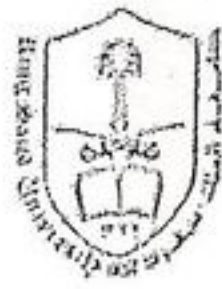


King Saud University
College of Sciences
Department of Statistic and
Operations Research



STAT 106: Biostatistics (Exercises)

Supervised by
Dr.Hanan Ali
Dr.Rae'esa

Prepared by
Amal Al-Mohisen
Kholoud Al-Sultan

Q#1) Answer the following questions:

- [1] The variable is a
- (a) subset of the population.
 - (b) parameter of the population.
 - (c) relative frequency.
 - (d) characteristic of the population to be measured.
 - (e) class interval.
- [2] Which of the following is an example of a discrete variable
- (a) the number of students taking statistics in this term at KSU.
 - (b) the time to exercise daily.
 - (c) whether or not someone has a disease.
 - (d) height of certain buildings.
 - (e) Level of education.
- [3] Which of the following is not an example of a discrete variable?
- (a) the number of students at the class of statistics.
 - (b) the number of times a child cry in a certain day.
 - (c) the time to run a certain distance.
 - (d) the number of buildings in a certain street.
 - (e) number of educated persons in a family.
- [4] Which of the following is an example of a qualitative variable?
- (a) the blood pressure level.
 - (b) the number of times a child brush his/her teeth.
 - (c) whether or not someone fail in an exam.
 - (d) weight of babies at birth.
 - (e) the time to run a certain distance.
- [5] Which of the following is an example of a parameter?
- (a) the mean age of every one living in Saudi Arabia
 - (b) the median of glucose level for all diabetic patients
 - (c) the variance of the birth weights of all babies born in Riyadh
 - (d) all of (a) to (c)
 - (e) None of (a) to (c)
- [6] Which of the following is an example of a statistic?
- (a) the mean age of all people in Saudi Arabia.
 - (b) the number of children in Saudi Arabia.
 - (c) the variance for the birth weights of all babies born in Riyadh.
 - (d) the median glucose level for a sample of diabetic patients.
 - (e) the type of disease a person has.
- [7] The continuous variable is a
- (a) variable with a specific number of values.
 - (b) variable which can't be measured.
 - (c) variable takes on values within intervals.
 - (d) variable with no mode.
 - (e) qualitative variable.
- [8] The width of the interval equals
- (a) the number of intervals.
 - (b) the distance from the lower limit of an interval to the lower limit of the next one.

- (c) the distance from the lower limit of an interval to the upper limit of it.
- (d) the difference between the frequency of an interval and the frequency of the next one.
- (e) the relative frequency.

[9] Which of the following is an example of a continuous variable?

- (a) The number of visitors of the clinic yesterday.
- (b) The time to finish the exam.
- (c) Whether or not the answer is true
- (d) The number of patients suffering from a certain disease
- (e) Level of education

[10] The y-axis in the graph of the histogram may represent the

- (a) frequency.
- (b) relative frequency.
- (c) either (a) or (b).
- (d) neither (a) nor (b).
- (e) we can't say.

[11] the discrete variable is a

- (a) Qualitative variable
- (b) Variable with no mode
- (c) Variable takes on values within intervals
- (d) Variable with a specific number of values
- (e) Variable which can not be measured

[12] Which of the following is an example of a nominal variable?

- (a) Age of visitors of the clinic
- (b) The time to finish the exam
- (c) Whether or not the person is infected by influenza
- (d) Weight of a sample of girls
- (e) The number of educated patients

[13] The nominal variable is a

- (a) variable with a specific number of values.
- (b) qualitative variable which can't be ordered.
- (c) variable takes on values within intervals.
- (d) variable with no mode.
- (e) quantitative variable.

[14] Which of the following is an example of an ordinal variable?

- (a) The number of persons who are injured in accidents
- (b) The time to finish the exam
- (c) Whether or not the medicine is effective
- (d) Blood type of a sample of patients
- (e) Socio-economic level

[15] The ordinal variable is a

- (a) variable with a specific number of values.
- (b) variable that takes on values within intervals.
- (c) qualitative variable which can be ordered.
- (d) variable with several modes.
- (e) quantitative variable.

[16] we calculate statistics instead of parameters because

- (a) Populations are too large to measure the parameters
- (b) Statistics are more popular than parameters

- (c) Parameters are very large but statistics are small
- (d) Statistics give more information than parameters
- (e) We are only interested in values of statistics

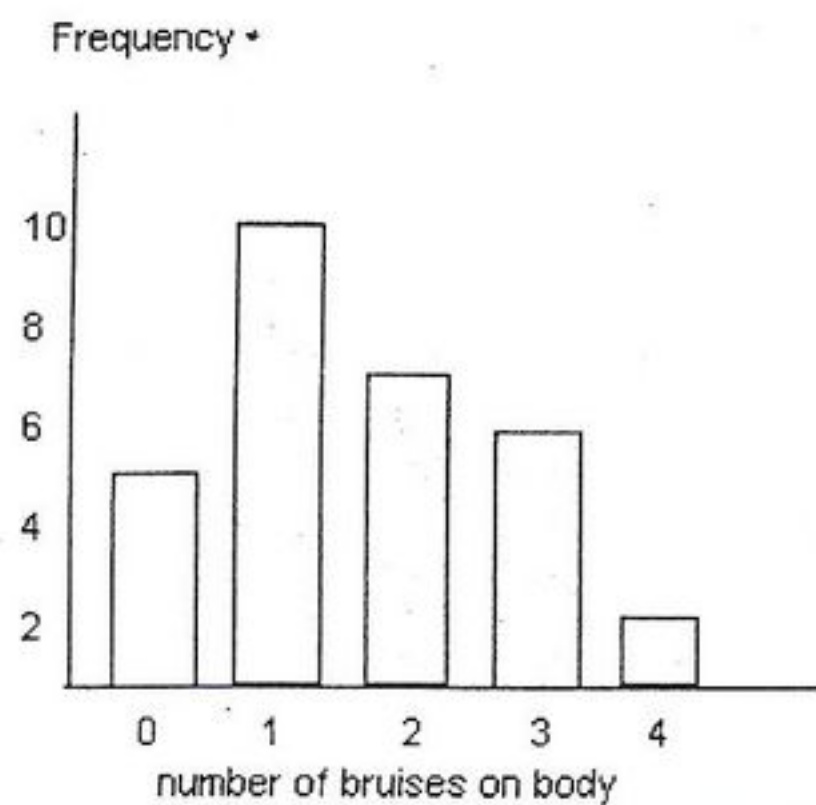
- [17] Which of the following is an example of a statistic?
- (a) the population variance.
 - (b) the sample median.
 - (c) the population mean.
 - (d) the population mode.
 - (e) none of these.

Q#2) For a sample of children with frequent toothache, we measure the number of times a child has been to a doctor in the last year as follows:

Number of times	1	2	3	4	5
Number of child	15	10	8	4	3

- A) What is the name of the variable?
- B) What is the type of the variable?
- C) What is the sample size?
- D) How many children have been to a doctor 3 times?
- E) How many children have been to a doctor 4 times or more?
- F) How many children have been to a doctor from 2 to 4 times?
- G) What is largest value of the variable?
- H) What is the percentage of the children has been to a doctor less than 3 times?
- I) What is the number of times a child has been to a doctor with highest percentage?

Q#3) For a sample of children age 5, we obtain the following graph:



- A) The type of the graph is:
- B) The name of the variable is:
- C) The type of the variable is:
- D) The sample size is:
- E) How many children had 3 or more bruises?
- F) What is the percentage of children has at most 2 bruises?
- G) What is the number of bruises has the largest percentage?
- H) What is the smallest value of the variable?

