

Please provide computational details for questions and problems to get any credit.

The following problem is associated with questions 1 to 5.

Most presidential polling institutions focus on the likely voters' support for either President Bush or Senator Kerry in the 2004 November 2nd election. A polling organization wishes to estimate the current likely voters' support for the two candidates. To do so, 1050 likely voters nationwide are selected randomly and found that their support for either candidate is tied at 47%.

1. What is the sample statistics?
 - a. The sample percentage of support from the 1050 likely voters for either candidate is at 47%.
 - b. The population percentage of support from all likely voters for either candidate.
 - c. Randomly selected 1050 likely voters national wide
 - d. The percentage of support from likely voters for either candidate..
2. What is the population parameter?
 - a. The sample percentage of support from the 1050 likely voters for either candidate is at 47%.
 - b. The population percentage of support from all likely voters for either candidate.
 - c. Randomly selected 1050 likely voters national wide
 - d. The percentage of support from likely voters for either candidate.
3. What is the sample of this study?
 - a. The sample percentage of support from the 1050 likely voters for either candidate is at 47%.
 - b. The population percentage of support from all likely voters for either candidate.
 - c. Randomly selected 1050 likely voters national wide
 - d. The percentage of support from likely voters for either candidate.
4. What is the variable of interest in the study?
 - a. The sample percentage of support from all likely voters for either candidate.
 - b. Randomly selected 1050 likely voters national wide
 - c. The percentage of support from likely voters for either candidate.
 - d. All likely voters national wide.
5. What is the population from which the study was sampled?
 - a. The population percentage of support from all likely voters for either candidate.
 - b. Randomly selected 1050 likely voters national wide

- c. The percentage of support from likely voters for either candidate.
- d. All likely voters national wide.

The following problem is associated with Questions 6 to 15

The Best Hotel asks its customers to fill out a questionnaire to rate the quality of meals served in the last 12 months. Two of the questions are listed here.

- 1) Would you be willing to pay additional \$10 for better meals?
 _____Yes _____No
- 2) How would you rate the meals served?
 _____Good _____Fair _____Poor

As a consultant to Best Hotel, you are asked to evaluate the design of the questionnaire. Among many shortcomings are the following four potential problem areas. For each question below, please identify the problem area to which each error is best viewed as belonging. Pick one answer for each question. Answers may be used more than once, or not at all. Choices are:

- A). sampling error**
- B). measurement error**
- C). selection bias (or coverage error)**
- D). nonresponse bias (or nonresponse error)**

- 6 What is the type of error it may suffer from if a questionnaire is handed out to each customer who stayed in Best Hotel over the next 12 months?
 A) sampling error
 B) measurement error
 C) selection bias (or coverage error)
 D) nonresponse bias (or nonresponse error)
- 7 What is the type of error if Best Hotel is only interested in the people he calls “Business Travelers”, but his survey may be filled out by customers who stay in Best Hotel for none business purposes?
 A) sampling error
 B) selection bias (or coverage error)
 C) measurement error
 D) nonresponse bias
- 8 What is the type of error it may cause if you decide to ask a sample of customers who have stayed in Best Hotel and filed a complaint over the past year to fill out the survey?
 A) measurement error
 B) sampling error
 C) nonresponse bias
 D) selection bias (or coverage error)
- 9 What is the type of error if customers who do not stay in Best Hotel cannot provide information at all?
 A) sampling error
 B) measurement error
 C) selection bias (or coverage error)
 D) nonresponse bias

