1. Explain **experimental** and **observational data** and give an example.
2. Explain types of data and give at least one separate example related to each of them.
3. Find …….. show all steps explicitly and interpret your result.

4. Look at the table and find …… and …… for  and 

|  | *x=0* | *x=1* |
| --- | --- | --- |
| *y=0* | ….. | ….. |
| *y=1* | …… | …… |

5. Suppose that ….. and … … Determine ……..and discuss your finding.

6. Imagine ***M*** is any discrete random variable with five possible outcomes starting from 0. The associated probabilities of each outcomes are followings; …., …, …, …., ..… Find …… …

7. How we interpret the shape of the any given distribution? Explain the idea behind the moments of the distribution.

8. Given ……..find probability density function for P(X>2).

9. Look at the table and find …. and …..

|  | *x=0* | *x=1* |
| --- | --- | --- |
| *y=0* | … | …. |
| *y=1* | … | … |

10.  *X* and *Y* are discrete random variables with the following joint distribution:

|  | *y=14* | *y=22* | *y=30* | *y=40* | *y=65* |
| --- | --- | --- | --- | --- | --- |
| *x=1* | .. | .. | … | … | … |
| *x=5* | … | … | …. | … | … |
| *x=8* | … | … | … | … | …. |

Calculate the covariance and correlation between *X* and *Y* .

11. *X* is a random variable with moment.. … Show …

12. *X* is a random variable with moment .… Show ……

13. Compute the following probabilities:

 a. If *Y* is distributed *N*(1,4), find …..

 b. If Y is distributed N(3,9), find……

14. Compute the following probabilities:

 a. If *Y* is distributed *N*(50,25), find ….

 b. If *Y* is distributed *N*(5,2), find…

15. Compute the following probabilities:

 a. If *Y* is distributed , find…

 b. If *Y* is distributed , find ….

16. Compute the following probability:

 If *Y* is distributed , find ….

17. Compute the following probabilities:

 a. If *Y* is distributed , find…..

 b. If *Y* is distributed , find…..

18. True or False? If   and and……

19. True or False? If  and …..then *e* is independent of *x*.

20. Using matrix algebra, show …. Where  is the OLS residual from a regression of *y* on *X*.

21. Let  be the OLS residual from a regression of *y* on *X.* Find the OLS coefficient from a regression of  on *X* .

22. Under homoskedasticity : Find …..Interpret your finding.

23. Show that .

24. Let …..Find the OLS coefficient from a regression of  on *X* .

25. A dummy variable takes on only values 0 and 1. It is used for categorical data, such as an individual’s gender. Let  and  be vectors of 1’s and 0’s, with the  element of  equaling 1 and that of  equaling 0 if the person is male, and the reverse if the person is female. Suppose that there are  men and women in the sample.

In the OLS regression  show that  is the sample mean of the dependent variable among the men of the sample, and that  is the sample mean among the women.

**Q1.** Explain the difference between the sample average and the population mean. Explain the difference between an estimator and an estimate. Provide an example of each.

**Q2.** A population distribution has a mean of 10 and a variance of 16. Determine the mean and variance of from an i.i.d. sample from this population for ….and ..… Relate your answers to the law of large numbers.

**Q3.** What role does the central limit theorem play in statistical hypothesis testing? In the construction of confidence intervals? Why does a confidence interval contain more information than the result of a single hypothesis test?

**Q4.** What is the difference between a null and alternative hypothesis? Among size, significance level, and power? Between a one-sided and two-sided alternative hypothesis?

**Q5.** Explain tools of statistical inferences (SI) in detail and with examples.

**Q6.** Sketch a hypothetical scatterplot for a sample of size 10 for two random variables with a population correlation of …..

**Q7.** In a population ……, …….Use the central limit theorem to answer the following;

 a. In a random sample of size , find…..

 b. In a random sample of size , find…..

 c. In a random sample of size , find……

**Q8.** Proof that the sample variance is an unbiased estimator of the population variance when  with mean and variance .

 **Hints:**

 a. ……

 b…..…

 c. ……

**Q9.** a.  is an unbiased estimator of . Is an unbiased estimator of 

 b.  is a consistent estimator of . Is a consistent estimator of 

**Q10.** Explain the difference between and ; between the residual and the regression error ; and between the OLS predicted value and 

**Q11.** For each least squares assumptions, provide an example in which the assumption is valid, and then provide an example in which the assumption fails.

**Q12.** Sketch a hypothetical scatterplot of data for an estimated regression with…….

**Q13.** Suppose that a researcher, using data on class size (CS) and average test scores from 100 third-grade classes, estimates the OLS regression,

 ……

 a. A classroom has 22 students. What is the regression’s prediction for that classroom’s average test score?

 b. What is the sample standard deviation of test scores across the 100 classrooms?

**Q14.** Suppose that a researcher, using data on class size (CS) and average test scores from 100 third-grade classes, estimates the OLS regression,

 ……..

 a. Last year a classroom had 19 students, and this year it has 23 students. What is the regression’s prediction for the change in the classroom average test score?

 b. The sample average class size across the 100 classrooms is 21.4. What is the sample average of the test scores across the 100 classrooms?