Q#4) For a sample of 60 women having children, we measure the age (in year). Complete the following table to give the values of the items mentioned:

Age	True classes	Midpoints	Frequency	Relative Frequency	Cumulative Frequency
15-20			3		
21-26				0.2	
			18		
			17		

- [1] 4th class (a)32-38 (b)32.5-38.5 (c)33.5-38.5
 (d)33-38 (e)32-37 (f)33-39
 [2] 5th true class: (a)39-<44 (b)39.5-<45.5 (c)38.5-<44.5
 (d)39.5-<44.5 (e)40-<45 (f)38.5-<45.5
 [3] 3rd midpoint: (a)29 (b)29.5 (c)28 (d)28.5 (e)27.5 (f)59
 [4] 2nd frequency: (a)12 (b)11 (c)10 (d)60 (e)38 (f)0
 [5] 3rd relative frequency: (a)0.2 (b)0.25 (c)0.18 (d)0.7 (e)0.5 (f)0.3
 [6] 2nd cum. frequency: (a)15 (b)12 (c)16 (d)13 (e)48 (f)60

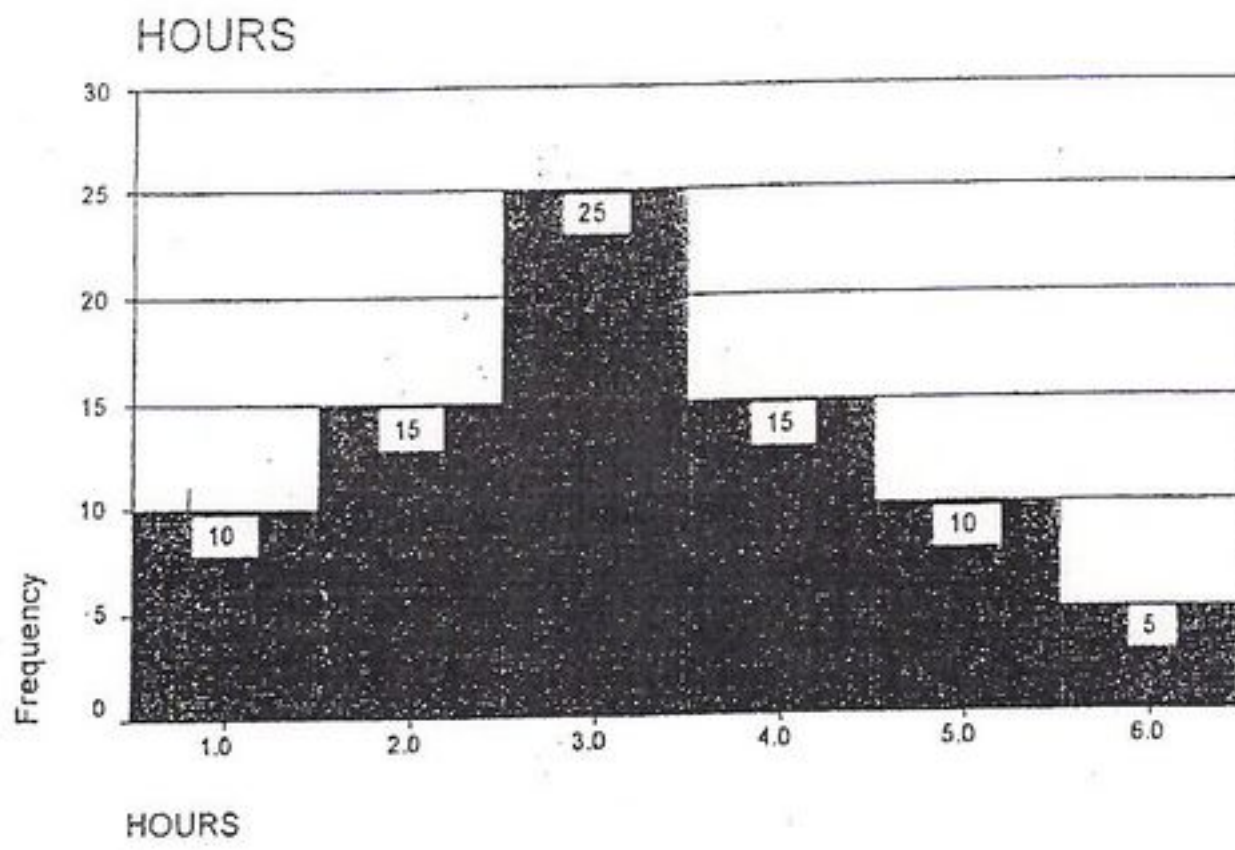
- [8] Which age class had the largest percentage of women?
 [9] How many women had of ages 21 or more?
 [10] The type of the variable is:
 [11] The percent of women had age 27 to less than 45:
 [12] The sample mean is
 [13] the sample standard deviation is

Q#5) For a sample of 200 mothers at delivery, we measure a certain serum level (in µg/l). Complete the following table to give the values of the items mentioned:

Time with doctor	True classes	Midpoints	Frequency	Relative Frequency	Cumulative Frequency
5-24		14.5	50		
25-44			63		
-				0.305	
-			24		
-					

- [1] 5th class (a)90-111 (b)84-104 (c)85-104
 (d)86-105 (e)87-107
 [2] 4th true class: (a)66-<86 (b)65-<85 (c)64.5-<84.5
 (d)66.5-<85.5 (e)64-<84
 [3] 3rd midpoint: (a)55 (b)56 (c)55.5 (d)55.75 (e)54.5
 [4] 3rd frequency: (a)2 (b)30 (c)63 (d)60 (e)61
 [5] 5th frequency: (a)12 (b)61 (c)2 (d)30 (e)100
 [6] 2nd relative frequency: (a)0.685 (b)0.63 (c)0.315 (d)0.25 (e)0.305
 [7] 4th cum. frequency: (a)200 (b)198 (c)174 (d)2 (e).99

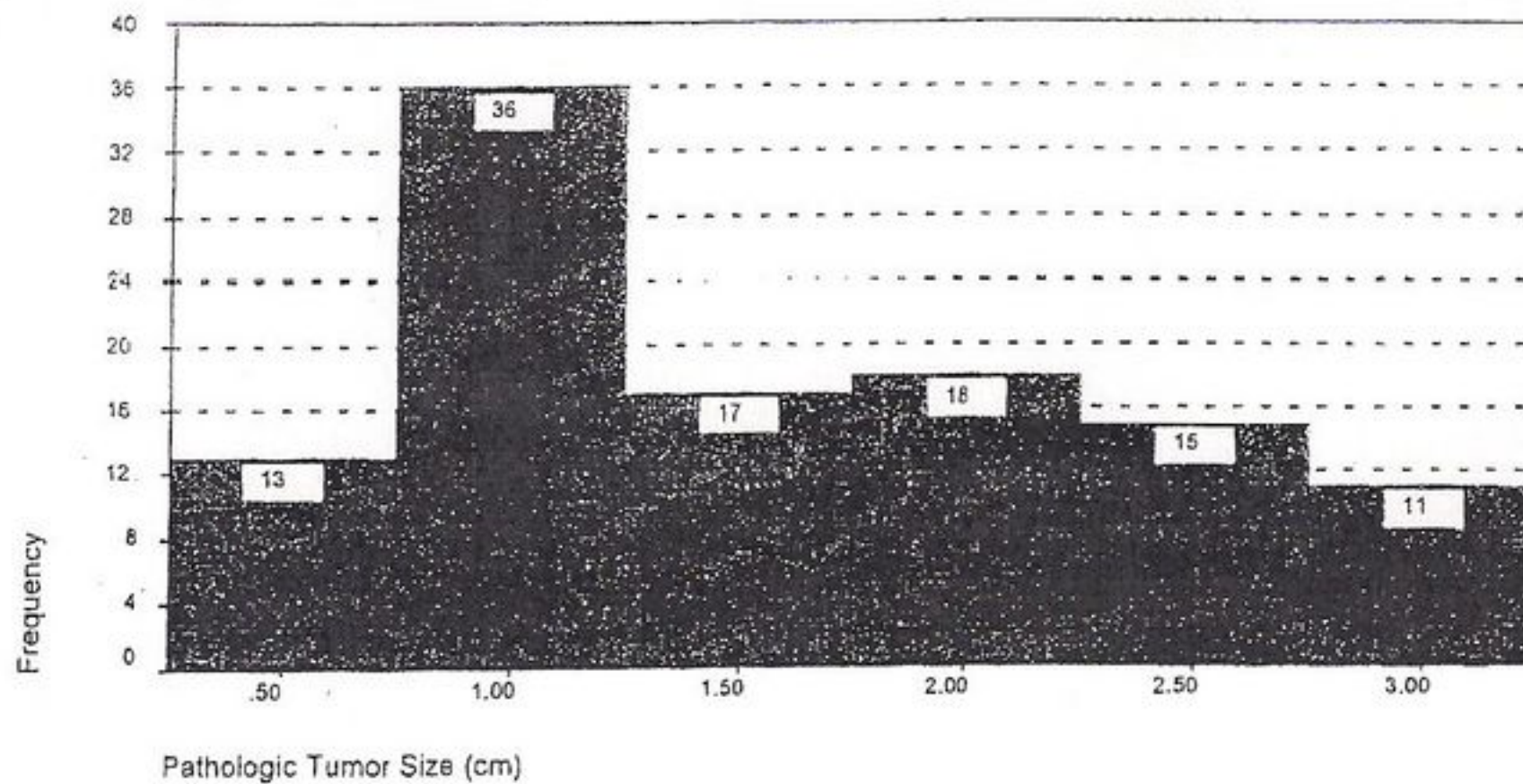
{...} For a sample of patients, we obtain the following graph for approximated hours spent without pain after a certain surgery (جراحة),



