1) Given the data set 10, 5, 2, 6, 3, 7, 20, the median value is 5.
   a) True
   b) False

2) In simple regression, at least one of the observed data points must be on the fitted regression line.
   a) True
   b) False

3) The residual is the difference between the observed value and the predicted value
   a) True
   b) False

4) Suppose that X is a uniform random variable on the interval (-1,1). Then Var(X)=1/3.
   a) True
   b) False

5) On an exam, Tom scored 8 points above the mean and had a z-score of +2.00. The standard deviation for the set of exam scores must be equal to 4.
   a) True
   b) False

6) If A and B are two events, then P(A)+P(B)=1.
   a) True
   b) False

7) The standard deviation of a data set is 4. If each measurement is multiplied by -2, the standard deviation of the resulting data set is -8.
   a) True
   b) False

8) Suppose $X \sim Bin(20,.8)$. Then $E(X^2) = 3.2$
   a) True
   b) False

   This can be seen using Var(X)=E(X^2)-E(X)^2

9) In general, which of the following statements are FALSE?

   a) The sample mean is more sensitive to extreme values than the median.
   b) The sample standard deviation is a measure of spread around the sample mean.
   c) If a distribution is symmetric, then the mean will be equal to the median.
   d) The sample standard deviation is a measure of central tendency around the median.

10) A new medicine has an 85% success rate. Twenty patients are treated with it. What is the mean and the standard deviation of this distribution?

For binomial mean=np and variance = np(1-p)

   a) mean = 17, standard deviation = 2.55
   b) mean = 17, standard deviation = 1.597
   c) mean = 3, standard deviation = 1.597
   d) mean = 3, standard deviation = 2.55
   e) none of the above

11) If the variance of random variable X is greater than the variance of random variable Y, then:

   a) the mean of X must be greater than the mean of Y.
   b) the median of X must be equal to the median of Y.
   c) the probability that X is equal to Y is zero.
   d) the standard deviation of X is greater than the standard deviation of Y.
   e) all of the above are true.

12) One out of three mini-vans sold by a nationwide auto dealer has a hidden defect in its transmission. What is the probability that a randomly selected purchaser of two mini-vans will wind up with at least one mini-van with a defective transmission?

Binomial-desired probability = 1-none defective = 1-(2/3)^2

a) 0.333
b) 0.500
c) 0.667
d) 0.250
e) 0.556

13) Assume that weights of school children are normally distributed with a mean of 80 pounds and a standard deviation of 20 pounds. Given that a student weighs more than 75 pounds, what is the probability that the student weighs more than 81 pounds?

We want P(X>81|X>75) = P(X>81 and X>75)/P(X>75)=P(X>81)/P(X>75)

a) 0.5
b) 0.2
c) 0.7
d) 0.8
e) 0.6

14) Suppose X is a uniformly distributed random variable on the interval 0 to 1. What is $P(X^2 < 0.25)$?

$$P(X^2 < 0.25) = P(X < \sqrt{0.25}) = P(X < 0.5) = 0.5$$

a) 0.25
b) 0.75
c) 0.0625
d) 0.5
e) 0.33

15) Suppose there are 10 measurements, half of which take negative values and the other half take positive values. Then which of the following statements must be true about this data (choose the most appropriate one)?

    a) variance is zero
    b) median is zero
    c) mean is zero
    d) standard deviation is positive
    e) both b and d

16) Suppose that the amount of time that it takes a clerk to process an employment application is uniformly distributed between 5 minutes and 12 minutes. What is the probability that the clerk will take more than 7 minutes to process a randomly selected application?

    a) 3/7
    b) 4/7
    c) 5/7
    d) 6/7
    e) cannot be determined

17) The sign of the correlation r between two variables X and Y (is)

    a) always positive
    b) always negative
    c) depends on the sign of the standard deviation of X
    d) depends on the sign of the standard deviation of Y
    e) depends on the sign of the covariance

18) A discrete random variable X denotes the number of times a month (0,1,2,3,…) that a machine fails, and $P(X \geq 1) = 0.8$ and $P(X \leq 1) = 0.3$. Which one of the following must be true?

We know P(X=0)+P(X=1)=0.3 and P(X=0)=1-0.8 = 0.2 so P(X=1) = 0.1.

    a) $P(X = 1) = 0.2$
    b) $P(X > 3) < 0.6$
    c) $P(X = 1) = 0.1$
    d) $P(X \leq 0) = 0.5$

19) A discrete random variable X takes values 1, 2 and 3 with P(X=2)=0.4 and E(X)=2. What is P(X=3)?