**Q15.** Suppose that a random sample of 200 twenty-year-old mean is selected from a population and that these men’s height and weight are recorded. A regression of weight on height yields

 ………..

 where *Weight* is measured in pounds and *Height* is measured in inches.

 What is the regression’s weight prediction for someone who is 70 inches tall? 65 inches tall? 74 inches tall?

**Q16.** Suppose that a random sample of 200 twenty-year-old mean is selected from a population and that these men’s height and weight are recorded. A regression of weight on height yields

 ……..

 where *Weight* is measured in pounds and *Height* is measured in inches.

 a. A man has a late growth spurt and grows 1.5 inches over the course of a year. What is the regression’s prediction for the increase in this man’s weight?

 b. Suppose that instead of measuring weight and height in pounds and inches, these variables are measured in centimetres and kilograms. What are the regression estimates from this new centimetre-kilogram regression ? (Give all results, estimated coefficients, , and SER.)

**Q17.** A regression of average weekly earnings (AWE, measured in dollars) on age (measured in years) using a random sample of college-educated full-time workers aged 25-65 yields the following:

 ……

 a. Explain what the coefficient values 696.7 and 9.6 mean.

 b. The SER is 624.1. What are the units of measurement for the SER(dollar?years? or is SER unit-free)?

**Q18.** A regression of average weekly earnings (AWE, measured in dollars) on age (measured in years) using a random sample of college-educated full-time workers aged 25-65 yields the following:

 ………

 a. What are the units of measurement for the (dollars? years? or is unit- free)?

 b. Will the regression give reliable prediction for 99-year-old worker? Why or why not?

**Q19.** A regression of average weekly earnings (AWE, measured in dollars) on age (measured in years) using a random sample of college-educated full-time workers aged 25-65 yields the following:

 …….

 a. What is the regression’s predicted earnings for a 25-year-old worker? A 45-year-old worker?

 b.Given what you know about the distribution of earnings, do you think it is plausible that the distribution of errors in the regression is normal? (**Hint:** Do you think that the distribution is symmetric or skewed? What is the smallest value of earnings, and is it consistent with a normal distribution?)

**Q20.**  A regression of average weekly earnings (AWE, measured in dollars) on age (measured in years) using a random sample of college-educated full-time workers aged 25-65 yields the following:

 ………

 The average age in this sample is 41.6 years. What is the average value of AWE in the sample?

**Q21.** Show that the first least squares assumption, , implies that

 …….

**Q22.** Show that is an unbiased estimator of .

**Q23.**  a. A linear regression yields . Show that .

 b. A linear regression yields . Does this imply that ?

**Q24.** Consider the regression model …….

 a. Suppose you know that … … Derive a formula for the least squares estimator of 

 b. Suppose you know that .… Derive a formula for the least squares estimator of 

**Q25.**  a. Show that the regression in the regression of *Y* on *X* is the squared value of the sample correlation between *X* and *Y*. That is, show that 

 b. Show that the from the regression of *Y* on *X* is the same as the from the regression of *X* on *Y*.

1. Show that …..
2. Show that ….
3. Show that……
4. Show that …..
5. Show that ……
6. Find  from…..
7. Find  from……
8. Explain **in detail** the times series regression !
9. Provides **two main reasons** why we transform data in to their natural logarithms !
10. Explain and show both **numerically** and **theoretically** relation between logarithms and percentage changes !

| Inflation in the USA in | 2004 Q1 and 2005 Q1 |  |
| --- | --- | --- |
| Quarter | U.S. CPI | Inf\_t |
| 2004 Q1 | 186.57 |  |
| 2004 Q2 | 188.60 |  |
| 2004 Q3 | 189.37 |  |
| 2004 Q4 | 191.03 |  |
| 2005 Q1  | 192.17 |  |

a. Find the possible annualised inflation rate for the following quarters. And show that it is the same with difference of logarithms approximation.

12. What is autocorrelation and auto covariance ? And explain the possible interpretation of the autocorrelation !

13. Explain **in detail** forecasts and forecast error versus OLS predicted values and OLS residuals !

14. How should be lag selection and **why** ?

| *p* | BIC(*p*) |
| --- | --- |
| **0** | 1.095 |
| **1** | 1.067 |
| **2** | 955 |
| **3** | 957 |
| **4** | 986 |
| **5** | 1.016 |
| **6** | 1.046 |

15. Explain **in detail** lag length selection using the Information Criteria (IC) in autoregressive model!

16. Explain non-stationarity with **trends** and with **breaks in detail** !

17. Explain the idea behind the spurious regression and provide geometric findings !

18. Explain the difference between and **in detail** !

19. Provide and explain the three useful property about !

20. Why degrees of freedom correction is important in econometrics and how you actually interpret it ?

21. Drive ……

22. Given that and and 

 Find ……

23. Given that …..and ……for any predefined single regression interpret measurement values.

24. Inter-corporate  and from single OLS regression !

25. Explain economic meanings for TSS, ESS, and SSR !