- [1] The type of the graph is:
 (a) bar chart (b) polygon (c) histogram (d) lines (e) curve (f) not known
- [2] The number of patients stayed the longest time without pain is:
 (a) 10 (b) 15 (c) 6 (d) 5 (e) 80 (f) 1
- [3] The percent of patients spent 3.5 hours or more without pain is:
 (a) 37.5% (b) 68.75% (c) 18.75% (d) 50% (e) 25% (f) 30%
- [4] The lowest number of hours spent without pain is:
 (a) 10 (b) 1 (c) 0.5 (d) 5 (e) 25 (f) 6.5
- [5] What is the approximate value of the sample mean?
 (a) 2.55 (b) 255 (c) 3 (d) 3.1875 (e) 40 (f) we can't find it
- [6] The mode equals
 (a) 80 (b) 3 (c) 15 (d) 2, 4 (e) 6 (f) we can't find it

$\bar{X} =$
 $S =$
 $S^2 =$

{**} The following histogram shows the frequency distribution of pathologic tumor size (in cm) for a sample of 110 Cancer patients:



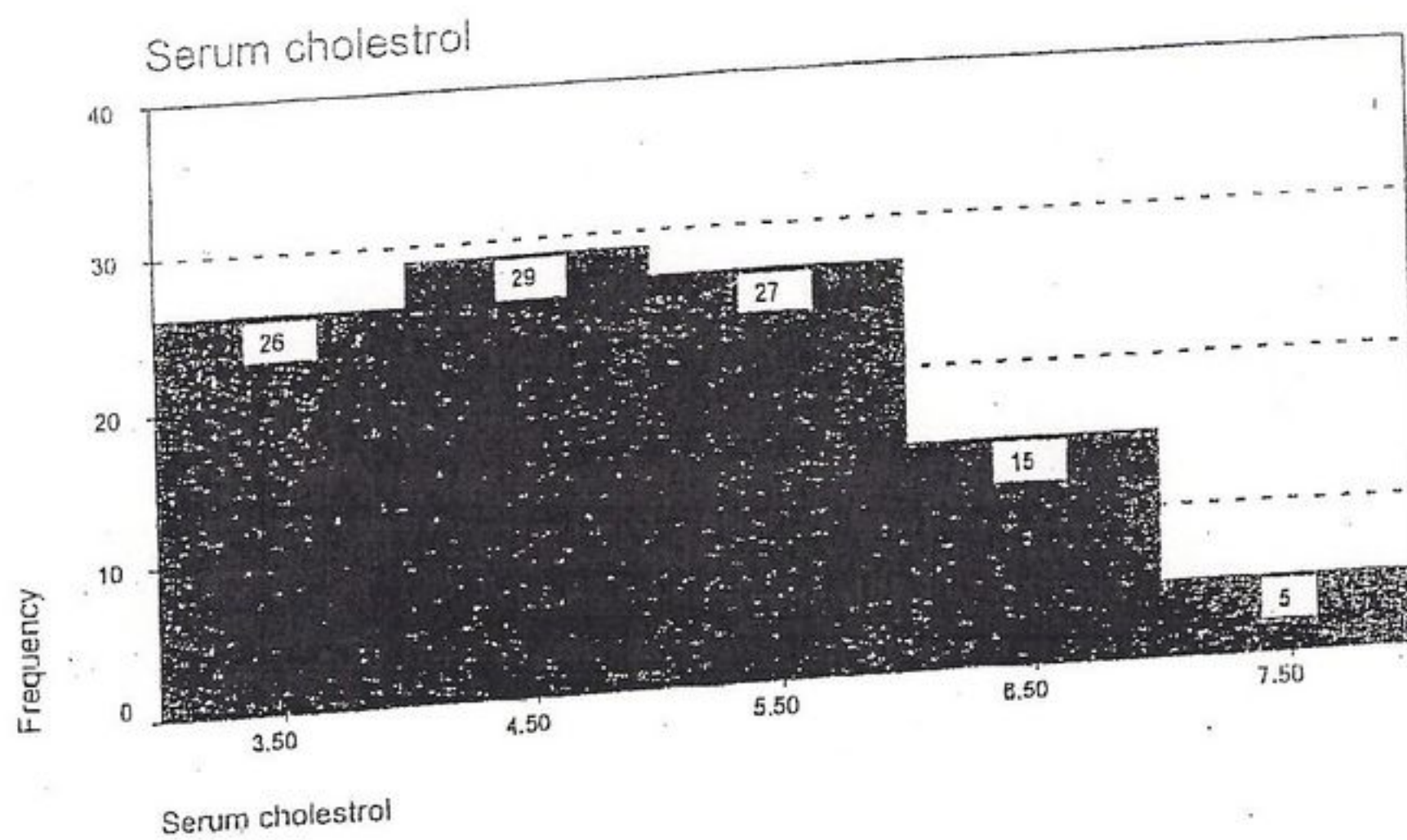
- [2.3] The percent of Cancer patients with approximate level of pathologic tumor size =2 cm is:
 (a) 18% (b) 50% (c) 16.36% (d) 32.72% (e) 36% (f) 0%
- [2.4] The number of Cancer patients with the lowest pathologic tumor size is:
 (a) 0.5 (b) 3 (c) 11 (d) 13 (e) 15 (f) 24
- [2.5] The approximate size of pathologic tumor with the highest percentage of patients is:
 (a) 3 (b) 1 (c) 36 (d) 110 (e) 32.72% (f) 16.36%
- [2.6] What is the approximate value of sample mean?
 (a) 18.33 (b) 36 (c) 1 (d) 1.75 (e) 1.523 (f) we can't find it
- [2.7] The mode takes a value near
 (a) 1 (b) 3 (c) 36 (d) 55 (e) 110 (f) we can't find it

$$\bar{x} =$$

$$s =$$

$$s^2 =$$

For a sample of Saudi women, we obtain the following graph for the serum cholesterol level (in mmol/l),



- [1] The type of the graph is:
 (a) curve (b) polygon (c) histogram (d) bar chart (e) lines (f) not known
- [2] The number of women having the highest serum cholesterol level is:
 (a) 4.5 (b) 29 (c) 5 (d) 26 (e) 102 (f) 82
- [3] The sample size =
 (a) 29 (b) 40 (c) 100 (d) 102 (e) 27.5 (f) 5
- [4] The approximate serum cholesterol level with the lowest percentage of women is:
 (a) 3.5 (b) 7.5 (c) 4.9% (d) 25.5% (e) 4.5 (f) 5
- [5] What is the approximate value of the sample mean?
 (a) 20.4 (b) 4.5 (c) 4.95 (d) 505 (e) 5.5 (f) we can't find it
- [6] The mode takes a value near
 (a) 102 (b) 29 (c) 40 (d) 4.5 (e) 7.5 (f) we can't find it

$\bar{X} =$

$S =$

$S^2 =$

Q#9) We have the score of student in stat106:

Freq	Stem	Leaf
1	2	7
1	3	5
2	4	2 9
5	5	0 1 3 7 9
6	6	2 3 4 7 8 9
5	7	0 4 5 8 8
3	8	4 5 8
2	9	3 2

- [1] What is the sample size?
- [2] How many student have less than 50?
- [3] How many student have 70 or more?
- [4] How many student have less than 72?
- [5] What is the percentage of students have 50 to less than 84?