- 10 What is the type of error if customers lie about how much they would truly pay for better meals?
- A) **sampling error**
 B) **measurement error**
 C) **selection bias (or coverage error)**
 D) **nonresponse bias**
- 11 What is the type of error if only 90 people can fill out the survey? Even if the other problems are straightened out, there are many customers who will not be surveyed, so the results are only an approximation of the real interest?
- A) **sampling error**
 B) **measurement error**
 C) **selection bias (or coverage error)**
 D) **nonresponse bias**
- 12 The possible responses to the question "How many times have you stayed in Best Hotel?" are values from a
- A). parameter.
 B). categorical random variable.
 C). continuous random variable.
 D). discrete random variable.
- 13 What type of data does the variable for the response to the questions "Would you be willing to pay extra for better meals?" belong to (circle the best one you think fit)?
- A) Categorical and On a nominal scale
 B) Numerical and on a interval scale
 C) Categorical and On a ordinal scale
 D) Discrete and on a nominal scale
- 14 The possible responses to the question "How would you rate the meals served?" in this survey can be best described as _____
- A). parameter.
 B). discrete random variable.
 C). continuous random variable.
 D). categorical random variable.
- 15 If Best Hotel has two types of customers: business or none business, the best sampling strategy would be to use
- A). a cluster sample.
 B). a systematic sample.
 C). a stratified sample.
 D). a simple random sample.

The following problem is associated with Questions 16 to 25.

A City Bus Company has categorized complaints from riders for last August and September as follows with the corresponding ROW, Column and Total Percentages, respectively. Assume the sample is a good representation of the population.

Type of Complaints	Buses sometimes don't stop at bus stops	Bus schedules make transfer difficult	Route maps not available	Col Total
August	25	35	5	65
September	50	70	15	135
RowTotal	75	105	20	200
Row Percentages (%)	Buses sometimes don't stop at bus stops	Bus schedules make transfer difficult	Route maps not available	Col Total

August	0.33	0.33	0.25	0.33
September	0.67	0.67	0.75	0.68
RowTotal	1	1	1	1
Column Percentages (%)	Buses sometimes don't stop at bus stops	Bus schedules make transfer difficult	Route maps not available	Col Total
August	0.38	0.54	0.08	1
September	0.37	0.52	0.11	1
RowTotal	0.38	0.53	0.1	1
Total Percentages (%)	Buses sometimes don't stop at bus stops	Bus schedules make transfer difficult	Route maps not available	Col Total
August	0.13	0.18	0.03	0.33
September	0.25	0.35	0.08	0.68
RowTotal	0.38	0.53	0.1	1

Referring to Tables above when you answer the following questions. You must first identify the relationship between or among the events in question, such as, **Union, Joint or Intersection, Conditional, Simple or Marginal**, and so forth.

- 16 Referring to the Tables above, what is the probability that a random selected rider in the population who complained Route maps not available and the complaint was filed in August?
- A) 0.250 B) 0.425 C) 0.400 D) 0.025
- D
- 17 Referring to the Tables above, what is the probability that a random selected rider filed a complaint in September?
- A) 0.675 B) 0.667 C) 0.525 D) 0.370
- 18 Referring to the Tables above, of those riders complained for Bus schedules make transfer difficult, what is the probability that a random selected rider filed the complaint in August?
- A) 0.333 B) 0.325 C) 0.175 D) 0.538
- 19 Referring to the Tables above, what is the probability that a random selected rider complained for Buses sometimes don't stop at bus stops or filed the complaint in September?
- A) 0.800 B) 0.667 C) 1.000 D) 0.250
- 20 Referring to the Tables above, what is the probability that a random selected rider complained for Buses sometime don't stop at bus stops or complained Route maps not available?
- A) 0.100 B) 0.375 C) 0.475 D) 0.250
- 21 Referring to the Tables above, the events of complaining in August and complaining in September are:
- A). mutually exclusive.
 B). collective exhaustive.
 C). statistically dependent.
 D). all of the above
- 22 Are the events of complaining Bus schedules make transfer difficult and complaining Route maps not available statistically independent?
- A) Yes, because $P(\text{Map}) = P(\text{Map} | \text{Transfer})$

5/7

- B) Yes, because $P(\text{Map}) = P(\text{Transfer})$
- C) No, because $P(\text{Map}) \neq P(\text{Map} | \text{Transfer})$
- D) No, $P(\text{Map}) \neq P(\text{Transfer})$

The following problem is associated with Questions 26 to 40.

