**“Beynəlxalq İqtisadiyyat Məktəbi”**

**Statistika**

**Question Part 1**

1. What is a Inferential statistics and Descriptive Statistics?
2. Discuss and provide example of ordinal scale
3. Discuss and provide example of nominal scale
4. Discuss and provide example of Interval Scale
5. Discuss and provide example of Ratio Scale
6. A subset of a population selected to represent the population is called ……. . Also provide an example to your answer in different examples.
7. What is a histogram? Provide an example to your answer.
8. What is range, width of interval and frequency distribution?
9. What is a range and provide an example to show your answer
10. What is the Bar chart and Pie chart? Show the in graphs.
11. Which of the following pair answers can be classified as quantitative data? Explain your answer.



1. In a sample of 800 students in a university, 240, or 30%, are International Business majors. The 30% is an example of which of the below answers? Justify your answers.



1. Which of the measure of the location is the most likely to be influenced by extreme values in the data set? Explain and provide an example.
2. The following data set shows the number of hours of sick leave that some of the employees of ABC Company have taken during the first quarter of the year (rounded to the nearest hour).



Develop a frequency distribution for the above data. (Let the width of your classes be 10 units and start your first class as 10 - 19.)

1. The following data set shows the number of hours of sick leave that some of the employees of ABC Company have taken during the first quarter of the year (rounded to the nearest hour).



Develop a relative frequency distribution for the above data. (Let the width of your classes be 10 units and start your first class as 10 - 19.)

16. KMV hired the following number of Class 1 developers during the first six months of the past year. (Assume the data represent a sample.)

|  |  |
| --- | --- |
| **Month**  | Number of Class 1 Developers |
| **January**  | 3 |
| **February**  | 2 |
| **March** | 4 |
| **April** | 2 |
| **May** | 6 |
| **June** | 0 |

Determine the mean, the median, the mode, and the range for the above data

17. KMV hired the following number of Class 1 developers during the first six months of the past year. (Assume the data represent a sample.)

|  |  |
| --- | --- |
| **Month**  | Number of Class 1 Developers |
| **January**  | 3 |
| **February**  | 2 |
| **March** | 4 |
| **April** | 2 |
| **May** | 6 |
| **June** | 0 |

Compute the variance and the standard deviation

18. KMV hired the following number of Class 1 developers during the first six months of the past year. (Assume the data represent a sample.)

|  |  |
| --- | --- |
| **Month**  | Number of Class 1 Developers |
| **January**  | 3 |
| **February**  | 2 |
| **March** | 4 |
| **April** | 2 |
| **May** | 6 |
| **June** | 0 |

Compute the first and the third quartiles

19. Organize below data for day students only using **Stem and Leaf** display.

|  |  |
| --- | --- |
| **Age of Surveyed College Students** | **Day Students** |
| 16 | 17 | 17 | 18 | 18 | 18 |
| 19 | 19 | 20 | 20 | 21 | 22 |
| 22 | 25 | 27 | 32 | 38 | 42 |
| **Night Students** |
| 18 | 18 | 19 | 19 | 20 | 21 |
| 23 | 28 | 32 | 33 | 41 | 45 |

20. Describe times series data. What is time series data? Could you distinguish it from Cross Sectional and Panel data?

21. Define with formula and provide an example for Mean, Median and geometric mean.

22. Define with formula and provide an example for variance, and standard deviation.

23. KMV hired the following number of Class 1 developers during the first six months of the past year. (Assume the data represent a sample.)

|  |  |
| --- | --- |
| **Month**  | Number of Class 1 Developers |
| **January**  | 3 |
| **February**  | 2 |
| **March** | 4 |
| **April** | 2 |
| **May** | 6 |
| **June** | 0 |

Compute the z-scores for the months of May and June.

24. Define with formula and provide an example for covariance, and correlation. Explain types of correlation.

 25. Discuss and provide examples on followings: *Skewnes; Kurtosi;, Coefficient of variations; Chebyshev Rule.*

**Questions PART 2**

1) The following is a set of data for a population with N=5:

5,90;-9;4,30;7,80;4,40

a) Compute the population mean.

b) Compute the population variance and standard deviation.

