CE311S PROBLEM EXAMPLES

The following are quick-test, straightforward problems for students to hone their skills, for rapid response on five different topic areas. These were developed by TAs for Dr Kara Kockelman's CE311S students, and for all other interested students of probability and statistics. Originally developed as adaptive problems (changing for each new student, online) using Brownstone software, the problems have been fixed here (since Brownstone software no longer exists).

MODULE 1: PROBABILITY

1. Let A_1, A_2, ..., A_k,, be a collection of k mutually exclusive & exhaustive events, each of which has a non-zero probability. Then given any other event B which also has a non-zero probability, the following formula holds:

$$P(A_j|B) = \frac{P(B|A_j)P(A_j)}{\sum_{i=1}^{k} P(B|A_i)P(A_i)}$$

What probability concept does this represent?

a) The Fundamental Theorem
b) Bayes' Theorem
c) Strong Law of Large Numbers
d) The Weak Law of Large Numbers

2. The probability that a building inspector will pass a building is exactly 0.5. What is the probability that he passes 9 buildings in a row?

3. A certain contractor has to decide on the type of dye to color the concrete walkway for a project. If there are 84 dyes numbered from 1 to 84, what is the probability of randomly selecting dye number 51 from the series? Assume that all dyes have an equal probability of being selected.

4. A sound wall is being built next to a freeway. 7 different colored bricks are being used, and the design stipulates that every six bricks (across) must be of a different color. How many ways (orders) could the six brick colors be placed?

5. Suppose *A* and *B* are mutually exclusive events for which P(*A*)=0.27 and P(*B*)=0.45. What is the probability that both *A* and *B* occur?

6. 10 buildings to be inspected today and there are 10 inspectors. How many different ways could the buildings and inspectors be paired up?

7. You are in charge of a construction project and must cancel work whenever it rains. 3 days (not necessarily concurrent) have been selected to be the days when an expensive crane and operator are to be rented. At least 50% of the cost of the crane and operator fee must be paid, even if it rains, to reserve their services. The probability that it rains on a certain day is 0.54 (which is independent of whether it rained on any other day). What is the probability that it rains all 3 days?

8. A crazy undergraduate is taking the FE exam and does not know the answer to 4 questions. He decides to bubble in a random answer, but, because he has already read the questions, his answers are biased and the probability that he gets the (a particular) question right 1/3. What is the probability that he gets all 4 questions correct?

9. Using the standard English alphabet, how many 4 letter words can be created? Assume that letters can be repeated. Note: Two words with the same letters but in different *orders* are considered

different words.
10. A planner is designing a really fancy subdivision with two different house styles: "A" and "B." How many different ways can 1 "A" and 3 "B" styles be placed in a row along a street? Assume two houses of the same style are identical; that is, if two houses of style "A" are next to each other and then they are switched, these two permutations *are not* different for the purposes of this question.

11. There are 52 types of culvert-to-reservoir connectors available. A test facility can test up to 5 at a time. How many different (unordered) groups of size 5 could be tested first at this facility?

12. Suppose A and B are mutually exclusive events for which $P(A)=0.31$ and $P(B)=0.34$. What is the probability that either A or B occurs?

13. Suppose A represents the event that a reservoir is contaminated with raw sewage, B represents the event that the reservoir has swimmers on any given day, and that A and B are mutually exclusive events. Let $P(A)=0.34$ and $P(B)=0.2$. What is the probability that nobody swims in the reservoir *and* no raw sewage contaminates it?

14. Say you have 9 fair coins (each has P(heads) = P(tails) = 0.5)). If you flip each of the coins, what is the probability that 6 of them are heads and 3 of them are tails?

15. Your engineering firm is deciding between two huge projects to select and has decided to poll their employees. As the projects are very similar, the probability that any one employee votes for any project is 0.5. The firm has 3 senior engineers and 4 junior engineers. What is the probability that all of them vote for the same project?

16. A certain unfair 6-sided die turns up even numbers ("2", "4", and "6") twice as often as odd numbers ("1", "3", and "5"). If the die is rolled 4 times, what is the probability that every roll turns up an even number?

17. 63 houses are numbered from 1 to 63, with odd numbers on one side of the street, evens on the other. A building inspector plans to inspect all of these buildings. However, for variety, he wants to start from a random spot and go from there. Because of the way he drives in, he will definitely select an odd numbered house first. If all of the odd numbered houses have equal probability of being selected, what is the probability that he selects house #3?

18. An online dating service is holding a contest in which they are going to select a single man and a single woman to go on a dream date to Malta. If 17 men and 20 women have entered the contest, how many different winning pairs are possible?

19. A new type of road surface is being tested. The surface consists of layers of asphalt and concrete, each of equal thickness. The only stipulations are that the top and bottom layers must be asphalt and the rest of the layers must alternate. In order to get a good feel for the strength of the layers, 2 kinds of asphalt and 1 concrete mixtures are going to be used. How many different ways can these materials be put together in accordance with the stipulations, if each type of concrete/asphalt can be used for, at most, one layer?

20. Suppose Erroll will pass his statistics test with probability 0.7, pass his fluid mechanics test with probability 0.5, and pass both with probability 0.37. What is the probability that Erroll will pass either the statistics or fluid mechanics test, but not both?

21. Suppose Erroll will pass his statistics test with probability 0.7, pass his fluid mechanics test with probability 0.5, and pass both with probability 0.34. What is the probability that Erroll will pass neither the statistics or fluid mechanics test?

22. A certain unfair 6-sided die turns up even numbers ("2", "4", and "6") twice as often as odd numbers ("1", "3", and "5"). If the die is rolled 8 times, what is the probability that 3 rolls turn up even numbers and 5 rolls turn up odd numbers?

23. A certain type of engineering license # is composed of 6 digits, the first 4 of which are numbers (0 through 9) and the final 2 of which are letters (a through z). How many license numbers can be used?

24. 7 large concrete blocks are going to make up one side of a foundation of a shed. Two of the blocks have a half circle hole in the side of them, so when they are put together, electrical conduits can be run under the building. If these two blocks must always be next to each other, how many ways can the 7 blocks be placed in a row to make up the foundation section?

25. How many different 5 letter "words" without any repeated letters can be created from the standard 26 letter English alphabet? Note: Two words with the same letters but in different *orders* are considered different words.

26. A 52-tract home subdivision has 4 different house styles, with 13 houses of each style. A crazy planner has placed the houses in a completely random manner. Given this information, what is the probability that along a certain street, 7 houses in a row are all of the same style?

27. A certain engineering license number is made up of 6 characters: the first 3 are numbers (0 through 9), while the final 3 are letters (a through z). How many different license numbers can be made where *no* characters are repeated?

28. A government board is to select, from 8 houses and 5 commercial buildings, 4 houses and 2 commercial buildings to be inspected. How many different combinations are possible?

29. Given a series of integers:
1,2,3,...,k,...,105
Let S_k represent the following sum $S_k = \sum_{i=1}^{k} i$
If each number in the series has an equal probability of being chosen, what is the probability that, for a randomly drawn k, S_k= 435?