Q#10) The following are the haemoglobin levels(g/100ml) of a sample of 10 children who are receiving treatment for a certain disease

6.7, 9.1, 10.0, 11.4, 12.4, 9.8, 8.3, 9.9, 9.1, 7.5

- (a) Compare the sample mean and variance
- (b) Suppose you have the following about haemoglobin levels for another sample of 15 healthy children

$$\sum_{i=1}^{15} X_i = 150, \quad \sum_{i=1}^{15} X_i^2 = 1521$$

Determine whether haemoglobin levels of this group are more variable than those of group (a)

Q#11) The weights (in kg) of 8 pregnant women gave the following results:

$$\sum_{i=1}^8 X_i = 495, \quad \sum_{i=1}^8 X_i^2 = 30659$$

- Find: (a) the mean, (b) the variance,
 (c) the standard deviation,
 (d) the coefficient of variation.

Q#12) The following table gives the results of a survey to study the weights of students of two schools.

	Mean Weights (kg)	Standard Deviation
School-A	30	6
School-B	60	10

Determine whether the weights of students from School A are more variable the weights of students from School B.

Q#13) If we measure the haemoglobin level and RBC for a sample of women:

	Mean	S.d	C.V
Hemoglobin	13.7	1.2	
RBC	4.5	0.5	

Which variable with less variation?

Q#14) The following table gives the results of a survey to study the ages and haemoglobin level of patients of a certain clinic.

	Mean	Standard Deviation
Age (Years)	30	6
Haemoglobin level (g/dl)	60	10

Determine whether haemoglobin levels of patients are more variable than ages.

Q#15) For a sample of size 15 women, here are the numbers of their children:

3 5 2 1 4 3 5 4 0 1 2 6 4 0 2

[1] The type of the variable is:

(a) qualitative (b) continuous (c) number (d) normal (e) discrete (f) statistic

[2] The sample mean is:

(a) 3 (b) 6 (c) 7.84 (d) 2.8 (e) 2.5 (f) 0

[3] The sample median is:

(a) 3 (b) 2 (c) 3.5 (d) 7.5 (e) 2.5 (f) 1.5

[4] The sample mode is:

[5] The sample (a) 6 (b) 2 and 4 (c) 2, 3 and 4 (d) 3 (e) 5 (f) no mode

standard deviation is:

(a) 166 (b) 3.457 (c) 1.364 (d) 1.796 (e) 1.859 (f) 1

[7] The coefficient of variation is:

(a) 18.59% (b) 15% (c) 123.42% (d) 150.62% (e) 66.39% (f) 100%

*If we add 2 to each number of children, then:

[8] The sample mean is:

(a) 2 (b) the same (c) decreased by 2 (d) increased by 2 (e) divided by 2
(f) multiplied by 2

[9] The sample variance is:

(a) the same (b) 4 (c) multiplied by 4
(d) increased by 4 (e) divided by 4 (f) multiplied by 2

[10] The coefficient of variation become:

(a) the same (b) 50% (c) smaller (d) larger (e) 0% (f) non of these

Q#16) We measure the number of asthma cases seen in the past months for a sample of hospital:

20 16 30 14 20 35 6 29 20 25 49 15

[1] The type of the variable is:

(a) qualitative (b) continuous (c) number (d) normal (e) discrete (f) statistic

[2] The sample mean is:

(a) 20 (b) 30 (c) 24 (d) 23.25 (e) 20.5 (f) 2.5

[3] The sample mode is:

(a) 16 (b) 49 (c) 35 (d) 25 (e) 20 (f) no mode

[4] The sample median is:

(a) 16 (b) 25 (c) 20 (d) 35 (e) 30 (f) 49

[5] The sample standard deviation is:

(a) 10 (b) 23.25 (c) 128.93 (d) 7905 (e) 279 (f) 11.355

[7] The coefficient of variation is:

(a) 204.76% (b) 48.84% (c) 554.53% (d) 100% (e) 84.82% (f) non of these

*If we multiply each number of asthma of stairs by 2, then:

[8] The sample mean is: (a) 2 (b) the same (c) the mode (d) increased by 2
(e) divided by 2 (f) multiplied by 2

[9] The sample standard deviation is:

- (a) the same (b) 2 (c) multiplied by 4
(d) increased by 2 (e) divided by 2 (f) multiplied by 2

[10] The coefficient of variation become:

- (a) the same (b) more than 100% (c) smaller (d) larger (e) 0% (f) non of these

Q#19) For any two events, A and B:

[1] $P(A \cap B) =$

- (a) $P(A)P(B)$ (b) $P(B) - P(B \cap A^c)$ (c) $P(A) + P(B)$
(d) $P(A) - P(B \cap A^c)$ (e) none of these

[2] $P(A \cup B) =$

- (a) 1 (b) $1 - P(A^c \cap B^c)$ (c) $P(A) + P(B)$
(d) $P(A)P(B)$ (e) none of these

[3] If A and B are independent, then $P(B|A) =$

- (a) $P(B)$ (b) $P(B \cap A)$ (c) $P(A)$
(d) $P(B \cap A) / P(B)$ (e) none of these

[4] $P(B) = P(A \cap B) +$

- (a) $P(B^c)$ (b) $P(A \cup B)$ (c) $P(A^c \cap B)$ (d) $P(A - B)$ (e) none of these

[5] $P(A \cap B) = 0$, then we know that A and B are

- (a) empty (b) disjoint (c) independent (d) events (e) none of these

Q#20) if $\Omega = \{A, B, C, D\}$, and the outcomes are equally likely, then

[1] $P(A) =$ (a) 1 (b) 1/2 (c) 1/4 (d) 1/3 (e) none of these

[2] $2P(D) =$ (a) $P(A) - P(B)$ (b) $P(A) / 2$ (c) $P(C)$ (d) $P(A) + P(B)$ (e) none of these

Q#21) A and B are events defined on the same sample space.

If $P(\bar{A}) = 0.6, P(B) = 0.5$ and $P(A \cap B) = 0.1$,

Find: (i) $P(A \cup B)$

(ii) $P(A \cap \bar{B})$

(iii) $P(\bar{A} \cap B)$

(iv) $P(\bar{A} \cap \bar{B})$

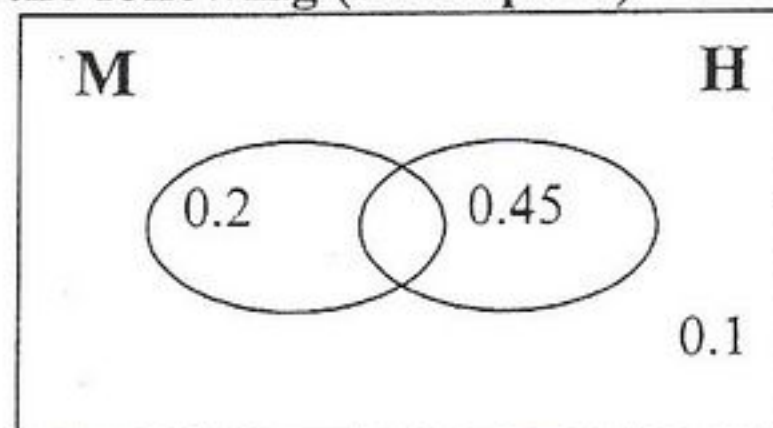
Q#22) If $P(A \cap \bar{B}) = 0.3, P(A \cap B) = 0.2$ and $P(\bar{A} \cap \bar{B}) = 0.1$,

Find: (i) $P(A)$

(ii) $P(\bar{A} \cap B)$

(iii) $P(A \cup B)$

Q#23) In a population of adult patients with a certain disease, let M = "is a man" and H = "has a heart disease". We have the following (incomplete) Venn diagram:



If we randomly choose one patient, find the probabilities that the patient chosen:

[1] is a man and had a heart disease:

- (a) 0.75 (b) 0.20 (c) 0.25 (d) 0.45 (e) 0.10

[2] is a women :