2) The following is a set of data from a sample of n=11 items:

X 7; 5;8;3;6;10;12;4;9;15;18

Y 21;15;24;9;18;30;36;12;27;45;54

a) Compute the covariance

b) Compute the coefficient of correlation.

c) How strong is the relationship between X and Y?Explain.

3) A Dubai-based media company is trying to decide whether to bid for a major TV advertising contract.In the past,the company’s main competitor,based in Abu Dhabi,has submitted bids 70% of the time.If the Abu Dhabi competitor does not bid on a job,the probability that Dubai Media Company will get the job is 0,50.If the Abu Dhabi competitor bids on a job,the probability that the Dubai-based company will get the job is 0,25.

a) If the Dubai-based company gets the job,what is the probability that the Abu Dhabi Company did not bid?

4) The manager of the commercial mortgage department of a large bank has collected data during the past two years concerning the number of commercial mortgages approved per week.The results from these two years(104 weeks)indicated the following:

|  |  |
| --- | --- |
| Number of Commercial Mortgages Approved | Frequency |
| 0 | 13 |
| 1 | 25 |
| 2 | 32 |
| 3 | 17 |
| 4 | 9 |
| 5 | 6 |
| 6 | 1 |
| 7 | 1 |

a) Compute the expected number of mortgages approved per week.

b)Compute the standard deviation.

5) You are trying to develop a strategy for investing in two different stocks.The anticipated annual return for a $1000 investment in each stock under four different economic conditions has the following probability distribution:

|  |  |  |  |
| --- | --- | --- | --- |
| Probability | Economic Condition | Stock X | Stock Y |
| 0,1 | Recession | -50 | -100 |
| 0,3 | Slow growth | 20 | 50 |
| 0,4 | Moderate growth | 100 | 130 |
| 0,2 | Fast growth | 150 | 200 |

Compute the

a)expected return for stock X and for stock Y.

b)standard deviation for stock X and for stock Y.

c)covariance of stock X and stock Y.

d)Would you invest in stock X or stock Y?Explain.

6)The following is a set of data for a population with N=5:

5,90;-9;4,30;7,80;4,40

a)Compute the population mean.

b)Compute the population variance and standard deviation.

7)The following is a set of data from a sample of n=11 items:

X 7; 5;8;3;6;10;12;4;9;15;18

Y 21;15;24;9;18;30;36;12;27;45;54

a)Compute the covariance

b)Compute the coefficient of correlation.

c)How strong is the relationship between X and Y?Explain.

8)A Dubai-based media company is trying to decide whether to bid for a major TV advertising contract.In the past,the company’s main competitor,based in Abu Dhabi,has submitted bids 70% of the time.If the Abu Dhabi competitor does not bid on a job,the probability that Dubai Media Company will get the job is 0,50.If the Abu Dhabi competitor bids on a job,the probability that the Dubai-based company will get the job is 0,25.

a)What is the probability that the Dubai-based company will get the job?

9)The manager of the commercial mortgage department of a large bank has collected data during the past two years concerning the number of commercial mortgages approved per week.The results from these two years(104 weeks)indicated the following:

|  |  |
| --- | --- |
| Number of Commercial Mortgages Approved | Frequency |
| 0 | 13 |
| 1 | 25 |
| 2 | 32 |
| 3 | 17 |
| 4 | 9 |
| 5 | 6 |
| 6 | 1 |
| 7 | 1 |

a)Compute the expected number of mortgages approved per week.

b)Compute the standard deviation.

c) provide relative and percentage frequencies

10)You are trying to develop a strategy for investing in two different stocks.The anticipated annual return for a $1000 investment in each stock under four different economic conditions has the following probability distribution:

|  |  |  |  |
| --- | --- | --- | --- |
| Probability | Economic Condition | Stock X | Stock Y |
| 0,1 | Recession | -20 | -50 |
| 0,3 | Slow growth | 30 | 40 |
| 0,4 | Moderate growth | 200 | 180 |
| 0,2 | Fast growth | 50 | 100 |

Compute the

a)expected return for stock X and for stock Y.

b)standard deviation for stock X and for stock Y.

c)covariance of stock X and stock Y.

d)Would you invest in stock X or stock Y?Explain.