Module 1 Solutions:

1) b	2) 0.001953	3) 0.0119	4) 5,040
5) 0	6) 3,628,800	7) 0.157464	8) 0.01235
9) 456,976	10) 4	11) 2,598,960	12) 0.65
13) 0.46	14) 0.001953	15) 0.0156	16) 0.19753
17) 0.03125	18) 340	19) 2	20) 0.46
21) 0.14	22) 0.06828	23) 6,760,000	24) 1,440
25) 7,893,600	26) 5E-5	27) 11,232,000	28) 700
29) 0.009524			

MODULE 2: DISCRETE DISTRIBUTIONS

1. You are working for a civil engineering firm doing work on a large water project. You are examining data which records the average water level of various reservoirs. It is found that the maximum reservoir level is 300m and the minimum is 25m. Is this data discrete or continuous?

2. On any given day, a certain construction site will have more than one supervisor on site with probability P. If you visit the site 13 times and there is more than one supervisor there 7 times, what is the value of P?

3. The number of cars passing ECJ during the middle of the day can be described as a Poisson process. On the average, 9 cars pass by the building every minute. What is the expected number of cars passing the building in 4 hours?

4. A pile of bricks at a construction site consists of 11 red bricks and 13 black bricks. A brick is randomly selected from the pile, its color noted, and then placed back in the pile. What is the expected number of draws until a black brick is selected?

5. The Electric Acidtest Company produces batteries that will be defective (independently) with probability 0.002. If the company sells them in boxes with 452 batteries each, what is the expected number of defective batteries in a box?

6. An average of 30 fish per day enter the intake pipe of a water treatment plant. A hydraulic engineer designing the plants intake system (as well as its fish release system) decides to model the number of entering fish using a Poisson distribution. In this case, what is the variance of the number of fish entering per day?

7. What is the expected value of the random variable, X? The pmf of X is shown below:

$$P(x) = \begin{matrix} 0.30 & x = 1 \\ 0.06 & x = 2 \\ 0.35 & x = 3 \\ 0.29 & x = 4 \end{matrix}$$

8. The number of earthquakes in a given month in California is described by a Poisson distribution. If the average number of earthquakes per month is 0.9, what is the probability that there will be no earthquakes next month?

9. A particular soil density experiment, which can result in either a success or a failure, consists of 14 independent trials, each with a probability of success equal to 0.13. What is the exact probability of 2 successes in the experiment?

10. Judy isn't the greatest cook. In fact, the probability that a meal she cooks will be edible is 0.06. If she cooks 46 meals in a certain month, what is the probability that exactly 5 of them will be edible?

11. An engineering office has 7 women and 12 men working in it. If someone calls the office, any one of the employees may answer the phone with equal probability. If you call the office repeatedly, what is the probability that exactly 7 calls are needed till a man answers the phone?

12. A building has 13 above ground floors and 9 below ground floors. With his eyes closed, a bored student randomly picks a level in the elevator, waits until the floor is reached, and then randomly picks another. What is the expected number of random selections needed for the student to select 7 *below ground* floors?

13. When a water quality test is applied to a certain effluent stream, the effluent sample will pass with probability 0.8. What is the variance of passing samples if this test is performed on only 1 sample?

14. What is the variance of the random variable, X?

$$P(x) = \begin{matrix} 0.30 & \text{for x} = 1 \\ 0.18 & \text{for x} = 2 \\ 0.35 & \text{for x} = 3 \\ 0.17 & \text{for x} = 4 \end{matrix}$$

15. The number of contaminants passing through a detector in a reservoir during a minute can be described by a Poisson distribution. On average, 6 parts per million (ppm) contaminants pass through the detector at any given minute. What is the probability that *at least* 3 ppm will pass through the detector during a given minute?

16. The number of cars pulling into a parking garage every sixty minutes can be described as a Poisson process. If, on average, 5 cars enter the garage every ten minutes, what is the probability that *at most* 1 car will arrive in the next hour?

17. Deepwater, FL experiences a flood an average of 1.6 times a year. If the number of floods Deepwater experiences in a given year can be described by a Poisson process, what is the probability that there will be at least 2 floods next year?

18. A roulette wheel consists of 38 numbers: 1 through 36, and the single and double zero (0 & 00), each of which can occur with equal probability. If a player always bets that the outcome will be between 1 and 11 (she bets every number included in this range), what is the probability that the player will lose her first 15 bets in a row? A win is when the number the player bets on is selected by the wheel.

19. In performing a simple test of flow from a reservoir, a rubber duck is placed in the middle of the reservoir and its path charted. There are two (overflow) outlets from the reservoir, one north and one south, and the probability that a rubber duck will go through either outlet is equal (i.e. P(north) = P(south) = 0.5). If 10 ducks are placed in the middle of the reservoir, what is the probability that 8 of them flow out of the north outlet?

20. A government agency has 15 building inspectors and 9 home inspectors. 8 inspectors are randomly selected for testing. What is the expected number of building inspectors in the sample?

21. Suppose that a random variable, X, has an expected value of 8 and a variance of 1.8. What is the expected value of $-0.4 + 3X$?

22. When a water quality test is applied to a certain effluent stream, the effluent sample will pass with probability 0.86. What is the variance of passing samples if this test is performed on 5 samples?

23. A karate white belt thinks he is pretty cool and wants to show how strong he is by breaking bricks with swift hand chops. Despite his self-confidence, the poor sap breaks the brick only 12% of the time. His karate sensei, who happens to be a statistics whiz, has determined that his rate of success can be described by a binomial distribution, with 1 representing a break and 0 a failure. Now, using a Poisson distribution to approximate, what is the approximate probability of the white belt breaking two bricks if he is given 8 to break?

24. A roulette wheel consists of 38 numbers: 1 through 36, and the single and double zero - 0 & 00, all of which can occur with equal probability. If a player always bets that the outcome will be between 1 and 8 (he bets on all of the numbers in that range), what is the probability that the player's first win will occur on the 12th bet? A win is when the number the player bets on is selected by the wheel.

25. A certain plot of land is made up of two types of dirt: a clay and a sand, each of different densities. The plot is divided into 28 sections; 13 of these sections are all clay, while the other 15 are all sand. A random section of the land is selected and its density tested. Then another section is selected (this selection could include the one just tested) and its density checked. What is the probability that at least 5 tests are needed before a section with sand is selected?

26. At a state Department of Transportation, a pile of signs contains 11 stop signs and 12 yield signs (and no other signs). A bored government worker randomly selects a sign form the pile, notes its shape, and then replaces it. What is the probability that, after 5 random selections, only stop signs have been selected?

27. The Like Bunnies company produces batteries which will be defective 0.0071 percent of the time. The company sells the batteries in boxes of size 412, and has a policy that it will replace a box if it contains a single defective battery. What proportion of boxes should the company expect to replace?

28. A geotechnical engineer believes that almost all homes can be built with only a slab for a foundation. Thus, she believes that each home for which she designs the foundation can be built with a slab with probability 0.91. If she is designing foundations for a new neighborhood development, what is the variance of the number of home foundations she will design before she encounters a home for which a slab cannot be used?

29. A structural engineer tests concrete cylinders under a certain axial load; each cylinder has a probability, 0.21, of failing this test. If she decides to keep testing beams until she observes 3 failures, what is the standard deviation of the number of total cylinders she will test?

30. A roulette wheel consists of 38 numbers: 1 through 36, and the single and double zero - 0 & 00, each occurring with equal probability. If a player always bets that the outcome will be between 1 and 14 (i.e., she bets on every number in that range), what is the probability that the player's 5th *win* will occur on bet 8? A win is when the number the player bets on is selected by the wheel.

31. A child has a crushed velvet bag full of marbles. 7 of the balls are worthless commons, but 10 of them are pricey cat's eyes. If the child randomly pulls a marble from the bag, notes its color, and then returns the marble to the sack, what is the probability that exactly 14 draws are needed before the 7th cat's eye is drawn?