- (a) 0.55 (b) 0.70 (c) 0.45 (d) 0.10 (e) 0.80

[3] does not have a heart disease : (a) 0.55 (b) 0.10 (c) 0.65 (d) 0.30 (e) 0.70

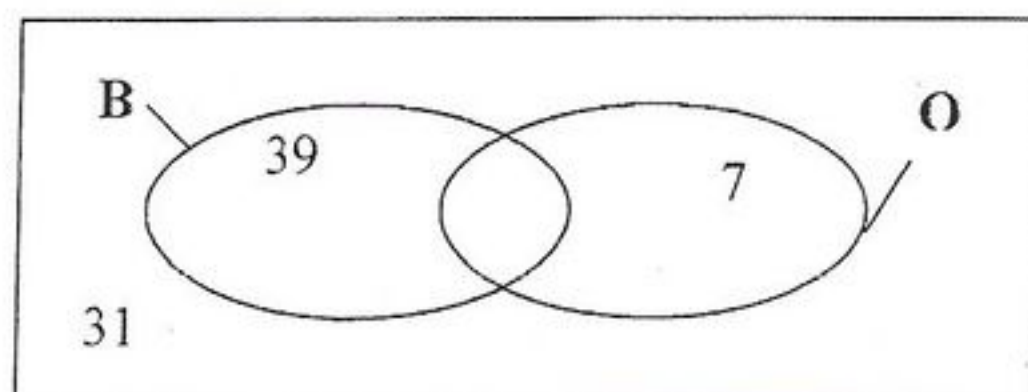
[4] is a man or has a heart disease : (a) 0.45 (b) 0.70 (c) 0.55 (d) 0.25 (e) 0.90

[5] she has heart disease : (a) 0.45 (b) 0.55 (c) 0.6429 (d) 0.8182 (e) 0.70

[6] if we know that the patient has heart disease , what effect does this have on the probability that he is a man ? (a) decreases (b) increases (c) has no effect

- (d) it is the probability of a man (e) person has heart disease

Q#24) From a population of 80 babies in a certain hospital the last month, let B = "is a boy" and O = "is over weight". We have the following incomplete Venn diagram:



If we randomly choose a baby born in this hospital, find the probabilities that the baby chosen:

[1] is a boy

- (a) 0.4875 (b) 0.525 (c) 0.6125 (d) 0.9286 (e) 0.5

[2] is not an over weight baby

- (a) 0.875 (b) 0.9125 (c) 0.0875 (d) 0.4875 (e) 0.3875

[3] is a girl or an over weight baby

- (a) 0.475 (b) 0.125 (c) 0.875 (d) 0.5125 (e) 0.6125

[4] is an over weight baby knowing that he is a boy

- (a) 0.0714 (b) 0.525 (c) 0.125 (d) 0.0375 (e) 0.3

Q#25) The following table classifies 400 people according to their smoking habits and whether or not they have cancer

	Smoker (A)	Non-Smoker(\bar{A})
Has cancer (C)	200	50
Dose not have cancer(\bar{C})	50	100

If an individual is selected at random from this group, find the probability that he/she is

(a) a smoker and has cancer,

(b) a smoker or has cancer

(c) a non-smoker or has cancer

(d) If an individual is selected at random from the group, find the probability that the person selected:

(i) has cancer given that he/she is a smoker.

(ii) is not a smoker given that he/she dose not have cancer

(e) determine whether smoking and having cancer are independent

Q#26) A group of people age 40 or less with a certain injury is classified by the place where the injury occurred and the age of the person:

Place where injury occurred (Age)	Home (H)	Work (W)	School (S)	Total
17 or less (A)	90	15	80	185
18 to 26 (B)	30	100	35	165
27 to 40 (C)	5	140	5	150
Total	125	255	120	500

If one of these people is randomly chosen give:

(1) In symbols, the event "injured at work or at school and the age is between 18 to 26"

- (a) $(W \cap S) \cup B$ (b) $(W \cup S) \cap B$ (c) $(W \cup S) \cup B$
 (d) $(W \cap S) \cap B$ (e) none of these

- (2) $P(B \cup W) =$ (a) 0.33 (b) 0.51 (c) 0.64 (d) 0.84 (e) 0.18
 (3) $n(A \cap W) =$ (a) 185 (b) 15 (c) 225 (d) 0.03 (e) 0.25
 (4) $P(B^c) =$ (a) 0.67 (b) 0.33 (c) 0.37 (d) 0.3 (e) 0.07
 (5) $P(S^c \cup C^c) =$ (a) 0.53 (b) 0.76 (c) 0.7 (d) 0.47 (e) 0.99
 (6) $P(S | B) =$ (a) 0.2917 (b) 0.2121 (c) 0.07 (d) 0.57 (e) 0.5
 (7) $n(A \cup C) =$ (a) 0 (b) 335 (c) 185 (d) 150 (e) 500

Q#27) A population of people is classified by the calcium intake and whether the person is a man, woman, or child:

	Calcium Intake			Total
	Below needed (B)	Enough (E)	Above needed (A)	
Man (M)	72	200	48	320
Woman (W)	104	160	16	280
Child (C)	64	106	30	200
Total	240	466	94	800

If one person is randomly chosen from the population,

[1] Give the event $B^c \cap C$ in words.

[2] Give the event "above needed calcium given a woman" in symbols.

- [3] $P(B) =$ (a) 0.09 (b) 0.15 (c) 0.3 (d) 0.7 (e) 240
 [4] $P(W^c) =$ (a) 0.13 (b) 0.35 (c) 0.4 (d) 0.65 (e) 0.87
 [5] $P(W \cup B) =$ (a) 0 (b) 0.13 (c) 0.52 (d) 0.65 (e) 0.73
 [6] $P(M \cap C) =$ (a) 0 (b) 0.13 (c) 0.35 (d) 0.65 (e) 0.69
 [7] $P(W \cap B^c) =$ (a) 0.18 (b) 0.22 (c) 0.54 (d) 0.83 (e) 1.05
 [8] $P(A \cup E) =$ (a) 0 (b) 0.31 (c) 0.39 (d) 0.7 (e) 0.83
 [9] $P(B | C) =$ (a) 0.08 (b) 0.267 (c) 0.32 (d) 0.47 (e) 0.55

Q#28) If X is a discrete random variable with probability distribution:

X	0	1	2	3	4	5
P(X=x)	0.05	0.15	0.15	0.25	0.3	0.1

- 1) What value of X has the highest probability? (a) 1 (b) 2 (c) 3 (d) 4 (e) 5
 2) $P(X < 3) =$ (a) 0.25 (b) 0.35 (c) 0.4 (d) 0.6 (e) 0.65
 3) $P(1 \leq X < 4) =$ (a) 0.1 (b) 0.15 (c) 0.3 (d) 0.55 (e) 0.85
 4) $\mu =$ (a) 2.5 (b) 2.9 (c) 3.0 (d) 3.5 (e) 11.3

Q#29) For a population of families, x = the number of children in primary school. We randomly choose one and the cumulative distributed is given below:

X	P(X ≤ x)
0	0.12
1	0.36
2	0.72
3	0.95
5	1

- (1) $P(X = 2) =$ (a) 0.24 (b) 0.36 (c) 0.48 (d) 0.64 (e) 0.72
 (2) $P(X = 3) =$ (a) 0.23 (b) 0.36 (c) 0.47 (d) 0.59 (e) 0.95
 (3) $P(1 < X \leq 4) =$ (a) 0 (b) 0.36 (c) 1.08 (d) 0.59 (e) 0.72
 (4) The expected number of children in primary school for a family from this population:
 (5) (a) 1.9 (b) 1.09 (c) 0.19 (d) 9.65 (e) we can't find it

Q#30) For a probability of children, X = the number of sweets eaten on a certain day. We randomly chose a child and the number of sweets he ate was from 0 to 4 sweets.