We know 1P(X=1)+2(0.4)+3P(X=3) = 2 and P(X=1)+P(X=3)=1-0.4=0.6.
Replace P(X=1) with 0.6-P(X=3) and we have 0.6-P(X=3)+3P(X=3)=1.2,
Or 2P(X=3)=0.6 or P(X=3)=0.3.

    a) 0.1
    b) 0.2
    c) 0.3
    d) 0.4
    e) 0.5

20) You roll a far six-sided die. If it comes up a 6 you win $100. If not, you get to roll again. If you get a 6 the second time, you win $50. Otherwise you get nothing. What is the expected value for the amount you win?

E(X)=100(1/6)+50(5/6)(1/6)+0

    a) 27.4
    b) 20.7
    c) 33.7
    d) 23.6
    e) 25.4

21) It rains in Cambridge 30% of the time in October. If it is raining, then Tommy will wear his hat with probability 0.8. If it is not raining, then Tommy will wear his hat with probability 0.2. If Tommy is seen wearing his hat, what is the probability that it is raining?

2x2 table. P(R)=0.3,P(H|R)=0.8,P(H|Rbar)=0.2 want P(R|H).
P(H and R)=0.8(0.3)=0.24 and P(H and Rbar)=(0.2)(0.7)=0.14.
P(H)=P(H and R)+P(H and Rbar) = 0.38 so P(R|H)= 0.24/0.38=0.63.

   a) 0.72
   b) 0.84
   c) 0.22
   d) 0.63
   e) 0.34

22) The heights of WNBA players are normally distributed with mean 73 inches, variance 4. Of which number x (rounded to one decimal place) is it true to say that 90% of all WNBA players are at least x inches tall?

Find x so that P(X<x) = 0.9.
> qnorm(.9,73,2)
[1] 75.5631

   a) 76.3
   b) 75.6
   c) 67.8
   d) 69.7
   e) 70.4

23) Two Physics classes at Jefferson High School took the same quiz. Mr. Spears had 20 students in his class with a mean score of 80. Mrs. Guyton's class of 30 students had a mean score of 90. Overall, what was the mean score for all students on the quiz?

   a) 85
   b) 84
   c) 86
   d) 87

24) What does the shape of a binomial distribution with n=10 and p=0.20 look like?

    a) Skewed right
    b) Skewed left
    c) Uniform
    d) Symmetric
    e) Normal

25) Which of the following statements is true:

a can be false-data set = {-1,-2} for example
b can be false-data set = {1,1,1} for example
c can be false-data set = {1,1,1} for example

    a) the mean of a sample is always greater than the std. dev. of the sample.
    b) the std. dev. for a data set with 3 observations must always be positive.
    c) the Q3 of a data set must always be greater than the Q1 of the data set.
    d) the mean of a data set can be negative but the std. dev. of a data set can never be negative

26) The five-number summary of credit hours for 24 students in an introductory statistics class is:

Min	Q_1	Median	Q_3	Max
13	15	16.5	18	22

Using the boxplot rule for outlier detection we know that

    a) there are no outliers in the data.
    b) there is at least one low outlier in the data.
    c) there is at least one high outlier in the data.
    d) None of the above.

27) Suppose A and B are two events such that P(A)=0.4, P(A or B) = 0.7 and P(B)=p. What value does p have to be for A and B to be independent?

P(A or B) = P(A) +P(B)-P(A and B) so 0.7 = 0.4+p-P(A and B).
We want A and B to be independent so P(A and B) = P(A)P(B)=0.4p
So we have 0.7 = 0.4+p-0.4p or 0.3 = p(1-.4) or p = .3/.6 = 0.5

    a) 0.2
    b) 0.7
    c) 0.8
    d) 0.5
    e) 0.6

28) Suppose the mean age of lead actresses from the top 50 grossing movies of 2016 was 29.6 years with a standard deviation of 6.35 years. Assume also the dotplot of the actresses' ages is mound shaped. Between what two values would you expect to find about 95% of the lead actresses ages?