32. In a certain development, there are 5 two-story houses and 11 one-story houses. An inspector is to inspect all of them, but to keep things interesting, he randomly selects them as he goes along. Once he has inspected 9 houses, what is the probability that 7 of them were one-story? Assume that all houses have an equal chance of being selected.

33. Suppose that a random variable, X, has an expected value of 5.5 and a variance of 0.6. What is the variance of $Y = 1.4 + 5X$?

34. A construction contractor has a big box full of 71 nails, of which 38 nails are 4 inches long and 33 are 6 inches long. If a construction worker randomly takes 29 nails out of the box, what is the variance of the number of 4-inch-long nails the worker took?

35. The percentage of defective bolts used in a building is 0.2. If the building has 13,862 bolts total, what is the standard deviation of the number of defective bolts?

Module 2 Solutions:

1)	Continuous	2)	7/13	3)	2,160
4)	24/13	5)	0.904	6)	30
7)	2.63	8)	0.40657	9)	0.28917
10)	0.08432	11)	0.001579	12)	11.846
13)	0.16	14)	1.1779	15)	0.9380
16)	0.04043	17)	0.47507	18)	0.005939
19)	0.04394	20)	5	21)	23.6
22)	0.602	23)	0.17644	24)	0.01563
25)	0.04647	26)	0.02502	27)	0.9469
28)	0.1086	29)	7.33086	30)	0.05985
31)	0.0839	32)	0.28846	33)	15
34)	4.3284	35)	5.26009		

MODULE 3: CONTINUOUS DISTRIBUTIONS

1. If two random variables (say, X and Y) both have a standard deviation = 1, then the correlation of the two random variables is equal to their covariance. True or False?

2. A random variable X is normally distributed with a mean 8.6 and variance 5.8. What is the value of the probability density function of X when $x = 10.2$?

3. Suppose that the length of a geotechnical soil test, in minutes, can be modeled as an exponential random variable with parameter $\lambda = 0.29$. If a geotechnical engineer goes down to her company's lab to perform this test, what is the expected length of time she will spend performing the test?

4. The remaining number of hours a crane can be operated before it needs maintenance is exponentially distributed with an average value of 7,083 hours. How many more hours of operation should a project manager expect to get out of this crane before it will need maintenance?

5. The number of cars pulling into a particular driveway can be described as a Poisson process. On average, 11 cars arrive at this driveway each hour. For a transportation engineer analyzing traffic into the driveway, what is the expected time between entering cars in minutes?

6. At a certain construction site, there is a collection of 2x4 planks. If the length of the planks is normally distributed with mean 2.3 m and variance 1.2 m^2, what is the value of the probability density function of the planks when a specific plank is 1.6 meters long?

7. The alkalinity level of a lake (measured in hydroxide ion parts per million (OH^- ppm)) is found to be Gamma distributed with parameters $\alpha = 0.3$ and mean = 0.45 OH^- ppm. What is the variance of the alkalinity level of the lake?

8. For the past 50 years, the height, in meters, of a reservoir has been tabulated. It is found that, for a given year, the height has a normal distribution with mean 19.3 and standard deviation 5.8. What is the probability that, in a given year, the reservoir will have a height less than or equal to 31.6?

9. Suppose that the time it takes to pour a certain section of concrete, in minutes, can be modeled as an exponential random variable with parameter $\lambda = 0.11$. A firm needs to create a cost estimate for a project and needs to do a statistical analysis of concrete pours. Given the above information, what is the variance of the length of time it takes to pour the section of concrete?

10. The number of people entering ECJ can be described as a Poisson process. On average, 6 people enter the building each minute. Based on this information, what is the parameter, □, to describe the exponential distribution governing gap time length between entering persons, in *seconds*?

11. The proportion of a particular airborne contaminate in a sample of the atmosphere (in parts per million (ppm)) has a gamma distribution with parameters $\alpha = 3.8$ and $\beta = 0.8$. What is the average amount of the contaminate (in ppm) that will be in a random atmospheric sample?

12. The number of birds that crash into a particular high-rise's windows on a given day is found to have a Weibull distribution with parameters $\alpha = 2$ and $\beta = 2.3$. What is the average number of birds that hit the building per day?

13. The mortar thickness (in cm) between two randomly selected bricks in a brick wall is found to be lognormally distributed with parameters $\mu= 0.5$ and $\sigma = 1.5$. What is the value of the probability density function when the mortar thickness between two randomly selected bricks is 0.8?

14. A series of yearly rainfall amounts for Deerwood, MI over the last 50 years has been collected and found to have a normal distribution with mean 30.3 (in inches) and variance 14.9. What is the value of the probability density function for the rainfall when the total rainfall for a given year is 27.5 inches?

15. Suppose that the length of a geotechnical soil test, in minutes, can be modeled as an exponential random variable with parameter $\lambda = 0.12$. If a geotechnical engineer goes down to her company's lab to perform this test, and notices another engineer just started using the test equipment, what is the probability that the first engineer's wait is more than 5 minutes?

16. The number of cars pulling into a particular driveway can be described as a Poisson process. On average, 7 cars arrive at this driveway each hour. As the transportation engineer in charge of analyzing the driveway noted another entering car, he stopped working to take a phone call from the office. What is the probability that he did not miss any more entering cars if the call lasted exactly 5 minutes?

17. The error in the measurement that a soil moisture gauge returns has a standard normal distribution (measurements are in g/cm^3). What is the probability that the measurement returned by the gauge will be under the true moisture level by at least 0.7 g/cm^3?

18. In order to make freshly poured concrete more workable on a job site, a plasticizer compound is used. If a certain fixed amount is added to a given batch of concrete, it is found that the amount of plasticizer, in grams, in a given kg sample of concrete has a gamma distribution with parameters $\alpha = 5.4$ and $\beta = 4.1$. What is the variance of the distribution of plasticizer in a one kg sample of concrete?

19. The error in the measurement that a soil moisture gauge returns has a standard normal distribution (measurements are in g/cm^3). What is the probability that the measurement returned by the gauge will be under the true moisture level by at least 0.8 g/cm^3?

20. The length of a randomly selected tube from a pile of aluminum tubes in a junkyard is found to be lognormally distributed with parameters $\mu = 4.4$ and $\sigma = 0.5$. What is the average length of the pipes in the pile?

21. The number of houses in a given 0.5 km^2 section of a city is found to have a Weibull distribution with parameters $\alpha = 1.8$ and $\beta = 2.6$. What is the probability that a random 0.5 km^2 section of the city has 2 houses or less in it?

22. In California, some buildings are built on a bed of large springs so as to minimize the damage caused by earthquakes. At any given time for a certain building, the deviation of a randomly selected spring (from its rest point in mm) is found to be lognormally distributed with parameters $\mu = 3.4$ and $\sigma = 0.4$. What is the standard deviation, in mm, of this distribution?

24. Suppose that the length of time, in minutes, that it takes a certain bricklayer to finish a section in a building can be estimated as an exponential random variable with parameter $\lambda = 0.32$. If a supervisor wants to see the bricklayer finish a section in 15 to 20 minutes, what is the probability of this happening? That is, what is the probability that the bricklayer finishes a section within 15 and 20 minutes?

25. The remaining number of hours an air-conditioning system can be used before it needs maintenance is exponentially distributed with an average value of 2,008 hours. If it is expected that over the next 5 weeks, the system will be operated day and night continuously (it is going to be very hot), what is the probability that it will need maintenance during the job?