The following probabilities are given:

$$P(X=0) = 0.08, P(X=2) = 0.31, P(X > 2) = 0.42, P(1 < X \leq 3) = 0.63.$$

Then,

- [1] $P(X = 1) =$ (a) 0.09 (b) 0.23 (c) 0.42 (d) 0.19 (e) 0
 [2] $P(X = 3) =$ (a) 0.68 (b) 0.22 (c) 0.32 (d) 0.51 (e) 0
 [3] $P(X \leq 2) =$ (a) 0.92 (b) 0.58 (c) 0.99 (d) 0.27 (e) 1
 [4] The expected number of sweets eaten daily by a child from this population is:
 (a) 1 (b) 2 (c) 0.217 (d) 4.7089 (e) 2.17

Q#31) In a large population of people, 25% of them use dental floss regularly. If we randomly choose 15 of these people and let X = the number of these 15 people who use dental floss regularly, then

[1] The probability distribution of X, is $P(X = x) =$

(a) $\binom{15}{x} (0.75)^x (0.25)^{15-x}$ (b) $\binom{15}{x} (0.25)^x (0.75)^{15-x}$ (c) $\binom{25}{x} (0.15)^x (0.85)^{25-x}$
 (d) $e^{-15} (15)^x / x!$ (e) $e^{-0.25} (0.25)^x / x!$

[2] The values that x takes are:

(a) 1, 2, ..., 15 (b) 1, 2, ..., 25 (c) 0, 1, ..., 15
 (d) 0, 1, ..., 25 (e) 0, 1, ..., ∞

[3] $P(X = 4) =$

(a) 0.0634 (b) 0.2252 (c) 0.0202
 (d) 0.1651 (e) none of these

[4] The probability that at most one person uses dental floss regularly?

(a) 0.9198 (b) 0.0134 (c) 0.0668
 (d) 0.0802 (e) none of these

[5] The expected value of X =

(a) 3.75

(b) 11.25

(c) 15

(d) 25

(e) none of these

Q#32) In a large population of young Arab children, 11 percent of them have trouble saying some letters. If we randomly choose 8 children and let X = the number of these 8 children who have trouble saying some letters, then

[1] The probability distribution of X , is $P(X = x) =$

- (a) $\binom{8}{x} (0.11)^x (0.89)^{8-x}$ (b) $\binom{11}{x} (0.08)^x (0.92)^{11-x}$ (c) $\binom{8}{x} (0.11)^x (0.89)^{8-x}$
 (d) $e^{-11} (11)^x / x!$ (e) $e^{-8} (8)^x / x!$

[2] The values that x takes are:

- (a) 1,2,...,8 (b) 1,2,...,11 (c) 0,1,...,89
 (d) 0,1,...,8 (e) 0,1,..., ∞

[3] $P(X = 4) =$

- (a) 0.057 (b) 0.0064 (c) 0.0303
 (d) 0.0075 (e) none of these

[4] The probability that at most there are one children who have trouble saying some letters =

- (a) 0.7829 (b) 0.2171 (c) 0.3892
 (d) 0.3937 (e) none of these

[5] The expected value of $X =$

- (a) 11 (b) 8 (c) 0.88 (d) 7.12 (e) none of these

Q#33) At a certain hospital, X = the number of kidney transplants in a year has a Poisson(5.4) distribution.

[1] The formula for $P(X = 4)$ is:

- (a) $\binom{4}{x} (0.54)^x (0.46)^{4-x}$ (b) $e^{-4} (5.4)^4 / 4!$ (c) $\binom{4}{x} (0.46)^x (0.54)^{4-x}$
 (d) $e^{-4} (4)^{5.4} / 4!$ (e) $e^{-5.4} (5.4)^4 / 4!$

[2] $P(X \geq 2) =$

- (a) 0.0045 (b) 0.9711 (c) 0.54 (d) 0.0289 (e) none of these

[3] What is the distribution of Y = the number of kidney transplants in 3 years.

- (a) Poisson(1.87) (b) Poisson(3) (c) Poisson(16.2)
 (d) Poisson(8.4) (e) Binomial(3,0.54)

Q#34) 'in a certain population, it is known that the average number of deaths per year from cancer is 3 if the number of deaths from this disease has the Poisson distribution, find the probability that, from this disease, there are:

- (a) not more than 2 deaths in a year,
 (b) at least two deaths in a year,
 (c) exactly 3 deaths in a year,
 (d) exactly 4 deaths in 4 months.

Q#35) At a certain hospital in Riyadh, X = the number of food poisoning cases in a month has a Poisson(5.6) distribution.

[1] The possible values of X are:

- (a) $x = 0,1,2$ (b) $x = 0,1,2,\dots,6$ (c) $x = 1,2,\dots,30$
 (d) $x = 0,1,2,3,4,\dots$ (e) $x = 1,2$

[2] $P(X = 7) =$

- (a) 0 (b) 0.0312 (c) 0.1267 (d) 0.4109 (e) none of these

[3] What is the expected number of food poisoning cases in 3 months?

- (a) 0.56 (b) 1.87 (c) 5.6
 (d) 7.47 (e) 16.8

[4] What is the distribution of Y = the number of food poisoning cases in 1/2 a month?

- (a) Poisson(0.373) (b) Poisson(2.8) (c) Poisson(5.6)
(d) Poisson(11.2) (e) Poisson(84)

Q#36) In a certain population, an average of 6 new cases of breast cancer are diagnosed each year. If the number of new diagnosed cases of this disease in the population follows the Poisson distribution, find the probability that, in this population, there are:

- (a) no new cases of the disease in a year,
(b) at most two new cases of the disease in 6 months.

Q#39) If Z a standard normal distribution, then

- [1] $P(z > 1.34) =$ (a) 0.0901 (b) 0.34 (c) 0.3669 (d) 0.6331
[2] $P(z \leq -0.78) =$ (a) 0.0375 (b) 0.2177 (c) 0.4681 (d) 0.7823
[3] $P(-1.89 \leq z < 2.56) =$ (a) 0.7486 (b) 0.9112 (c) 0.9654 (d) 0.9929

Q#40) If Z is $N(0,1)$, find

- (a) $P(Z \leq 1.36)$ (b) $P(Z \geq 2.4)$ (c) $P(Z < 1.81)$
(d) $P(Z > 2.7)$ (e) $P(-1.2 < Z < 2.1)$
(f) $P(Z = 1.4)$ (g) $P(-2.36 < Z < 1.45)$

Q#41) Let Z have the standard normal distribution. The value of z such that $P(Z \leq z) = \alpha$, is defined by z_α . Find the following z values:

- (a) $z_{0.5}$ (b) $z_{0.95}$ (c) $z_{0.025}$ (d) $z_{0.975}$

Q#42) In population of people, X = the body mass index (in kg/m^2) is normally distributed with mean $\mu = 25$ and standard deviation $\sigma = 2$. For a randomly chosen person,

- [1] $P(X < 21) =$
(a) 0.9772 (b) 0.4772 (c) 0.9821 (d) 0.0228 (e) none of these
[2] $P(19 < X < 28) =$
(a) 0.9332 (b) 0.9319 (c) 0.9345 (d) 0.5332 (e) none of these
[3] The value of x such $P(X > x) = 0.2578$ is:
(a) 0.65 (b) 1.3 (c) 26.3 (d) 23.7 (e) none of these

Q#43) In a population of heat stroke victims, X = the cooling time (in min.) is normally distributed with $\mu = 40$ and $\sigma^2 = 100$. For a randomly chosen victim,

- [1] $P(X \leq 25) =$
(a) 0.0062 (b) 0.0668 (c) 0.4404 (d) 0.5596 (e) 0.9332
[2] $P(20 \leq X \leq 70) =$
(a) 0.1587 (b) 0.8413 (c) 0.9544 (d) 0.9759 (e) 1.0215
[3] The value of x such $P(X > x) = 0.791$ is:
(a) 0.81 (b) 18.4 (c) 31.9 (d) 47.9 (e) 48.1

Q#44) if T has a student's t-distribution, then choose the correct answer for:

1- $P(T_{24} = 2.0639) =$

- (a) 0 (b) 0.025 (c) 0.975 (d) 0.95

2- $P(T_{18} > t) = 0.995$, then $t =$

- (a) 0.005 (b) -2.552 (c) -2.8784 (d) -1.7341

3- $P(T_9 < 1.8331) =$

- (a) 0.005 (b) 0.995 (c) 0.975 (d) 0.05 (e) 0.9

4- $P(T_{(17)} < 1.7396) =$

- (a) 0.005 (b) 0.01 (c) 0.95 (d) 0.995 (e) 1

Q#45) Answer 1 to 3 directly from the question. Answer 4 to 6 using new information as given

Statement: In a sample of 49 Saudis living in villas, the mean Vitamin D level was 16 with a variance of 12. Assume that the distribution is normal. **Question:** Find and interpret a 90% confidence interval for the average .