11) The distribution of the annual incomes of a group of middle-management employees at Compton Plastics approximates a normal distribution with a mean of $37200 and a standard deviation of $800.

a) what is the probability that income will be $38000?

b) what is the probability that income will be less than $35000?

c) what it the probability that income will be more than 38800 and less than $39000?

12) The percentage change in the Russel 2000 Index of the stock prices of 2000 small companies was

 -33,79% in 2008 and 27,17% in 2009.Compute the geometric mean rate of return per year.

13) Define and discuss continues probability distribution. Explain

14) Define and discuss various method of sampling.

15) Define and discuss discrete and continues probability distributions.

16) The probability that a person has a certain disease is 0,03.Medical diagnostic tests are available to determine whether the person actually has the disease.If the disease is actually present,the probability that the medical diagnostic test will give a positive result(indicating that the disease is present)is 0,90.If the disease is not actually present,the probability of a positive test result(indicating that the disease is present)is 0,02.Suppose that the medical diagnostic test has given a positive result(indicating that the disease is present).What is the probability that the disease is actually present?

17) Consider a population of 1024 investment funds that primarily invest in large companies.You have determined μ,the mean one-year total percentage return achieved by all the funds,is 8,20 and that σ,the standard deviation,is 2,75.

- According to the Chebyshev rule,what percentage of these funds are expected to be within +-1,+-2,or +-3 standard deviations of the mean?

18) Consider a population of 1024 investment funds that primarily invest in large companies.You have determined μ,the mean one-year total percentage return achieved by all the funds,is 8,20 and that σ,the standard deviation,is 2,75.

- According to the empirical rule,what percentage of these funds is expected to be within +-1 standard deviation of the mean?

19) After deducting grants based on need, the average cost to attend the University of Agriculture is $37,300. Assume the population standard deviation is $7400. Suppose that a random sample of 60 USC students will be taken from this population.

- What is the value of the standard error of the mean?

- What is the probability that the sample mean will be more than $37,300?

- How would the probability in part (c) change if the sample size were increased to 200?

20) After deducting grants based on need, the average cost to attend the University of Agriculture is $37,300. Assume the population standard deviation is $7400. Suppose that a random sample of 60 USC students will be taken from this population.

- What is the value of the standard error of the mean?

- What is the probability that the sample mean will be within $1000 of the population mean?

- How would the probability in part (c) change if the sample size were increased to 150?

21) What is the probability that a simple random sample of 40 applicants will provide an estimate of the population mean GMAT score that is within +/-50 of the actual population mean *m* ? Please note that mean and standard deviation are 750 and 17.9. Also show your answer in graphical way. Suppose we select a simple random sample of 20 applicants instead of the 40 originally considered. What is changed? Explain.

22) What is the probability that a simple random sample of 50 applicants will provide an estimate of the population mean GMAT score that is within +/-50 of the actual population mean *m* ? Please note that mean and standard deviation are 750 and 17.9. Also show your answer in graphical way. Suppose we select a simple random sample of 60applicants instead of the 50 originally considered. Use central limit to explain your answer.

23) Discuss various methods of Assigning Probabilities. Explain your answer using arguments.

24) Define and provide explanation for Complement of an Event; Union of Two Events; Intersection of Two Events; Mutually Exclusive Events. Use graphs/diagrams as much as you can.

25) Define and explain Mutual Exclusiveness and Independence of events. Define and explain Bayes’ Theorem.

