frequency
$400 < 800$	7
$800 < 1200$	21
$1200 < 1600$	26
$1600 < 2000$	29
$2000 < 2400$	30

The relative frequency for class $800 < 1200$ is:

(1) 0.47.

(2) 0.19.

(3) 21.

(4) 0.7.

(5) 14.

Question 14

What is the frequency of class $2000 < 2400$?

(1) 0.03.

(2) 30.

(3) 1.

(4) 0.97.

(5) None of the above.

Question 15

Which one of the following statements is incorrect?

(1) Graphs provide an overview of the profile of random variables.

(2) Descriptive statistics identify the location, dispersion and shape of the data.

(3) The location (central tendency) refers to where the data values are spread about the middle values.

(4) Three commonly used central location statistics are: average, median and mode.

(5) The second quartile equals the median.

Question 16

The middle value in an ordered array of numbers is the:

(1) Midpoint.

(2) Median.

(3) Mean.

(4) Mode.

(5) None of the above.

Question 17

The sum of deviations about the mean is always?

(1) Zero.

(2) Total standard deviation.

(3) Positive.

(4) Negative.

(5) Range.

Question 18

If the population variance is 4, then the population standard deviation must be?

(1) 4.

(2) 16.

(3) 2.

(4) 0.

(5) 12.

The data below shows an ordered array of the numbers of times fake news posts were shared on social media. Use this information to answer questions 19 to 24.

159	170	172	173	192	193	199	201
201	201	216	217	230	235	256	

Question 19

Which one of the following statements is incorrect?

(1) The number of observations is 15.

(2) The sum of all observations is 2759.

(3) The range is 97.

(4) The mean is 201.

(5) None of the above.

Question 20

Consider the statements below:

(A) The median is the 9th observation.

(B) There are three modes.

(C) The median is 201.

(D) The mode is 201.

Which statement(s) are incorrect?

(1) Only A.

(2) Only B.

(3) Only A and B.

(4) Only A, B and C.

(5) A, B, C and D.

Question 21

The distribution of the number of times the fake news posts are shared is _____?

(1) Positively skewed.

(2) Negatively skewed.

(3) Asymptotic.

(4) Asymmetric.

(5) Symmetric.

Question 22

What is the sample standard deviation?

(1) 25.8.

(2) 26.7.

(3) 374.2.

(4) 666.8.

(5) 714.4.

Question 23

Which one of the following statements is incorrect?

(1) The position of Q_1 is 4.

(2) The position of Q_2 is 8.

(3) The position of the median is 8.

(4) The position of Q_3 is 12.

(5) The value of Q_3 is 216.

Question 24

Which one of the following statements is incorrect?

(1) The range is 97.

(2) The interquartile range is 44.

(3) The coefficient of variation is 13.3%

(4) The range and interquartile range are measures of variation.

(5) The coefficient of variation is a measure of central tendency.

Question 25

Which one of the following statements is incorrect?

(1) If an event has a probability of 0.25, then the event has a 25% chance of occurring.

(2) A probability is a chance of an event not occurring.

(3) When two events are mutually exclusive, they cannot occur together.

(4) When two events are independent, they can occur together but they do not influence each other.

(5) The intersection of two events A and B is denoted as $P(A \text{ and } B)$.

Question 26

Consider the sample space $S = \{1, 2, 3, 4, 5\}$ and let A be a set of numbers less or equal to 4 and B be set of numbers greater than 4. Both A and B are subsets of S.
Which one of the following statements is incorrect?

(1) $A = \{1, 2, 3, 4\}$ and $B = \{5\}$.

(2) $P(A) = \frac{4}{5}$.

(3) $P(B) = \frac{1}{5}$.

(4) Events A and B are mutually exclusive.

(5) Events A and B are independent.

Table 1.1 below shows fake news medium and the type of personality the postings were about. Use this information to answer questions 27 to 30.

Table 1.1

	Celebrity (C)	Politician (P)	Sport Star (SS)	Total
Social Media (SM)	1800	700	515	3015
Traditional Media (TM)	485	350	150	985
Total	2285	1050	665	4000

Question 27

Which one of the following statements is incorrect?

(1) $P(SM) = 0.75$.

(2) $P\left(SM^C\right) = P(TM) = 0.25$.

(3) $P(C) = 0.57$.

(4) $P(SS) = 0.17$.

(5) None of the above.

Question 28

Which one of the following statements is incorrect?

(1) $P(C \text{ and } SS) = 0$.

(2) $P(C \text{ and } SM) = 0.45$.

(3) $P(C \text{ and } TM) = 0.12$.

(4) $P(SM|C) = 0.60$.

(5) $P(SM|SS) = 0.77$.

Question 29

Which one of the following statements is incorrect?

(1) Events Social Media and celebrity are dependent.

(2) Event Social Media and Sport Star are dependent.

(3) Events Traditional Media and Sport Star are independent.

(4) Events Social Media and Traditional Media are mutually exclusive.

(5) Events Celebrity and Politician are mutually exclusive.

Question 30

What is the probability that a fake news posting is about a Sport Star given that the source is traditional media?

(1) 0.25.

(2) 0.23.

(3) 0.17.

(4) 0.15.

(5) 0.04.