[1] The number 12 is:

- (a) S^2 (b) S (c) σ^2 (d) σ (e) μ_0

[2] the value of α is:

- (a) 0.01 (b) 0.10 (c) 0.45 (d) 0.95 (e) unknown

[3] The assumptions are:

- (a) normal, σ^2 known (b) not normal, σ^2 known, n large
(c) normal, σ^2 unknown (d) not normal, σ^2 unknown, n large
(e) none of these

[4] Suppose that the assumption are "normal, σ^2 known" then the correct formula to answer the question is:

- (a) $Z = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}}$ (b) $\bar{X} \pm z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$ (c) $T = \frac{\bar{X} - \mu_0}{s / \sqrt{n}}$ (d) $\bar{X} \pm t_{1-\frac{\alpha}{2}, n-1} \frac{s}{\sqrt{n}}$ (e) $\bar{X} \pm z_{1-\frac{\alpha}{2}} \frac{s}{\sqrt{n}}$

[5] using $\bar{X} \pm z_{1-\frac{\alpha}{2}} \frac{s}{\sqrt{n}}$, calculate the confidence interval with $\alpha = 0.05, s = 3.8$ and $n = 49$

- (a) (14.9066, 17.0934) (b) (14.9096, 17.0904) (c) (14.936, 17.064)
(d) (15.107, 16.893) (e) (15.3041, 16.6959)

[6] If a 90% confidence interval for the average gives (15.2, 16.9), what part is missing in the indicated place in the interpretation below:

Interpretation: we are 90% sure that the (a) the (b) Vitamin D level (c) of (d) Saudis living in villas is between 15.2 and 16.9.

- (a) If the mean was 16 (b) average (c) with a variance of 12
(d) 49 (e) all of these are missing

Q#46) Suppose we take 14 samples of a certain food we measure the vitamin A content obtaining the mean 128.214 and standard deviation 13.874

Assume a normal distribution.

[1] Test if the mean is different from 120 $\alpha = 0.05$

[2] Test if the mean is more than 120 $\alpha = 0.05$

Q#47) Answer 1 to 3 directly from the question. Answer 4 to 9 using new information as given.

Statement: In a sample of 60 lung cancer patients, the mean lifetime after diagnosis was 2.78. Assume a distribution that is not normal with a standard deviation of 1.5. use $\alpha = 0.05$. Test if the average lifetime is more than 2.5 year.

[1] The number 1.5 is:

- (a) S^2 (b) S (c) σ^2 (d) σ (e) μ_0

[2] The assumptions are:

- (a) normal, σ^2 known (b) not normal, σ^2 known, n large
(c) normal, σ^2 unknown (d) not normal, σ^2 unknown, n large
(e) none of these

[3] The hypotheses are:

- (a) $H_0: \mu_0 = 2.5$ (b) $H_0: \mu > 2.5$ (c) $H_0: \mu_0 = 2.5$ (d) $H_0: \mu_0 = 2.5$ (e) $H_0: \mu = 2.5$
 $H_a: \mu < 2.5$ $H_a: \mu = 2.5$ $H_a: \mu_0 > 2.5$ $H_a: \mu_a > 2.5$ $H_a: \mu > 2.5$

[4] Suppose that the assumption are "normal, σ^2 unknown", then the correct test statistic is:

- (a) $Z = \frac{\bar{X} - \mu_0}{s/\sqrt{n}}$ (b) $Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$ (c) $T = \frac{\bar{X} - \mu_0}{s/\sqrt{n}}$ (d) $\bar{X} \pm t_{1-\frac{\alpha}{2}, n-1} \frac{s}{\sqrt{n}}$ (e) none of these

[5] The distribution of the test statistic in [4] is:

- (a) $t_{1-\alpha, 59}$ (b) t_{60} (c) t_{59} (d) $z_{1-\alpha}$ (e) standard normal

[6] Suppose that $H_a: \mu \neq \mu_0$ and we are using a Z statistic and $\alpha = 0.05$ (n as in the statement), then the correct rejection rule is:

- (a) Reject H_0 if $Z > 1.645$ (b) Reject H_0 if $Z < -1.645$ or $Z > 1.645$
(c) Reject H_0 if $Z < -1.96$ (d) Reject H_0 if $Z < -1.96$ or $Z > 1.96$
(e) Reject H_0 if $Z < -2.0086$ or $Z > 2.0086$

[7] Use the statement along with $s = 2$ to calculate the value of $T = \frac{\bar{X} - \mu_0}{s/\sqrt{n}}$. The answer is:

- (a) -6.902 (b) 0.0723 (c) 1.084 (d) 2.135 (e) 4.957

[8] If the rule is "Reject H_0 if $Z > 1.645$ " and $Z = 3.642$, then the decision is:

- (a) Reject H_0 (b) Fail to reject H_0 (c) either one of them (d) we can not say (e) $\mu = 2.5$

[9] If we were testing $H_a: \mu \neq 2.5$, and our decision was "Fail to Reject H_0 " then **choose the underlined part** in the conclusion that is **incorrect**:

Conclusion: We cannot conclude that the average lifetime after diagnosis for 60 lung cancer patients is different from 2.5.

- (a) cannot (b) average (c) 60 (d) different from (e) none of these (all are correct)

Q#48) Statement: In a sample of 38 air samples in a certain city, the mean level of chemical A (in ppm) was 97.5. Assume a normal distribution with a variance of 10. Use $\alpha = 0.05$. Test if the average level is less than the recommended health limit of 100 ppm.

[1] The number 10 is:

- (a) S^2 (b) S (c) σ^2 (d) σ (e) μ_0

[2] The assumptions are:

- (a) normal, σ^2 known (b) not normal, σ^2 known, n large
 (c) normal, σ^2 unknown (d) not normal, σ^2 unknown, n large
 (e) none of these

[3] The hypotheses are:

- (a) $H_0: \mu < 100$ (b) $H_0: \mu = 100$ (c) $H_0: \mu = 100$ (d) $H_0: \mu = 100$ (e) $H_0: \mu = 100$
 $H_a: \mu = 100$ $H_a: \mu > 100$ $H_a: \mu < 100$ $H_a: \mu \neq 100$ $H_a: \mu < 100$

[4] Suppose that the assumption are "normal, σ^2 unknown" then the correct test statistic is:

- (a) $Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$ (b) $Z = \frac{\bar{X} - \mu_0}{s/\sqrt{n}}$ (c) $T = \frac{\bar{X} - \mu_0}{s/\sqrt{n}}$ (d) $\bar{X} \pm t_{1-\frac{\alpha}{2}, n-1} \frac{s}{\sqrt{n}}$ (e) none of these

[5] The distribution of the test statistic in [4] is:

- (a) $t_{1-\alpha, 37}$ (b) t_{37} (c) t_{38} (d) $z_{1-\alpha}$ (e) standard normal

[6] Suppose that $H_a: \mu \neq \mu_0$ and we are using a Z statistic and $\alpha = 0.1$ (n as in the statement), then the correct rejection rule is:

- (a) Reject H_0 if $Z < 1.282$ (b) Reject H_0 if $Z < -1.645$ or $Z > 1.645$
 (c) Reject H_0 if $Z < -1.282$ or $Z > 1.282$ (d) Reject H_0 if $Z < -1.6896$ or $Z > 1.6896$
 (e) Reject H_0 if $Z : z_{0.9}$

[7] Use the statement along with $s = 3$ to calculate the value of $Z = \frac{\bar{X} - \mu_0}{s/\sqrt{n}}$. The answer is:

- (a) -5.137 (b) -1.726 (c) -0.658 (d) 0.95 (e) 9.5

[8] If the rule is "Reject H_0 if $T < -1.689$ " and $T = -3.26$, then the decision is:

- (a) Reject H_0 (b) Fail to reject H_0 (c) either one of them (d) we can not say (e) $\mu = 100$

[9] If we were testing $H_a: \mu \neq 105$, and our decision was "Reject H_0 " then **choose the underlined part in the conclusion that is incorrect:**

Conclusion: We can conclude that the average level of chemical A for 38 air samples in the city is different from 105.