**Questions PART 3**

1. The average price of homes sold in Sweden in the past year was $120,000. A random sample of 81 homes sold this year showed an average price of $110,000. It is known that the standard deviation of the population is $36,000. At 95% confidence test to determine if there has been a significant decrease in the average price homes.

a. State the null and alternative hypotheses to be tested.

 b. Compute the test statistic.

c. Determine the critical value for this test. Also confirm the conclusion using p-value

d. What do you conclude?

e. Compute the p-value.

2.

The average price of homes sold in Sweden in the past year was $120,000. A random sample of 81 homes sold this year showed an average price of $110,000. It is known that the standard deviation of the population is $36,000. At 90% confidence test to determine if there has been a significant decrease in the average price homes.

a. State the null and alternative hypotheses to be tested.

 b. Compute the test statistic.

c. Determine the critical value for this test. Also confirm the conclusion using p-value

d. What do you conclude?

e. Compute the p-value.

1. In a test of the quality of two television commercials, each commercial was shown in a separate test area six times over a one-week period. The following week a telephone survey was conducted to identify individuals who had seen the commercials. Those individuals were asked to state the primary message in the commercials. the following results were recorded.



a. Use α 0.05 and test the hypothesis that there is no difference in the recall proportions for the two commercials.

b. Compute a 95% confidence interval for the difference between the recall proportions for the two populations.

1. For a random sample of 125 US entrepreneurs, the mean number of job changes was 1.91 and the sample standard deviation was 1.32. For an independent random sample of 86 US corporate managers, the mean number of job changes was 0.21 and the sample standard deviation was 0.53.

Test the null hypothesis that the population means are equal against the alternative that the mean number of job changes is higher for the US entrepreneurs than for US corporate managers

1. A company selling licenses for new e-commerce computer software advertises that firms using this software obtain, on average during the first year, a minimum yield of 10% on their initial investments. A random sample of 10 of these franchises produced the following yields for the first year of operation:

5.1, 9.2, 10.5, 8.6, 11.1, 3.9, 8.4, 10.1, 7.4, 9.9 Assuming that population yields are normally distributed,

Test the company's claim with a significance level of 5% (.05).

1. A company selling licenses for new e-commerce computer software advertises that firms using this software obtain, on average during the first year, a minimum yield of 10% on their initial investments. A random sample of 10 of these franchises produced the following yields for the first year of operation:

5.1, 9.2, 10.5, 8.6, 11.1, 3.9, 8.4, 10.1, 7.4, 9.9 Assuming that population yields are normally distributed,

Test the company's claim with a significance level of 1% (.01).

1. Sparr Investments, Inc., specializes in tax-deferred investment opportunities for its clients. Recently Sparr offered a payroll deduction investment program for the employees of a particular company. Sparr estimates that the employees are currently averaging $100 or less per month in tax-deferred investments. A sample of 40 employees will be used to test Sparr’s hypothesis about the current level of investment activity among the population of employees. Assume the employee monthly tax-deferred investment amounts have a standard deviation of $75 and that a .05 level of significance will be used in the hypothesis test.

a. What is the Type II error in this situation?

b. What is the probability of the Type II error if the actual mean employee monthly investment is $120?

c. What is the probability of the Type II error if the actual mean employee monthly investment is $130?

d. Assume a sample size of 80 employees is used and repeat parts (b) and (c).

8. In the past, 75% of the tourists who visited Chattanooga went to see Rock City. The management of Rock City recently undertook an extensive promotional campaign. They are interested in determining whether the promotional campaign actually **increased** the proportion of tourists visiting Rock City.

State the correct set of hypotheses. How would you test this hypothesis?

1. Your investment executive claims that the average yearly rate of return on the stocks she recommends is at least 10.0%. You plan on taking a sample to test her claim.