26. The thickness of a road is normally distributed with standard deviation 0.42 (in^2). If the probability that the road thickness is greater than or equal to 3.8 is 0.683031, what is the average of the road thickness?

27. If the volume of a randomly selected room in a house in Burbank, CA is normally distributed with mean 1.5 m^3 and the value of the probability density function of the distribution when a given room's volume is 1.2 m^3 is 0.712326, what is the variance of the distribution?

28. In a given cord of cut wood, the number of knots has a gamma distribution with parameter $\alpha = 5$ and variance = 5. What is the parameter β of the distribution?

29. The average width of an electrical conduit (in inches) in a randomly chosen house from the town of Tesla is found to have a Weibull distribution with parameters $\alpha = 1.8$ and $\beta = 0.7$. What is the probability that a randomly chosen house will have an electrical conduit that is less than or equal to 1.6 in?

30. The number of workers taking a smoke break at any given time on a construction site is found to be lognormally distributed with parameter $\mu = 1.3$. If the average number of workers taking a smoke break at any given time is 4.687972, what is the parameter σ of the distribution?

31. The remaining number of hours a crane can be operated before it needs maintenance is exponentially distributed with an average value of 1,623 hours. The project manager needs this crane to work on a big job for 5 straight weeks; but, more importantly, he has scheduled it for a critical activity for exactly one week, starting exactly 24 hours into the project. What is the probability that this piece of equipment will need maintenance during this critical activity, thus delaying the entire 5 week job?

32. In a certain housing development, the average volume of concrete used in a house's foundation is 206 $yards^3$. If the probability that a foundation has less than or equal to 100 $yards^3$ is 0.076, and the distribution of the volumes is normal, what is the standard deviation?

33. The thickness of paint (in mm) on a random 1 mm^2 section of the outside of a military hanger has a gamma distribution with parameters $\alpha = 1.5$ and $\beta = 1.1$. What is the probability that a random 1 mm^2 section has a thickness of 0.6 mm?

34. On a random weekday, the level of service of a particular bus route (measured in buses/hour) for mid-day times is found to have a Weibull distribution. If the parameter $\alpha = 1.7$ and the probability density function for distribution when the level of service is 2 buses/hour is 0.5, what is the parameter β?

35. For a certain engineering firm, the percentage change in profits between any two randomly chosen sequential years is found to be lognormally distributed with parameter $\sigma = 0.5$. If the standard deviation of this percentage is 0.331, what is the parameter μ of the distribution?

36. For a given square mile in Dewey, LA, the percentage of wood infested with termites (this includes both trees and houses) is found to be lognormally distributed with parameters $\mu = 0.4$ and $\sigma = 0.5$. Given this information, what is the probability that at least 0.3 percent of the wood in a randomly selected square mile in Dewey is infested with termites?

37. The number of hours per day spent goofing off at work by a certain engineer is found to have a beta distribution with parameters $\alpha = 1.1$, $\beta = 6.4$, and $B = 9.4$. If the average time the engineer goofs off per day of work is 2.147 hours, what is the parameter A of the distribution?

1)	True	2)	0.1328	3)	3.4483
4)	7,083	5)	5.4545	6)	0.2969
7)	0.675	8)	0.9830	9)	82.64
10)	0.1	11)	3.04	12)	2.0383
13)	0.2960	14)	0.07944	15)	0.5488
16)	0.5580	17)	0.2420	18)	90.774
19)	0.2118	20)	92.30	21)	0.3008
22)	13.521			24)	0.006568
25)	0.3418	26)	4	27)	0.2
28)	1	29)	0.9881	30)	0.7
31)	0.0969	32)	74	33)	0.43909
34)	2.4812	35)	-0.6	36)	6.684E-4
37)	0.9				

MODULE 4: CONFIDENCE INTERVALS

1. True or False: The variance of a parameter estimate falls as the sample size grows large.

2. The following numbers are a random sample from a normal distribution with a variance = 4.6. What is the lower end (left hand side) of the two-sided 95% confidence interval about the sample mean?

7.5	15.8	12.4
21.2	13.6	19.1
10	4	7.2
9.2	8.8	22.8
22.3	11.3	22.7

3. If a random sample of 251 numbers has a sample mean of -10.9 and a sample variance of 4.5, what is the upper level (right hand side) of a two-sided 95% confidence interval about the sample mean?

4. The following numbers are a random sample from a normal distribution with a variance = 7.1. What is the level of a one-sided *lower* 97.5% confidence interval about the sample mean?

7.9	7.5	12.6
5.7	8.4	15.1
25.1	14.3	2.3
0.2	20.7	17.1
9.5	7.6	17.5

5. Yes or No, say you have a sample from a normal distribution with known variance. If you know the width of a two-sided 95% confidence interval about a sample mean, as well as the size of the sample, can you determine the sample mean?

6. If a random sample of 31 numbers from a normal distribution has a sample mean of 34.4 and sample variance of 4.8, what is the upper level (right hand side) of a two-sided 90% confidence interval about the sample mean?

7. If the following numbers represent random draws from a normal distribution, what is the level of a one-sided *upper* 95% confidence interval about the sample mean?

5.3	17.5	-2	8
14.6	1.1	6.8	-1.1
16.7	2.2	-13	-8.3
2.5	7.5	-18.9	-20
-6.2	-5.9	-11.7	-7.6

8. Shown below are measurements of deviations, above or below, of a reservoir height (in feet) from its average level. Such measurements are determined to follow a normal distribution with a variance = 5.1. If you know that the level of the reservoir will, with 95% confidence, be within a region around the mean, what is the upper level of that region? That is, what is the lower end (left hand side) of the two-sided 95% confidence interval about the sample mean?

17.8	-2.6	-6.6
4.6	-10.8	8.7
26.8	8.1	-17.8
-18.3	-11.5	20
19.4	15.2	-1

9. A random sample of 277 adults is selected and their heights measured. If the sample has a mean of 5 (feet) and a sample variance of 1.8 ($feet^2$), what is the upper level (right hand side) of a two-sided 95% confidence interval about the sample mean?

10. The numbers below are forest service measurements of the distance above or below the tree line that snow has fallen this year. If the true distribution of these distances follows a normal distribution with variance = 4.8, what is level of a one-sided *lower* 97.5% confidence interval about the sample mean?

-1.3	27	1.3
15.1	2.7	20
11.2	26	22.2
21	19.5	-1
1.4	10.1	15.9

11. A random sample of 133 numbers is taken from an unknown distribution. If the sample mean is -6.8 and $E[X^2]$ is 147.8, what is the lower level (left hand side) of a two-sided 95% confidence interval about the sample mean?

12. If a random sample of 42 numbers from a normal distribution has a sample mean of 8.2 and sample variance of 3.8, what is the *width* of a two-sided 90% confidence interval about the sample mean?

13. If the following data is randomly drawn from a normal distribution, what is the *width* of a two-sided 95% confidence interval about the sample mean?

-17.5	10.4	-14	16.1
10.4	-1.9	-0.9	-4.6
-10.5	3.3	3.2	15.9
19.7	13.8	2.2	8.4
0	-18	-4.5	7.7

14. A random sample is selected from a normal distribution with variance 46.12336. If the width of a 95% confidence interval about the sample mean is 4.5, what is the size of the sample?