Empirical rule: xbar +/-2s = 29.6+/-2(6.35)

    a) 16.9 and 42.3 years
    b) 23.25 and 35.95 years
    c) 10.55 and 48.65 years
    d) None of these

29) A garage mechanic keeps a box of good springs to use as replacements on customer's cars. The box contains 5 springs. A colleague, thinking that the springs are for scrap, tosses three faulty springs into the box. The mechanic picks two springs out of the box while servicing a car. Find the probability that the second spring drawn is faulty. [hint define A=first spring is faulty and B = second spring is faulty]

Build a 2x2 table but we'll work it out without building a table. Using the hint, P(A)=3/8 and P(B|A) = 2/7 and P(B|Abar) = 3/7.
Then P(A and B) = (3/8)(2/7) = 6/56 and P(Abar and B) = (3/7)(5/8) = 15/56.
Then P(B) = (6/56)+(15/56) = 21/56 = 3/8.

    a) 1/8
    b) 2/8
    c) 3/8
    d) 4/8
    e) 5/8

30) Suppose linear regression was used to model annual professor salaries (in dollars) based on how many years they have worked. The resulting equation is Salary = 95000 +1280 (Years). Interpret the slope in the context of the data.

a) The slope is 95000. For every additional year a professor works at a college, his/her salary is predicted to increase by $95,000

b) The slope is 95000. If a professor has never worked at a college, his/her salary is expected to be $95,000.

c) The slope is 1280. If a professor has never worked at a college, his/her salary is expected to be $1,280.

d) The slope is 1280. For every additional year a professor works at a college, his/her salary is predicted to increase by $1,280.

31) Below is the standard deviation for 10k finish times for a randomly selected group of women and men. Choose the statement that best summarizes the meaning of the standard deviation. Women: s = 0.16 Men: s = 0.25.

a) On average, women's finish times will be 0.16 hours less than men's finish times.

b) On average, men's finish times will be 0.25 hours faster than the overall average finish time.

c) The distribution of women's finish times is less varied than the distribution of men's finish times.

d) The distribution of men's finish times is less varied than the distribution of women's finish times.

32) Researchers conducting a study to investigate the relationship between a mother's age in years (x) and her baby's birth weight in pounds (y) found that the relationship can be modelled by the following linear regression line yhat = 3.23+0.15x. Based on this equation what would be the residual for a 28 year old mother who gave birth to a baby weighing 8 pounds?

a) -1.14
b) -0.57
c) 0.57
d) 1.14
e) 1.34

33) The following is a payoff table giving profits for various situations. What is the maximin decision?

Alternatives	States of Nature		
	A	B	C
Alternative 1	120	140	120
Alternative 2	200	100	50
Alternative 3	100	120	180
Do Nothing	0	0	0

    a) **Alternative 1**
    b) Alternative 2
    c) Alternative 3
    d) Do Nothing

34) Suppose the length of a particular species of fish is normally distributed with a mean length of 15 inches and a standard deviation of 2 inches. If a fish is more than 15.5 inches long it is considered a legal catch. If you randomly catch 4 of these type of fish one afternoon, what is the probability that all of them are longer than 15.5 inches? [round to the closest answer]

    a) 0.03
    b) 0.05
    c) 0.07
    d) 0.09
    e) 0.11

35) Let $x_1, x_2, \cdots, x_n$ be n observations. Let $w_i = a + bx_i$ for i = 1, 2, ...n, where a and b are constants. If the mean of the x_i's is 48 and their standard deviation is 12, the mean of w_i's is 55 and standard deviation of the w_i's is 15, the values of a and b should be

This is the a+bX rule written in a slightly different way.
We know 55=a+ b(48) and 225 = b*b*144 [using the variances].
Then b = sqrt(225/144) = 1.25 so a = 55-1.25(48) = -5.

    a) **b = 1.25, a = − 5**
    b) b = − 1.25, a = 5
    c) b = 2.5, a = − 5
    d) b = 2.5, a = 5
    e) Not enough information provided

36) Suppose, using 36 months of monthly returns, that the stock of Company A, a tech startup, has an average monthly return of .02, with a standard deviation of .06. The stock of Company B, a tire manufacturer, has an average monthly return of .01, with a standard deviation of .02. The covariance between stocks A and B is .000675. What is the correlation between stocks A and B?

Use the relationship cor = cov/(s_x*s_y)

    a) 0.48
    b) 0.63
    c) 0.27
    d) 0.79
    e) 0.56

Answers:

1) B
2) B
3) A
4) A
5) A
6) B
7) B
8) B
9) D
10) B
11) D
12) E
13) D
14) D
15) D
16) C
17) E
18) C
19) C
20) D
21) D
22) B
23) C
24) A
25) D
26) A
27) D
28) A
29) C
30) D
31) C
32) C
33) A
34) A
35) A
36) E