State the correct set of hypotheses. How would you test this hypothesis? Explain

1. We are interested in determining whether or not the variances of the sales at two small grocery stores are equal. A sample of 21 days of sales at Store A and a sample of 16 days of sales at Store B indicated the following.



|  |  |
| --- | --- |
| a.  | Provide the hypotheses to be tested.  |
| b.  | Compute the test statistic.  |
| c.  | Determine the critical value of F at 95% confidence.  |
| d.  | Compute the *p*-value and use it to test the above hypotheses  |

11 he Consumer Reports National Research Center conducted a telephone survey of 2000 adults

to learn about the major economic concerns for the future (*Consumer Reports,* January 2009).

The survey results showed that 1760 of the respondents think the future health of Social

Security is a major economic concern.

a. What is the point estimate of the population proportion of adults who think the future health

of Social Security is a major economic concern.

b. At 90% confidence, what is the margin of error?

c. Develop a 90% confidence interval for the population proportion of adults who think the

future health of Social Security is a major economic concern.

d. Develop a 95% confidence interval for this population proportion

12. The percentage of people not covered by health care insurance in 2003 was 15.6% (*Statistical*

*Abstract of the United States*, 2006). A congressional committee has been charged

with conducting a sample survey to obtain more current information.

a. What sample size would you recommend if the committee’s goal is to estimate the current

proportion of individuals without health care insurance with a margin of error of

.03? Use a 95% confidence level.

b. Repeat part (a) using a 99% confidence level.

13. The 92 million Americans of age 50 and over control 50 percent of all discretionary income

(*AARP Bulletin,* March 2008). AARP estimated that the average annual expenditure

on restaurants and carryout food was $1873 for individuals in this age group. Suppose this

estimate is based on a sample of 80 persons and that the sample standard deviation is $550.

a. At 95% confidence, what is the margin of error?

b. What is the 95% confidence interval for the population mean amount spent on

restaurants and carryout food?

c. What is your estimate of the total amount spent by Americans of age 50 and over on

restaurants and carryout food?

d. If the amount spent on restaurants and carryout food is skewed to the right, would you

expect the median amount spent to be greater or less than $1873?

14. *Condé Nast Traveler* conducts an annual survey in which readers rate their favorite cruise

ship. All ships are rated on a 100-point scale, with higher values indicating better service.

Asample of 37 ships that carry fewer than 500 passengers resulted in an average rating of

85.36, and a sample of 44 ships that carry 500 or more passengers provided an average rating

of 81.40 (*Condé Nast Traveler,* February 2008). Assume that the population standard

deviation is 4.55 for ships that carry fewer than 500 passengers and 3.97 for ships that carry

500 or more passengers.

a. What is the point estimate of the difference between the population mean rating for

ships that carry fewer than 500 passengers and the population mean rating for ships

that carry 500 or more passengers?

b. At 95% confidence, what is the margin of error?

c. What is a 95% confidence interval estimate of error?

15. Are nursing salaries in Tampa, Florida, lower than those in Dallas, Texas? Salary data

show staff nurses in Tampa earn less than staff nurses in Dallas (*The Tampa Tribune*, January

15, 2007). Suppose that in a follow-up study of 40 staff nurses in Tampa and 50 staff

nurses in Dallas you obtain the following results.



a. Formulate hypothesis so that, if the null hypothesis is rejected, we can conclude that

salaries for staff nurses in Tampa are significantly lower than for those in Dallas.

Use *α* = .05.

b. What is the value of the test statistic?

c. What is the *p*-value?

d. What is your conclusion?

16. The manager of the Danvers-Hilton Resort Hotel stated that the mean guest bill for a weekend

is $600 or less. A member of the hotel’s accounting staff noticed that the total charges

for guest bills have been increasing in recent months. The accountant will use a sample of

weekend guest bills to test the manager’s claim.

1. Which form of the hypotheses should be used to test the manager’s claim? Explain



b. What conclusion is appropriate when *H*0 cannot be rejected?

c. What conclusion is appropriate when *H*0 can be rejected?