15. A transportation safety company is examining the destructive effects of car accidents. From a random sample of accidents, car speeds are found to be normally distributed. However, the destructive power of an accident is proportional to the square of the velocity. The distribution of such a sample of squared velocities is thus chi-squared. The company is going to apply a method of moments technique to the data. So, if the following measurements represent square velocities (in ft^2/sec^2), what is $E[X^3]$?

19.8	2.1	12.2	20
11.7	2.1	16.5	6.6
8.7	14.2	3	18.1
11.8	7.2	18.2	4
12.8	18.8	11.7	13.2

16. A random sample of 8 measurements of concrete strength is found to follow a normal distribution with a sample mean of 14.2 (kpsi) and sample variance of 7.6 ($kpsi^2$), what is the upper level (right hand side) of a two-sided 90% confidence interval about the sample mean?

17. A sample of farms is visited by a USDA inspector and the number of chickens not yet out of shell counted. If the following data represents these counts, and it is assumed to follow a normal distribution, with 95% confidence what is the largest number of chickens counted before they

hatched we would expect to see? That is, what is the level of a one-sided *upper* 95% confidence interval about the sample mean?

3.8	16.5	1.4	0.1
4	20	14.6	10.9
5.1	13.3	13.2	12.1
0.3	3.7	11.7	6.7
6	1.9	11.5	15.7

18. The following is a random sample drawn from a gamma distribution. What is your best estimate of the parameter α?

1.2	11.9	27.2
10.7	10.2	28.6
30.4	11	25.7
2.1	24.2	9.8
26.2	2.3	1.1

19. The following data is randomly drawn from an exponential distribution. What is your best estimate of the parameter lambda?

8	12.3	8.1	12.6
9.8	0.7	10.6	11.5
0	8	11.7	11.1
12.4	0.5	9.6	1
12.6	15.1	15.3	8.8

20. If the following data is randomly drawn from a normal distribution, what is the upper level (right hand side) of a two-sided 95% confidence interval about the sample mean?

18.1	-3.4	-11.4	19.6
6.5	10	11.4	-15.4
-15.3	4.4	0	13.8
-10.9	9.2	-19.6	12.4
13.2	4.5	-16.8	-15.3

21. A random sample of 211 numbers is taken from an unknown distribution. If the sample mean is -1.3 and $E[X^2]$ is 100.8, what is the level of a one-sided *upper* 95% confidence interval about the sample mean?

22. A series of 32 numbers have been randomly drawn from a normal distribution with variance 4.4. If the upper level (right hand side) of a 95% confidence interval about the sample mean is 12.6, what is the sample mean?

23. The number of windows on the facades of a random sample of 10 buildings follows a normal distribution with a sample mean of 16.8 and sample variance of 4.1, what is the *width* of a two-sided 90% confidence interval about the sample mean?

24. A FEMA analysis has shown that the average length of time between two flood-causing storms in the town of Lamotta follows a gamma distribution. If the following numbers represent the number of years between flood storms, find an estimate of the parameter α.

1.6	4.4	0.7
3.1	3	2.9
4	1.5	3.2
1.5	4.1	4.6
0.6	2.8	4.7

25. Shown below are measurements of the average time (in hours) a bus passes a given bus stop. If these measurements are assumed to be from an exponential distribution, find the best estimate of the parameter λ.

0.3	1.5	0.4	0.6
1	0.6	0.5	0.3
0.4	1.2	0.1	0.5
1	1	1.3	0.2
0.8	0.5	0.5	0.3

26. If the following data represent an estimate of the number of years left before a random sample of roads needs to be replaced. If the estimate is assumed to follow a normal distribution, what is the upper level (right hand side) of a two-sided 95% confidence interval about the sample mean? That is, what is, with 95% confidence, what is the longest we can expect our roads to last?

11.5	18.2	19.5	7.2
11.6	6.7	6	9.5
19.3	12.3	3.2	2.1
11.6	3.4	8.5	5.8
18	9.5	13.7	4.6

27. At 144 random times, the height (in feet) that an elevator is above or below the ground level of a building is measured. If the sample mean is -4.6 and $E[X^2]$ is 226.7, what is the size of a one-sided *upper* 95% confidence interval about the sample mean?

28. The following data represent the number of trailer homes destroyed from randomly selected tornado occurrences. If these values are assumed to be normally distributed (assume they are continuous), what is the *width* of a two-sided 95% confidence interval about the sample mean?

29	10	47	34
38	22	11	16
39	17	44	34
20	30	26	48
19	18	15	4

29. The number of rooms in a random sample of houses is normally distributed with a variance 1.7. If the upper level (right hand side) of a 95% confidence interval about the sample mean is 15.1, what is the sample mean?

30. A random sample of tide levels is selected from a normal distribution with variance 15.147334. If the width of a 95% confidence interval about the sample mean is 2.3, what is the size of the sample?

1)	True	2)	12.77	3)	-10.638
4)	10.085	5)	No	6)	35.045
7)	3.597	8)	2.3238	9)	5.158
10)	11.631	11)	-8.519	12)	0.9866
13)	10.45	14)	35	15)	2,710.4
16)	15.80	17)	10.96	18)	1.9424
19)	0.1113	20)	6.868	21)	-0.1699
22)	11.873	23)	2.100	24)	4.5225
25)	1.5386	26)	12.67	27)	-2.628
28)	12.046	29)	14.514	30)	44

MODULE 5: Hypothesis Tests

1. You conduct a hypothesis test and end up rejecting the null hypothesis. Later, you find out that H_0 was actually true. What type of error did you commit? Typographic, Human, Type I, or Type II?

2. A sample of 20 numbers is drawn from a normal distribution with variance 2.3. If the sample mean is 16.2, is it possible to say, with 95% confidence, that the actual mean is 16.5 (the the alternative being the mean is ≠ 16.5)? Yes or No?

3. Given that the following numbers are sampled from a normal distribution with variance 2.1, is it possible to say, with 95% confidence, that the actual mean is 15.005 (the the alternative being the mean is ≠ 15.005)? Yes or No?

1.4	6	17.9	14.4	Sample Mean: 14.405
16.6	22.9	13.8	3.3	
21.1	6.1	20.8	14.5	
23.4	1.9	20.4	10.8	
24.9	10.6	24.1	13.2	

4. A sample of 238 numbers is drawn from an unknown distribution. The sample mean is 57.4 and the sample variance is 1.3. With 90% confidence (use a 2-tailed test), what is the *largest* value the true mean of the underlying distribution could be?

5. 8 numbers are randomly sampled from a normal distribution. The *sample* has a mean of 4.4 and a variance of 3.9. With 90% confidence, can you accept the hypothesis that the true mean of the underlying distribution is 3.6 (with the alternative hypothesis being that the true mean is ≠ 3.6)? Yes or No?

6. The following numbers have been randomly sampled from a normal distribution. With 90% confidence, can you accept the hypothesis that the true mean of the underlying distribution is -8.553 (with the alternative hypothesis being that the true mean is ≠ -8.553)? Yes or No?

-7.6	-3.3	-6.6	Sample Mean: -5.353
-5.3	-7	-4.4	Sample SD: 1.756
-6.9	-4.4	-3	
-2.1	-6.6	-7.3	
-5.4	-6.6	-3.8	

7. A sample of 51 numbers from a normal distribution is made. The sample mean is 63.8 and the width of a two-sided 95% confidence interval about the sample mean is 4.9. With 95% confidence, what is the largest value that the actual mean of the underlying distribution could take (use a two-tailed test)?