A city bus company receives complaints from its riders in the last few months as follows

	January	February	March	April	May	June	July	Total
Complaints X	11	22	18	8	9	27	24	119
$(X - \bar{X})$	-6	5	1	-9	-8	10	7	
$(X - \bar{X})^2$	36	25	1	81	64	100	49	356

Please answer the following questions:

23 What is the relative frequency of number of complaints in March?

- A) $18/119=15.13\%$
- B) 51
- C) 18
- D) $51/119=42.86\%$

24 What is the Cumulative Relative Frequency of number of complaints in March?

- A) 18
- B) 51
- C) $51/119=42.86\%$
- D) $18/119=15.13\%$

25 What is the arithmetic mean (\bar{X}) or the average number of complaints?

- A) 19.0
- B) 20
- C) 18.0
- D) 17.0

26 What is the Mode?

- A) 7.7
- B) 16.5
- C) 13.0
- D) Not available

27 What is the First Quartile Q_1 ?

- A) 9
- B) 8
- C) 11
- D) 10

28 What is the Median or Second Quartile Q_2 ?

- A) 20.0
- B) 18.0
- C) 17.0
- D) 23.0

29 What is the Excel equation to computer the median?

- I. =MEDIAN(range)
- II. =QUARTILE(range,3)
- III. =QUARTILE(range,2)

- A) I only
- B) I and II only
- C) I, II and III
- D) I and III only

30 What is the Interquartile Range?

- A) 15.0
- B) 23.0
- C) 19.0
- D) 13.0

31 What is the Midhinge?

- A) 13.0
- B) 16.5
- C) 19.0
- D) 59.3

32 What is the Sample Standard Deviation ?

- A) 18.0 B) 19.0 C) 7.7 D) 59.3

33 What is the Coefficient of Variation?

- A) 17.0 B) 0.4531 C) 7.7 D) 59.3

34 Construct a Box-and-Whisker Plot for the number of complaints from January to July? Is the data skewed to the right or to the left _____?

- A) Skew to the right, because the mean is greater than the median
 B) Skew to the left, because the mean is greater than the median
 C) Skew to the left, because the mean is less than the median
 D) Skew to the right, because the mean is less than the median

35 In right-skewed distributions, which of the following is the correct statement?

- A). The distance from the smallest observation to Q_1 is smaller than the distance from Q_3 to the largest observation.
 B). The distance from the smallest observation to Q_2 is larger than the distance from Q_2 to the largest observation.
 C). The distance from Q_2 to Q_3 is smaller than the distance from Q_1 to Q_2 .
 D). The distance from Q_1 to Q_3 is twice the distance from Q_2 to Q_3

36 Construct a Steam and Leaf Graph for the number of complaints from January to July?

Stem	Leaf
(0)	8 9
(1)	1 8
(2)	2 4 7

Referring your Stem and Leaf display in part a) of this questions, what percentage of the observations had the number of complaints between ten and twenty?

- A) $2/7=28.6\%$ B) $1/7$ C) 20 D) 2

37 Among the values computed above, which ones are the measures for central tendency?

- A) Mean, Median, and Range
 B) Mean, Mode, and Median
 C) Mean, Range and Standard deviation
 D) Mode, Median and Coefficient of Variation

38 Among the values computed above, which ones are the measures for variation?

- I. Mean II. Interquartile Range
 III. Standard deviation IV. Coefficient of Variation

7/7

- A) I and II only B) II, III and IV only C) I, III and IV only D) I, II AND III only

39 Which of the values computed above is/are sensitive (not resistant) to extreme values?

I. Mean

II. Median

III. Standard deviation

IV. Coefficient of Variation

- A) I and II only B) II and III only C) I, III and IV only D) III and IV only

40 Which of the following is NOT sensitive to extreme values?

- A). The range.
- B). The interquartile range.
- C). The coefficient of variation.
- D). The standard deviation.