- (a) can (b) average (c) 38 (d) different from (e) none of these (all are correct)

Q#49) Answer 1 to 3 directly from the question. Answer 4 to 6 using new information as given

Statement: In a sample of 13 male smokers, the mean basal acid output was 2.23 with standard deviation of 1.35. Assume a normal distribution. Question: Find and interpret a 90% confidence interval for the average acid output.

[1] The number 1.35 is:

- (a) S^2 (b) S (c) σ^2 (d) σ (e) μ_0

[2] the value of α is:

- (a) 0.05 (b) 0.10 (c) 0.45 (d) 0.95 (e) unknown

[3] The assumptions are:

- (a) normal, σ^2 known (b) not normal, σ^2 known, n large
(c) normal, σ^2 unknown (d) not normal, σ^2 unknown, n large
(e) none of these

[4] Suppose that the assumption are "not normal, σ^2 unknown, n large" then the correct formula to answer the question is:

(a) $Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$ (b) $\bar{X} \pm z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$ (c) $T = \frac{\bar{X} - \mu_0}{s/\sqrt{n}}$ (d) $\bar{X} \pm t_{1-\frac{\alpha}{2}, n-1} \frac{s}{\sqrt{n}}$ (e) $\bar{X} \pm z_{1-\frac{\alpha}{2}} \frac{s}{\sqrt{n}}$

[5] Using $\bar{X} \pm t_{1-\frac{\alpha}{2}, n-1} \frac{s}{\sqrt{n}}$, calculate the confidence interval with $\alpha = 0.05, s = 1.2$ and $n = 13$.

- (a) (1.505, 2.955) (b) (1.511, 2.949) (c) (1.578, 2.882) (d) (15.107, 16.893) (e) (1.803, 2.657)

[6] If a 90% confidence interval for the average gives (1.518, 2.943), **what part is missing in the indicated place** in the interpretation below:

Interpretation: we are 90% sure that the (a) the (b) basal acid output (c) of (d) male smokers is between 1.518 and 2.943.

- (a) If the mean was 2.23 (b) average (c) with a standard deviation of 1.35
(d) 13 (e) all of these are missing

Q#50) Answer 1 to 3 directly from the question. Answer 4 to 9 using new information as given

Statement: In a sample of 50 female students, the mean daily caloric intake was 1849 calories. Assume a distribution that is not normal with a standard deviation of 75. use $\alpha = 0.05$. Test if the average intake is less than 2000 calories

[1] The number 12 is:

- (a) S^2 (b) S (c) σ^2 (d) σ (e) μ_0

[2] The assumptions are:

- (a) normal, σ^2 known (b) not normal, σ^2 known, n large
(c) normal, σ^2 unknown, n large (d) not normal, σ^2 unknown, n large
(e) none of these

[3] The hypotheses are:

- (a) $H_0 : \mu_0 = 2000$ (b) $H_0 : \mu_0 = 2000$ (c) $H_0 : \mu < 2000$ (d) $H_0 : \mu_0 = 2000$ (e) $H_0 : \mu = 2000$
 $H_a : \mu_0 < 2000$ $H_a : \mu < 2000$ $H_a : \mu = 2000$ $H_a : \mu_a < 2000$ $H_a : \mu < 2000$

[4] Suppose that the assumption are "not normal, σ^2 unknown, n large", then the correct test statistic is:

(a) $Z = \frac{\bar{X} - \mu_0}{s/\sqrt{n}}$ (b) $Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$ (c) $T = \frac{\bar{X} - \mu_0}{s/\sqrt{n}}$ (d) $\bar{X} \pm t_{1-\frac{\alpha}{2}, n-1} \frac{s}{\sqrt{n}}$ (e) none of these

[5] The distribution of the test statistic in [4] is:

- (a) $t_{1-\alpha, 49}$ (b) t_{50} (c) t_{49} (d) $z_{1-\alpha}$ (e) standard normal

[6] Suppose that $H_a : \mu > \mu_0$ and we are using a T statistic and $\alpha = 0.05$ (n as in the statement), then the correct rejection rule is:

- (a) Reject H_0 if $T < -1.96$ (b) Reject H_0 if $T > 1.645$
(c) Reject H_0 if $T > 1.6794$ (d) Reject H_0 if $T < -1.96$ or $T > 1.69$
(e) Reject H_0 if $T < -2.0141$ or $T > 2.0141$

[7] Use the statement along with $s = 80$ to calculate the value of $Z = \frac{\bar{X} - \mu_0}{s/\sqrt{n}}$. The answer is:

- (a) -170.84 (b) -84.41 (c) -13.347 (d) -6.412 (e) -2.159

[8] If the rule is "Reject H_0 if $T < -1.6759$ " and $T = -1.38$, then the decision is:

- (a) Reject H_0 (b) Fail to reject H_0 (c) either one of them (d) we can not say (e) $\mu = 35$

[9] If we were testing $H_a : \mu \neq 1900$, and our decision was "Fail to reject H_0 " then choose the **underlined part** in the conclusion that is **incorrect**:

Conclusion: We cannot conclude that the average daily caloric intake for 50 female students is different from 1900.

- (a) cannot (b) average (c) 38 (d) different from (e) none of these (all are correct)

Table B: Quantiles of Student's t-distributions with given degrees of freedom for selected cumulative probabilities.

df	Cumulative probability				
	0.90	0.95	0.975	0.99	0.995
1	3.078	6.3138	12.706	31.821	63.657
2	1.886	2.9200	4.3027	6.965	9.9248
3	1.638	2.3534	3.1825	4.541	5.8409
4	1.533	2.1318	2.7764	3.747	4.6041
5	1.476	2.0150	2.5706	3.365	4.0321
6	1.440	1.9432	2.4469	3.143	3.7074
7	1.415	1.8946	2.3646	2.998	3.4995
8	1.397	1.8595	2.3060	2.896	3.3554
9	1.383	1.8331	2.2622	2.821	3.2498
10	1.372	1.8125	2.2281	2.764	3.1693
11	1.363	1.7959	2.2010	2.718	3.1058
12	1.356	1.7823	2.1788	2.681	3.0545
13	1.350	1.7709	2.1604	2.650	3.0123
14	1.345	1.7613	2.1448	2.624	2.9768
15	1.341	1.7530	2.1315	2.602	2.9467
16	1.337	1.7459	2.1199	2.583	2.9208
17	1.333	1.7459	2.1098	2.567	2.8982
18	1.330	1.7341	2.1009	2.552	2.8784
19	1.328	1.7291	2.0930	2.539	2.8609
20	1.325	1.7247	2.0860	2.528	2.8453
21	1.323	1.7207	2.0796	2.518	2.8314
22	1.321	1.7171	2.0739	2.508	2.8188
23	1.319	1.7139	2.0687	2.500	2.8073
24	1.318	1.7109	2.0639	2.492	2.7969
25	1.316	1.7109	2.0639	2.485	2.7874
26	1.315	1.7081	2.0595	2.479	2.7787
27	1.314	1.7056	2.0555	2.473	2.7707
28	1.313	1.7033	2.0518	2.467	2.7633
29	1.311	1.7011	2.0484	2.462	2.7564
30	1.310	1.7011	2.0452	2.462	2.7500
35	1.3062	1.6991	2.0423	2.457	2.7239
40	1.3031	1.6973	2.0301	2.438	2.7045
45	1.3007	1.6896	2.0211	2.423	2.6896
50	1.2987	1.6839	2.0141	2.412	2.6778
60	1.2959	1.6794	2.0086	2.403	2.6603
70	1.2938	1.6759	2.0003	2.390	2.6480
80	1.2910	1.6707	1.9945	2.381	2.6316
90	1.2887	1.6669	1.9867	2.368	2.6175
100	1.2869	1.6620	1.9799	2.358	2.6070
120	1.2858	1.6577	1.9749	2.350	2.6006
160	1.2858	1.6545	1.9719	2.345	2.576
200	1.282	1.6525	1.96	2.326	
∞	1.282	1.645			