8. A sample of 15 numbers is selected from a normal distribution. The sample mean is 3.9 and the sample variance is 0.4. With 90% confidence, what is the *minimum* value that the true mean of the underlying distribution could take? Use a two-tailed test.

9. The following data has been sampled from a normal distribution with variance 117.483. What is the *p*-value of a two-tailed test that the true mean of the underlying distribution is 25.485?

24.6	32.6	26.6	4.6
26.3	37.4	12.2	36.3
37.5	19.8	25.3	41.4
11	3.4	25.5	34.4

23.8	22.1	16.5	18.4	Sample Mean: 23.985

10. A random sample of 32 numbers has been collected from a normal distribution whose variance is 22.2. If the mean of the sample is 31.4, what is the *p*-value of a two-tailed test that the true mean of the underlying distribution is 32?

11. You have been asked to investigate a series of budget overruns on public construction projects. Because of bad records, you can only analyze a sample of 19 of the projects for fiscal responsibility. For this sample, you record the amount that each project was over (or under) budget and determine the sample mean to be 4.8 (in millions of dollars) with variance 2.6 (in millions of dollars^2). You feel that this value of the mean is under-representing the true average. If you assume the true distribution of the budget overruns is normally distributed, then, with 95% confidence, is it possible to say that the actual mean is 6.2 million dollars (the the alternative being the mean is $\neq$ 6.2)? Yes or No?

12. You are asked to investigate the effects of rain on building foundations. As part of your analysis, you wish to get a good understanding of how much rain falls in an average year. However, the National Weather Service lost most of its records for your area because a roof leaked in their headquarters, so you can only use a sample in your investigation. The following data represent rainfall amounts (in inches) for a random sampling of years for your area. You think that the actual mean is higher than what this sample represents. If you assume that they come from a normal distribution with variance 5.4, is it possible to say, with 95% confidence, that the actual mean is 13.85 inches (the the alternative being the mean is $\neq$ 13.85)? Yes or No?

8.2	9.2	24.9	16.8	
9	7.4	8	14.1	Sample Mean: 12.55
25	0.4	3.4	13.6	
14.7	14.9	6.9	21	
21.9	12.5	5.7	13.4	

13. A house inspector is developing a report for the state board. In the last year, he has inspected 170 houses and found, on average, 59.2 problems per house (with a variance of 2.1). He would, in order to cover himself, like to give the board as pessimistic a report as possible. So, with 90% confidence (use a 2-tailed test), what is the *largest* value the true mean of how many problems per house could take?

14. The thickness of handmade windows is being investigated by a construction firm interested in using them, but concerned that they may be difficult to use in standardized openings. 16 of the windows are randomly selected and their thickness measured. The *sample* has a mean of 6 mm and a variance of 1.5 mm^2. With 90% confidence, can you accept the hypothesis that the true mean of the underlying distribution is 5.7 mm(with the alternative hypothesis being that the true mean is $\neq$ 5.7)? Yes or No?

15. The data below represent measurements of a toxic substance (in parts per million) from random spots in a lake. The company which manages the lake would like, for insurance reasons, to have the official average level of the substance to be 2.487. Can you, with 90% confidence, accept this hypothesis (with the alternative hypothesis being that the true mean is not 2.48)? Yes or No?

11.5	5.6	8.4	
1.9	4.7	1	Sample Mean: 6.0867
8	7	3.2	Sample SD: 3.1145
10.9	8.3	4.6	
3.1	5.3	7.8	

16. A sample of 51 numbers is drawn from a normal distribution. The sample mean is 42.9. It is found, with 95% confidence, that the actual value of the mean is 45.4 (as opposed to the sample mean $\neq$ 45.4). Given this, what is the *smallest* possible value of the variance of the underlying distribution?

17. A sample of 197 numbers from an unknown distribution is made. The sample mean is 75.6 and the width of a two-sided 95% confidence interval about the sample mean is 4.2. With 90% confidence, what is the largest value that the actual mean of the underlying distribution could take (use a two-tailed test)?

18. The following sample has been drawn from a normal distribution. With 95% confidence, what is the *minimum* value that the true mean of the underlying distribution could take? Use a two-tailed test.

47.7	65.5	29.2	54.8	
59.9	63.7	23.3	32.7	Sample Mean: 45.845
62.8	21.7	57.6	54	Sample SD: 15.070
54.8	28	23.1	51.4	
47.2	32.8	47.7	59	

19. The following data has been sampled from a normal distribution with variance 89.725158. What is the p-value of a *lower-tailed* test that the true mean of the underlying distribution is 26.59?

38.2	14.8	38.8	27.4	
19.6	13.1	29.2	31.6	Sample Mean 27.790
32.4	42.4	39.8	24.8	Sample SD: 9.4723
26	7.8	36.1	20.5	
30.9	25.1	34.8	22.5	

20. A sample of 17 numbers is drawn from a normal distribution with variance 2.5. If the sample mean is 9.8, is it possible to say, with 95% confidence, that the actual mean is 11.3 (the the alternative being the mean is less than 11.3)? Yes or No?

21. An investigation into the possible effects of global warming on lake levels is being conducted. A sample of 148 lakes are selected and their levels measured. A worst case scenario analysis is going to be conducted. The sample mean is 49.6 (in meters) and the width of a two-sided 95% confidence interval about the sample mean is 3.9 meters. With 95% confidence, what is the largest value that the actual mean of the underlying distribution could take (use a two-tailed test)?

22. The variation of a reservoir over time is being checked. The reservoir level is checked at 11 random times over a year, and the deviation from the average height checked. These deviations are assumed to be normally distributed, and if the average deviation is 1.3 feet and the variance 0.7 feet^2. Using a 90% confidence level, what is the lowest value that the true mean deviation could take? That is, with 90% confidence, what is the *minimum* value that the true mean of the underlying distribution could take? Use a two-tailed test.

23. A firm is compacting ground for placement of a highway. The following measurements are compaction values (densities) taken at random points around the site. They would like to see if the true value of the average density is 25.835, which is what is specified by their contract. If it assumed that these values follow a normal distribution with variance 132.655026, what is the *p*-value of a two-tailed test that the true mean of the underlying distribution is 25.835?

41.7	14.6	42.6	6.4	
24.4	32.5	12.4	15.9	Sample Mean: 24.835
39.9	10.4	16.8	26.2	Sample SD: 11.518
18	31	24.7	42.2	
35.5	11.6	26.6	23.3	

24. The following numbers have been drawn from a normal distribution. It is found, with 95% confidence, that the actual value of the sample mean is 23.313 (as opposed to the sample mean ≠ 23.313). Given this information, what is the *smallest* possible value of the variance of the underlying distribution? (Hint: Large sample interval, but solve for S^2)

36.9	33	40.4	Sample Mean: 22.813
15.2	14.6	7.5	
17.5	33.2	19.5	
12.5	34.9	3.7	
27	3.4	42.9	

25. A sample of numbers is randomly selected from a distribution with a sample mean of 34.164449 and the sample variance of 5.2. If 34.48785 is found, with 95% confidence (two-tailed test), to be the actual value of the mean of the underlying distribution, what is the *largest* size the sample could be? (Assume the sample size is large.)

26. The following data have been sampled from a normal distribution with variance 108.254737. What is the *p*-value of a *lower-tailed* test that the true mean of the underlying distribution is -13.6?