17. A production line operation is designed to fill cartons with laundry detergent to a mean

weight of 32 ounces. A sample of cartons is periodically selected and weighed to determine

whether underfilling or overfilling is occurring. If the sample data lead to a conclusion

of underfilling or overfilling, the production line will be shut down and adjusted to

obtain proper filling.

a. Formulate the null and alternative hypotheses that will help in deciding whether to shut

down and adjust the production line.

b. Comment on the conclusion and the decision when *H*0 cannot be rejected.

c. Comment on the conclusion and the decision when *H*0 can be rejected.

18. Carpetland salespersons average $8000 per week in sales. Steve Contois, the firm’s vice

president, proposes a compensation plan with new selling incentives. Steve hopes that the

results of a trial selling period will enable him to conclude that the compensation plan increases

the average sales per salesperson.

a. Develop the appropriate null and alternative hypotheses.

b. What is the Type I error in this situation? What are the consequences of making this error?

c. What is the Type II error in this situation? What are the consequences of making this error?

19. Suppose a new production method will be implemented if a hypothesis test supports the

conclusion that the new method reduces the mean operating cost per hour.

a. State the appropriate null and alternative hypotheses if the mean cost for the current

production method is $220 per hour.

b. What is the Type I error in this situation? What are the consequences of making this error?

c. What is the Type II error in this situation? What are the consequences of making this error?

20. Fowle Marketing Research, Inc., bases charges to a client on the assumption that telephone

surveys can be completed within 15 minutes or less. If more time is required, a premium

rate is charged. With a sample of 35 surveys, a population standard deviation of 4 minutes,

and a level of significance of .01, the sample mean will be used to test the null hypothesis

*H*0: *μ* \_ 15.

a. What is your interpretation of the Type II error for this problem? What is its impact

on the firm?

b. What is the probability of making a Type II error when the actual mean time is

*μ* \_ 17 minutes?

c. What is the probability of making a Type II error when the actual mean time is

*μ* \_ 18 minutes?

d. Sketch the general shape of the power curve for this test.

21. At Western University the historical mean of scholarship examination scores for freshman

applications is 900. Ahistorical population standard deviation *σ* \_ 180 is assumed known.

Each year, the assistant dean uses a sample of applications to determine whether the mean

examination score for the new freshman applications has changed.

a. State the hypotheses.

b. What is the 95% confidence interval estimate of the population mean examination

score if a sample of 200 applications provided a sample mean \_ 935?

c. Use the confidence interval to conduct a hypothesis test. Using *α* \_ .05, what is your

conclusion?

d. What is the *p*-value?

22. A random sample of 49 business statistics examinations was taken. The average score, in the sample, was 84 with a variance of 12.25. The 95% confidence interval for the average examination score of the population of the examinations is close to \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ What happen if you change confidence interval to 99%. Comment on it.

23. *BusinessWeek* conducted a survey of graduates from 30 top MBA programs (*BusinessWeek,*

September 22, 2003). On the basis of the survey, assume that the mean annual salary for male

and female graduates 10 years after graduation is $168,000 and $117,000, respectively.

Assume the standard deviation for the male graduates is $40,000, and for the female graduates

it is $25,000.

a. What is the probability that a simple random sample of 40 male graduates will provide

a sample mean within $10,000 of the population mean, $168,000?

b. What is the probability that a simple random sample of 40 female graduates will

provide a sample mean within $10,000 of the population mean, $117,000?

c. In which of the preceding two cases, part (a) or part (b), do we have a higher probability

of obtaining a sample estimate within $10,000 of the population mean? Why?

d. What is the probability that a simple random sample of 100 male graduates will provide

a sample mean more than $4000 below the population mean?

24. Explain and discuss followings: F- test; T-test; t-distribution; z-normal distribution; Simple linear regression; R-Square; TSS; SSE; Coefficient of determination;

25. Explain followings using simple linear regression;

a) explain dependent and independent variables

b) A regression analysis between sales (Y in $1000) and advertising (X in dollars) resulted in the following equation: = 30,000 + 4 X. EXPLAIN This in detail.

c) In regression analysis, the model in the form this : explain each variable and assumptions used for each variable (wherever possible) in equation it in detail.