-23	-26.4	-7.3	-8.7	
-32.4	-8	-13.6	-5.7	Sample Mean: -15.000
-1	-21	-17.1	-31.4	Sample SD: 10.405
-14.2	-7	1.6	-23.8	
-14.2	0.2	-27	-20	

27. A random sample of 21 numbers has been collected from a normal distribution whose variance is 12.3. If the mean of the sample is 13, what is the *p*-value of a *lower-tailed* test that the true mean of the underlying distribution is 13.7?

28. Given that the following numbers are sampled from a normal distribution with variance 3.8, is it possible to say, with 95% confidence, that the actual mean is 11.79 (the the alternative being the mean is less than 11.79)? Yes or No?

6.3	5.3	9.5	2	
1.5	20.9	15.6	17.4	Sample Mean: 11.190
2.2	11	7.1	7.3	Sample SD: 6.4466
14.5	17.6	14.7	8.7	
14.9	8.7	13.7	24.9	

29. A sample of 8 numbers is drawn from a normal distribution with variance 1.8. If the sample mean is 9.3, is it possible to say, with 95% confidence, that the actual mean is 8.3 (the the alternative being the mean is greater than 8.3)? Yes or No?

30. The sewer system in the town of Dregs is having problems. It keeps getting clogged up whenever it rains and cutting off service to some of its customers. Through a series of investigations, the officials of Dregs have determined that the number of homes cut off each time it rains follows a normal distribution. Through their statistical analysis, they have also shown that, for a sample of 15 rain events, though the sample average number of home cut off 17.3, with 95% confidence they believe the true average of the underlying distribution is 20.7. Because this situation is potentially embarrassing, the officials want to put the best spin on it as possible. So, given these results, what is the smallest value that the variance of the true distribution of the number houses that get cut off when it rains can take?

31. As part of an investigation into building practices at the first half of the 20th century, you are looking at the average height of foundations of old buildings. A sample of 193 buildings has a average foundation height of 65.4 centimeters, and the width of a two-sided 95% confidence interval about the sample mean is 4.1 centimeters. With 90% confidence, what is the largest value that the actual mean of the underlying distribution could take (use a two-tailed test)?

32. A firm is creating a cost estimate for a bid on a construction contract. The project will require piling in order to put the building on a firm foundation. To create an estimate for piling costs, the average pile depth must be determined. A sample of drill cores from random points on the site were collected and the following depths necessary for pile recorded. Since the company wishes to keep its bid as low as possible, they would like to create an average as low as statistically possible. If the data is assumed to follow a normal distribution, what is the *minimum* value that the true mean of the underlying distribution could take? Use a 95% interval and a two-tailed test.

33.7	62	17.7	32.1	
52	41.4	38.8	37.4	Sample Mean: 41.590
26.7	37.1	29.1	62.1	Sample SD: 13.250
26.7	64.7	43.3	45.4	
42	38.3	62.8	38.5	

33. The average wind speed during storms has been shown to follow a normal distribution with variance 132.202. Below are measurements of average wind speeds (mph) from a random sample of storms. Models have been developed assuming the true mean is 21.4 mph, and a p-test is to be carried out to test the hypothesis that this is the true mean. So, what is the p-value of a *lower-tailed* test that the true mean of the underlying distribution is 21.4 mph?

23.2	4.6	18.3	22.9	
7.1	8.5	32	29	Sample Mean: 22.700
4.5	36.6	17.8	26.4	Sample SD: 11.498
34.4	38.6	40.6	17.2	
28.9	15.3	33.9	14.2	

34. A study measuring energy efficient buildings is counting the number of skylights in buildings. A sample of 11 buildings is chosen and the number of skylights counted. It is assumed that the underlying distribution governing the number of skylights in the building is normal with variance 5.1. If the sample mean is 22 skylights, is it possible to say, with 95% confidence, that the actual mean is 22.5 (the alternative being the mean is less than 22.5)? That is, is it possible that there are 0.5 more skylights, on average, in buildings than what the sample represents? Yes or No?

35. Shown below are 20 measurements of the average time (in hours) a bus passes a given bus stop. If these measurements are assumed to be from an exponential distribution, find the maximum likelihood estimate for the parameter α.

0.8	1.6	0	0.8	
1.4	1.4	0.7	1.6	Sample Mean: 1.1050
0.5	0.9	1.5	1.2	Sample SD: 0.46620
0.8	1.6	1.6	1.5	
1.4	0.5	0.9	1.4	

36. If the following 20 values represent estimates of the number of years left before a random sample of roads needs to be replaced. If the estimate is assumed to follow a normal distribution, what is the upper level (right hand side) of a two-sided 95% confidence interval about the sample mean? That is, what is, with 95% confidence, what is the longest we can expect our roads to last? (Hint: t Distribution)

16	17.4	19.9	10.4	
11.6	12.6	5.5	18.4	Sample Mean: 11.770
8.2	12.8	11.4	13.8	Sample SD: 5.3364
4.8	19.9	6.5	13	
16.2	3.2	2.6	11.2	

37. At 213 random times, the height (in feet) that an elevator is above or below the ground level of a building is measured. If the sample mean is 8.3 and the 2nd sample moment (our estimate of $E(X^2)$ using the sample data) is 263.8, what is the size of a one-sided *upper* 95% confidence interval about the sample mean?

1)	Type I Error	2)	Yes	3)	Yes
4)	57.5	5)	Yes	6)	No
7)	66.25	8)	3.612	9)	0.536
10)	0.471	11)	No	12)	No
13)	59.4	14)	Yes	15)	No
16)	83.0	17)	77.4	18)	38.8
19)	0.286	20)	No	21)	51.55
22)	0.8429	23)	0.698	24)	0.976
25)	191	26)	0.274	27)	0.180
28)	Yes	29)	No	30)	45.1
31)	67.1	32)	35.4	33)	0.307
34)	Yes	35)	0.905	36)	14.3
37)	9.88				

MODULE 6: SIMPLE LINEAR REGRESSION

1. When conducting a t-test for a parameter estimated by simple linear regression using a formula which consists of the parameter estimate divided by its standard deviation, you are testing whether or not that parameter is equal to:

   a) Its actual value
   b) Anything but the value you estimated
   c) Zero
   d) One

2. A civil engineering firm is trying to relate the cost of a construction job (Y, measured in millions of dollars) to the number of people working on the job (X). Given the following data, in which each i represents one job from a database of jobs the firm has worked on, what is the slope of the regression line relating these two variables?

$\sum_i X_i = 113$ $\sum_i Y_i = 51$
$\sum_i X_i Y_i = 393$ $\sum_i X_i^2 = 769$
no. observations = 41

3. An analysis is being conducted comparing the height of a building (X, in meters) of a building and the age of the building (Y, in years). Given the data provided below, what is the least squares estimate of the β_0 coefficient?

$\sum_i X_i = 167$ $\sum_i Y_i = 49$
$\sum_i X_i Y_i = 366$ $\sum_i X_i^2 = 425$
no. observations = 62

4. One would expect that the longer the days (in terms of daylight), the faster a construction project will be done. An analysis concerning this idea yields the data below, where X represents the length of the day (in hours) and Y gives the length of time to complete the project (in days). Given the data provided below, what is the least squares estimate of the $\beta_1$coefficient?

$\sum_i X_i = 118$ $\sum_i Y_i = 953$
$\sum_i X_i Y_i = 2{,}853$ $\sum_i X_i^2 = 610$
no. observations = 26

5. 38 reservoirs were analyzed to determine how much water is lost per year to evaporation effects. The average temperature on the day a reservoir was investigated was recorded (X, in ^{O}C) as well as a measurement of how much water was lost to evaporation (X, in hundreds of gallons). Using data given below, which includes the estimation of the intercept of the least squares regression line, what is the slope (β_1) of the least squares line?

$\sum_i X_i = 1{,}063$ $\sum_i Y_i = 172$
$\sum_i X_i Y_i = 12{,}869$ $\sum_i X_i^2 = 616{,}644$
no. observations = 38

6. Assume you are creating a least squares regression with X as the independent variable and Y as the dependent variable. Given the data provided below, what is the least squares estimate of the β_0 coefficient?

$\sum_i X_i = 1{,}784$ $\sum_i Y_i = 54$
$\sum_i X_i Y_i = 698$ $\sum_i X_i^2 = 35{,}609$

no. observations = 32

7. A least squares regression is run comparing the age of houses with the "level-ness" of their house (measured in deviations from flat). From the regression, the sum of squared totals (SST) is found to be 1,633, while the sum of squared errors (SSE) is found to be 1,362. What is the R^2 of this analysis?

8. After working with a global positioning system (GPS) device which occasionally returns large errors, you think that you understand the problem - high humidity causes the device to work improperly. You then collect data on the error in the GPS device's returns (X, in meters) and the level of humidity on a given day (Y, in parts per million water). Given the data given below, which includes the estimation of the slope of the least squares regression line, what is the intercept (β_0) of the least squares line?

$\sum_i X_i = 172$ $\sum_i Y_i = 813$

$\beta_1 = 1.3376$

no. observations = 36

9. A least squares regression is run comparing the age of houses with the "level-ness" of their house (measured in deviations from flat). From the regression, the sum of squared totals (SST) is found to be 4,064, while the R^2 value is found to be 0.667077. What is the sum of squared errors (SSE) of this analysis?

10. A regression analysis is run comparing the power generated by a hydroelectric dam in a year and the average temperature in that same year. The results give a β_1 estimate of -0.6 and the estimate of that parameter's standard deviation is 1.25. Given this, running a t-test is one claim that the true value of the β_1 coefficient is, with 95% confidence, equal to 0 (use a two-tailed test), Yes or No?

11. A government agency is performing a simple regression analysis comparing building age with depreciation due to wear. The below data shows some of the results of this analysis, with X representing the building age and Y representing the cost, in millions, of depreciation. Given the data, what is the least squares estimate of the β_0 coefficient?

$\sum_i X_i = 254{,}019$ $\sum_i Y_i = 80{,}948$

$\sum_i X_i Y_i = 7{,}870{,}016$ $\sum_i X_i^2 = 24{,}056{,}945$

$\sum_i Y_i^2 = 2{,}608{,}221$

no. observations = 3,099

12. In a regression analysis comparing the total cost of accidents over the lifetime of an engineer (X, in thousands of dollars) and the age of the engineer (Y), the following results are recorded (the β_o and β_1 values are estimates). Given this information, what is the value of the sum of squared errors (SSE) of the anlysis?

β_o= -6.2189 β_1= 0.39285

$\sum_i X_i Y_i = 7{,}870{,}659$ $\sum_i Y_i = 80{,}901$

$\sum_i Y_i^2 = 2{,}608{,}877$

13. In a regression analysis comparing the thickness of insulation of a building (*X*, in mm) and the temperature inside the building (*Y*, in ^{o}C), the following results are recorded (the β_o and β_1 values are estimates). Given this information, what is the value of the total sum of squares (SST) of the analysis?

β_o= -5.251655 β_1= 0.38261

$\sum_i X_i Y_i = 7{,}870{,}300$ $\sum_i Y_i = 80{,}905$

$\sum_i Y_i^2 = 2{,}608{,}285$

no. observations = 3,100

14. It is suspected that the closer a certain type of factory is to a river or lake, the higher the lead levels in the sediment of the river or lakebed. Data are collected on the level of lead in the sediment (*Y*, in parts per billion (ppb)) and the distance from a factory that the lake is situated (*X*, miles). Given the data below, which includes the estimation of the slope and the intercept of the least squares regression line, what is the average level of lead (in ppb) in the sediment?

$\bar{X}$= 6.96154 β_o= 8.41053

β_1= 3.25042

no. observations = 26

15. A regression analysis is run comparing the lateral stress on a bridge with the average wind speeds recorded in the area. The results give a β_1 estimate of 0.6 and the estimate of that parameter's standard deviation is 1.65. Given this information, what is the t-statistic for the test that the true β_1 coefficient is 0 ?

16. The results of a certain regression analysis (comparing the hardness of soil with its clay content) give a β_1 estimate of -2.1. Given this information, what is the *largest* value of (the estimate of) the standard deviation of the estimate of the β_1 parameter such that you could say, with 95% confidence, that the true value of the coefficient *is not* equal to zero?

17. In a regression analysis of the structural strength of a beam, the following data has been tabulated, with *X* representing the length of the beam and *Y* representing the amount of force (in MN) that the beam could withstand before buckling. Given this information, what is the sum of squared errors (SSE) of the analysis?

$\sum_i X_i = 254{,}081$ $\sum_i Y_i = 80{,}902$

$\sum_i X_i Y_i = 7{,}870{,}787$ $\sum_i X_i^2 = 24{,}056{,}600$

$\sum_i Y_i^2 = 2{,}608{,}113$

no. observations = 2,985

18. A regression analysis has been performed comparing the average number of cars per hour on one mile stretches of highway (*X*) and the number of crashes per year on those same stretches (*Y*). Given the data below, which has been tabulated from the analysis, what is the variance of the error term in the least squares regression equation?

$\sum_i X_i = 254{,}061$ $\sum_i Y_i = 80{,}920$

$\sum_i X_i Y_i = 7{,}870{,}557$ $\sum_i X_i^2 = 24{,}056{,}905$

$\sum_i Y_i^2 = 2{,}608{,}403$

no. observations = 2,956

19. A firm has developed a regression analysis linking the average rainfall (X, in cm) in an area and how often, on average, roofs have to be replaced (Y, in years). The firm wishes to rate the performance of the analysis. Given the data below, which has been tabulated from the collected data, what is the value of the R^2 measure from this analysis?

$\sum_i X_i = 254{,}064$ $\sum_i Y_i = 80{,}936$

$\sum_i X_i Y_i = 7{,}870{,}000$ $\sum_i X_i^2 = 24{,}056{,}521$

$\sum_i Y_i^2 = 2{,}608{,}322$

no. observations = 3,098

20. A regression analysis is performed comparing the height of a reservoir (X, in meters) with the net water usage of the local community the previous year (Y, in billions of gallons). The firm performing the analysis wants to be sure the analysis is valid, so they will perform a t-test on the β_1 coefficient. In order to do this, they need an estimate of the standard deviation of that coefficient. Given the information provided below, which has been tabulated from the collected data, what is the least squares estimate for the *standard deviation* of the β_1 parameter?

$\sum_i X_i = 254{,}095$ $\sum_i Y_i = 80{,}942$

$\sum_i X_i Y_i = 7{,}870{,}186$ $\sum_i X_i^2 = 24{,}056{,}303$

$\sum_i Y_i^2 = 2{,}608{,}438$

no. observations = 2,944

1)	Zero	2)	0.552	3)	26.2
4)	-19.8	5)	4.14	6)	-0.332
7)	0.166	8)	16.2	9)	1,353
10)	Yes	11)	-5.16	12)	19,972
13)	496,784	14)	31.0	15)	0.364
16)	1.28	17)	16,502	18)	5.32
19)	0.955	20)	0.00156		